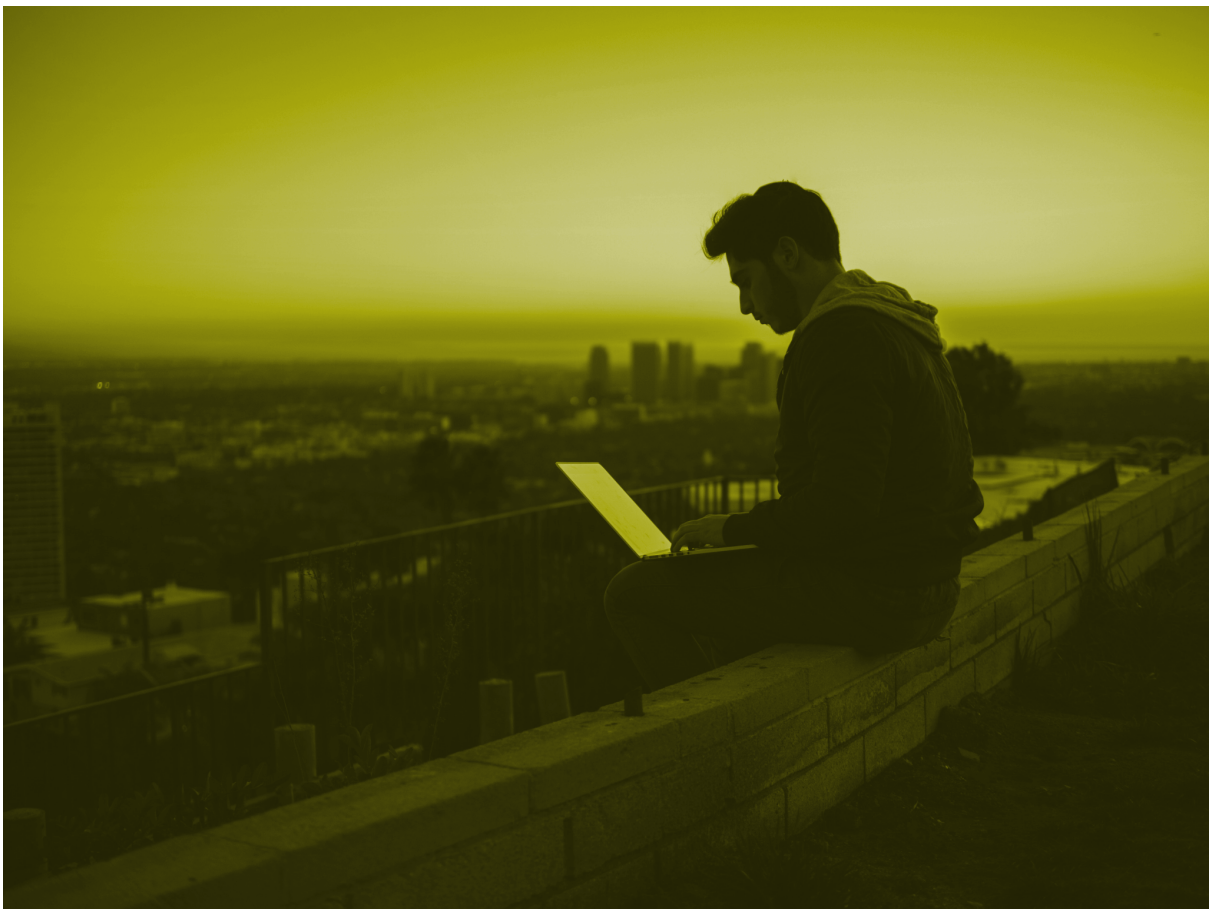


29 August 2023

Submission on “Safer Online Services and Media Platforms”

Department of Internal Affairs

internetnz

Table of contents

Introduction

Who we are and what we stand for

InternetNZ | Ipurangi Aotearoa operates the .nz domain space. We ensure all domain names ending with .nz are available for people and businesses in Aotearoa to function and thrive online.

We're a not-for-profit organisation, and the money we receive from .nz domain names goes back into the community. We provide grants, help to fund other organisations, and advocate for an accessible and safe Internet that benefits everyone in Aotearoa.

We welcome this opportunity to submit our feedback on the Safer Online Services and Media Platforms (SOSMP) discussion document.

We value our relationship with The Department of Internal Affairs Te Tari Taiwhenua and wish to continue general engagement, as well as engagement on this proposal and the specific issues we comment on in this submission.

Please continue to contact us via email at policy@internetnz.net.nz with opportunities for engagement.

Summary of submission

Te Tiriti o Waitangi is this nation's founding document. The partnership agreements contained within Te Tiriti must be the initial consideration by every organisation, particularly Crown agencies. InternetNZ | Ipurangi Aotearoa believe that equality starts with equity, which is one of the motivations behind our commitment to becoming a Te Tiriti o Waitangi centric organisation. We thus support the document's emphasis on upholding human rights and freedom of expression. However, as Te Tiriti partners, Māori need to be approached first with opportunities to achieve that equity, and be equal partners in the development and oversight of this process.

We recognise that many people are adversely affected by the content experienced on media platforms, and that harmful content disproportionately targets and affects some communities more than others. To help us prepare to write this document, we engaged with representatives from a variety of organisations and communities who confirmed the disproportionate harms regularly experienced by Māori, and also marginalised and disenfranchised groups. While we do not speak for these groups, their knowledge, experiences, and insights inform the points in the document below.

One of the most frequent issues we heard during our engagement on the SOSMP is that the voices of those most affected by online harms are often left out of conversations about how to regulate them. Many highlighted the lack of

consultation with these communities during the development of the Aotearoa New Zealand Code of Practice for Online Safety and Harms. The SOSMP, and the entity that emerges from it, is a chance to do better for Aotearoa.

In addition, Māori, and other groups disproportionately affected by harmful content and attempts to regulate that content, need to be involved in the development of the Independent Regulator so that their needs are met and their rights protected. The protection of human rights, freedom of expression, and freedom of the press also needs to be a priority. If done well, a strong regulatory framework can protect, rather than limit, freedom of expression by helping more people feel safe and free to participate.

Finally, this regulatory framework needs to not just be reactive, but also proactive, providing information, tools, and resources to better equip everyone to deal with, report, and minimise online harms. This new entity will require significant investment in order to keep up with the rapid development of technologies, systems, and tools, as evidenced by the relatively quick emergence and prominence of artificial intelligence (AI) and large language models (LLM) since the beginning of 2023.

In our submission below, we make several recommendations we hope can help achieve these goals. Some of our primary recommendations include:

- Making the proposed Independent Regulator one element in a new Aotearoa Media Commission, alongside two other bodies: an independent Recourse Council and an independent Advisory Board. The Recourse Council would be an entity to which individuals and platforms alike could appeal decisions made by the Independent Regulator. The independent Advisory Board should include Māori as Treaty partners; Māori and other representatives of communities most affected by harmful content; and technology, legal, and subject matter experts. The role of these members would be to advise and support the other two bodies.

- Having sector-specific “Commissioners” for different media within the Independent Regulator, including one each for news media and journalism, traditional or linear media, and online platforms.

- Taking a somewhat prescriptive approach to the development and enforcement of industry codes of practice, with the ability to institute industry standards when industry representatives cannot or will not develop satisfactory codes of practice.

- Making the Independent Regulator, and the proposed Advisory Board and Recourse Council, as well as the media platforms, subject to rigorous transparency and oversight requirements.

- Prioritising Māori and other community-led educational initiatives, and providing monetary support and other resources to them, to develop Aotearoa-specific educational content on how to report and minimise harmful content on media platforms.

We offer these recommendations as points of discussion, and would welcome the opportunity to explore them further with you. You can contact us at policy@internetnz.nz.

Summary of recommendations

R1 **Any development of regulation of media (both online and traditional) must be co-designed and co-governed with Māori.**

R2 **Further engagement with Māori should be undertaken, with a particular focus on wāhine Māori who have been identified as targets of harmful content.**

R3 **The Independent Regulator should initially consist of three "Commissioners", one each for traditional or linear media, news media and journalism, and online platforms.**

R4 **Government should establish an Advisory Board made up of Māori as Treaty partners, Māori and other representatives of communities most targeted and affected by harmful content, and technology, legal, and subject matter experts to advise and inform the Independent Regulator.**

R5 **Establish a Recourse Council to examine and respond to complaints about both platforms' and the Independent Regulator's decisions regarding content, with advice and support from the proposed Advisory Board.**

R6 **Maintain the current Office of Film and Literature Classification (Classification Office) as a separate and independent entity to investigate reported content and make rulings on 'objectionable' content.**

R7 **Review and revise the structure of the proposed Aotearoa Media Commission annually for the first three years of its existence, then every three years thereafter so that it can remain responsive to emerging technologies, issues, and threats; ensure it does not limit human rights, freedom of expression, or freedom of the press; and protects and supports everyone in Aotearoa, especially those from the communities most targeted and affected by online harms.**

R8 **Ensure that the Independent Regulator has the authority to develop a system that provides independent auditors and researchers with access to de-identified data to enable independent assessment of the results of algorithmic controls and platforms' adherence to industry codes, and to support general research that contributes to public knowledge.**

R9 **All entities within our proposed Aotearoa Media Commissioner model should, like the platforms it regulates, be subject to rigorous oversight and be required to publish regular transparency reports.**

R10 **The new framework should give the Independent Regulator the authority to institute industry standards, in cases where co-development of industry codes proves ineffective, subject to regular review and oversight.**

R11 **Revise the definition of "harmful content" so that it is clear that an evidence-based evaluation of risk of harm, rather than the experience of harm, is central to the definition or classification of harmful content.**

R12 **Require the Independent Regulator to seek out and incorporate information about context-dependent risks of harm as a part of its decision-making processes.**

R13 **The Independent Regulator should be provided with budget and other resources necessary to develop informational resources to help smaller platforms address harmful content and, if necessary, assists smaller platforms with monitoring for and removing harmful and objectionable content.**

R14 **Amend the definition of "Regulated Platform" so that there are no exclusions for charities, clubs, retailers, or professional services.**

R15 **The Independent Regulator should have takedown powers for material that is illegal under other New Zealand regimes, but only material that can result in harm to physical, social, emotional, and mental wellbeing.**

R16 **Develop a process to channel complaints handled by other agencies to those organisations to maintain a simplified complaints process.**

R17 **Prioritise Māori and other community-led educational initiatives (and provide monetary support and other resources to them) to develop Aotearoa-specific educational content on how to report and minimise harmful content on media platforms.**

R18 **Any voluntary online filters should incorporate the use of interstitials that provide links with education and support information to help minimise the effects of harmful and objectionable content.**

Discussion of key themes and issues

Human rights and freedom of expression

1. A central point of discussion in the SOSMP discussion document, and a significant concern of those with whom we consulted in preparing this document, is how the proposed framework will either protect or infringe upon human rights and freedom of expression.

2. We strongly believe upholding human rights and freedom of expression should be one the most important outcomes of this process.

3. We were therefore happy to read that the discussion document clearly states that "**New Zealand is committed to freedom of expression**, so it is important that New Zealanders can continue to create and share content and access the content and services they value" (p. 77, para 1).

    a. The UN Declaration of the Rights of Indigenous Peoples also notes that "control by indigenous peoples over developments affecting them and their lands, territories and resources will enable them to maintain and strengthen their institutions, cultures and traditions, and to promote their development in accordance with their aspirations and needs".[1]

    b. In addition, Article 37 argues that indigenous peoples "have the right to the recognition, observance and enforcement of treaties, agreements and other constructive arrangements concluded with States or their successors and to have States honour and respect such treaties, agreements and other constructive arrangements".[2]

    c. Having Māori as partners in this process would be both in line with that objective and Crown obligations under Te Tiriti.

4. We also are strongly supportive of the fact that DIA is "**not proposing any changes to the types of material that are currently considered illegal in New Zealand**" (p. 52)

    a. The goal of any platform or content regulation, we believe, should be to keep the definition of illegal or objectionable content as narrow as possible so that freedom of expression is protected.

---

[1] See *United Nations Declaration On The Rights Of Indigenous Peoples*, p. 4, available at https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf.

[2] See *United Nations Declaration On The Rights Of Indigenous Peoples*, pp. 25-26, para. 37, available at https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf.

5. However, we also recognise that many people are adversely affected by the content they experience on media platforms. In addition, we fully recognise, and were reminded frequently during our discussions in advance of writing this submission, that **harmful content disproportionately targets and affects some communities more than others**. These communities include Māori, women, ethnic and religious minorities, the LGBTTQIA+ (Lesbian, Gay, Bisexual, Transgender, Takatāpui, Queer, Intersex, Asexual +) community, and people with disabilities, among others.

6. The experiences of these most affected communities is why we also support the stated goal that the SOSMP seek to ensure that "freedom of expression is **balanced with other human rights** such as non-discrimination, security, and democratic rights, and that all people, from consumers to creators and publishers, have equitable opportunities and do not suffer unfair treatment" (p. 85, para. 32).

7. In short, regulation, if done well, can work to ensure freedom of expression and limit harm.

8. The remainder of our comments below are made with the goal of establishing that balance in mind.

Partnership with Māori

InternetNZ | Ipurangi Aotearoa recommends that the DIA speak directly to Māori further about this submission.

In our preparation for this submission we have spoken with our Māori members, partners, and stakeholders. Engaging with Māori voices and bringing these voices and views into the InternetNZ | Ipurangi Aotearoa submission is important to us in supporting Article Two of Te Tiriti o Waitangi, Tino Rangatiratanga, and Article Three Oritetanga. The following feedback concerning Māori issues is a result of our engagements with Māori.

Aspects InternetNZ | Ipurangi Aotearoa supports

9. InternetNZ | Ipurangi Aotearoa supports the acknowledgement of specific and targeted harm against Māori, in particular wāhine Māori.

    a. There is strong evidence of a notable increase in this kind of harm, and Māori researchers and activists believe this harm will only continue to grow.[3]

[3]More information on the increase of racism is avaliable at https://www.nzherald.co.nz/kahu/experts-say-increase-in-online-racism-towards-maori-especially-wahine-maori-is-concerning/X6MX2O23EG4GM2U5BLCK7ZQONY/.

b. Therefore, it is important that any regulation of harmful content takes into account the needs and whakaaro of tāngata whenua.

c. It is also important that all development in this space centres the Government's obligations under Te Tiriti o Waitangi. We note that these obligations were spoken to in the discussion document.

10. We also support the inclusion of Māori representation on the Board of the Independent Regulator. (Suggestions from our stakeholders for what this could look like are outlined below.)

Whakaaro InternetNZ | Ipurangi Aotearoa has received

11. The document states: "There is a role for Māori in the governance and decision-making of the regulator, as well as in developing the codes and delivering education." (p. 24, para. 29)

    a. The document also states that "The regulator will be responsible for meeting the Government's obligations under Te Tiriti" and that this obligation "would be built into the code development and approval process, including ensuring Māori participation and that codes reflect Māori social and cultural values." (p. 39, para. 58)

    b. The processes and mechanisms needed to meet these goals are unclear and must be clarified. How does ensuring codes reflect Māori social and cultural values equal equitable partnership? Can the codes reflecting Māori values be enough to fulfil the Te Tiriti obligation?

12. Many Te Tiriti o Waitangi obligations need to be met in this process.

    a. Equitable partnership with Māori is obligated as per article one, of Te Tiriti o Waitangi.

13. DIA acknowledges that Māori have a specific and vested interest in the issues this proposal seeks to address. A way to evolve this acknowledgement into action would be for Māori to have joint or equal governance. InternetNZ | Ipurangi Aotearoa requests that DIA embed co-governance into the creation and statute of any legislation.

14. InternetNZ | Ipurangi Aotearoa has heard from our Māori partners that the engagement with Māori up to this point may have some potential gaps.

    a. In accordance with the government's engagement framework the Crown must engage early, be inclusive, and think broadly. [4]

[4] Māori Crown relations capability framework for the public service: https://www.tearawhiti.govt.nz/assets/Tools-and-Resources/Whainga-Amorangi/TA013.02-MCR-capability-framework-guide.pdf

b. All aspects of legislation (but particularly legislation which directly affects Māori this way), should be co-designed with Māori.

c. Our Māori stakeholders have expressed concern that they are not aware of any co-design of this discussion document with Māori.

15. The Māori aspects of the education arm of the Independent Regulator is also an area that needs more consideration (see the section “Education and support” starting on p. 35).

    a. InternetNZ | Ipurangi Aotearoa is in support of having specific funds for Māori engagement with the education arm of the Independent Regulator, it is also vital that these kaupapa is set up alongside Māori, employing Māori methodologies.

    b. There are cases of government education programs not creating strong positive outcomes for Māori. A critical factor in programs succeeding is a base of positive relationships between Crown agencies and Māori so that Māori can clarify the needs of their communities. This must be the case with this program.

        i. To ensure this positive outcome, the education program should be co-designed with communities, co-governed by Māori, and delivery should be co-lead by Māori.

16. Māori representation within the regulator should be embedded into legislation.

    a. In our consultations, Māori groups in particular repeatedly recommended that the legislation should mandate a minimum of 50% Māori members on the Board.

    b. InternetNZ | Ipurangi Aotearoa supports the legislation of Māori co-governance of the Independent Regulator.

        i. This requirement is in accordance with the Crown's obligation to provide Māori with equal partnership and governance.

        ii. This advisory board should also include a Mātauranga Māori (Māori knowledge) expert to remove extra labour from other Māori staff with different roles.

        iii. In addition to Mātauranga Māori experts, efforts should be made to find other speciality Māori in all areas of the recommended Advisory Board on p. 9, such as Māori legal experts, Māori tech experts, etc.

    c. The Māori advisory structure must be well resourced, i.e., well funded, staffed, and supported.

    d. This structure should be permanent and not able to be dismantled.

e. This structure must also have a strong mandate to ensure the Māori concerns are heard and acted upon.

f. General Māori representation should be established throughout all aspects of the regulatory body, and equitable hiring practices must be in place to ensure this representation.

R1 Any development of regulation of media (both online and traditional) must be co-designed and co-governed with Māori.

17. As the discussion document states, wāhine Māori are uniquely harmed by content, particularly online content.

    a. This harm is contextualised within embedded and systemic colonial patriarchal practices.

    b. This harm is specific and ongoing. Wāhine Māori who are active in media spaces are regularly harmed.

    c. This harm has intergenerational implications and intergenerational transmission that will continue to impact wāhine Māori, whānau, hapū, and iwi for generations long after the harm has been perpetrated.

    d. Because of this, it is our recommendation that further targeted engagement is undertaken by DIA on this issue before legislation is drafted, with a particular focus on wāhine Māori who have been victims of harmful content through orchestrated and deliberate campaigns of hate.

R2 Further engagement with Māori should be undertaken, with a particular focus on wāhine Māori who have been identified as targets of harmful content.

Structure of the regulatory entity

18. We support the stated desire for an approach that addresses systemic harms, rather than focussing (solely) on individual 'pieces' of content, and gathers an evidence base to inform responses over time.

    a. This risk-based approach is in line with recent recommendations from the Democracy and Internet Governance Initiative at the Harvard Kennedy School.[5]

19. Depending on how it is structured, the establishment of a single Independent Regulator for both traditional and online media could be a beneficial approach.

[5] The full report, "Towards Digital Platforms and Public Purpose", is available at https://shorensteincenter.org/wp-content/uploads/2023/07/DIGI-Final-Report_July5_2023.pdf.

a. At an earlier point in this process, we stated that a unified framework could improve "ease of public understanding, upholding standards, and reporting of complaints".

b. We are, therefore, generally supportive of the intention to transition existing standards and codes, such as those in the Broadcasting Standards Act, to this new framework over time (p. 64, para. 130), including those that "focus on the practice of ethical journalism in matters such as intrusion into privacy and accuracy of information" (p. 78, para. 9).

    i. In some cases, however, these existing codes address individual pieces of content, which suggests some conflict with the claim that the Independent Regulator "would not be involved in moderating individual pieces of legal content" (p. 22, para. 26).

20. However, the Regulator needs to be flexible enough to respond to emerging technologies and harms, responsive to communities that are disproportionately affected by harmful content, and respectful of human rights, freedom of expression, and press freedom.

21. We therefore believe that the Independent Regulator requires sector-specific Commissioners within its structure, and needs to be balanced by separate entities that advise and oversee the Independent Regulator, and provide redress opportunities for both individuals and platforms affected by the Regulator's activities and decisions.

22. Below, we outline these elements which, in line with our previous comments, we envision as part of **a new Aotearoa Media Commission**.

Proposal for sector-specific Commissioners within the Independent Regulator

23. We also stated in our previous document that the variety of services and activities included in this unified framework may be "different enough to require varied approaches going beyond the scope of industry-specific codes, and understanding those requirements may require looking at features of each area".

    a. We used the example of news and reporting, "well served by specific standards for news and journalistic content backed by regulation" and recommended a "level of continuity, levelling up these well established standards with unified oversight as part of a broader system".

    b. To achieve this mix of sector specific standards and continuity with unified oversight, we suggest that the Independent Regulator consists of separate "Commissioners", one each for:

i. News media and journalism (including the subsumption of the duties of the Media Council);

ii. Traditional or linear media (e.g., television and radio, including the subsumption of the duties of the Broadcasting Standards Authority [BSA]) except for news media and journalism;

iii. Online platforms (e.g., social media platforms and websites) except for news media and journalism.

c. This model is similar to one included in Ireland's Online Safety and Media Regulation Act 2022, which outlined an integrated Media Commission with three commissioners, one each for linear media, online safety and video sharing platforms, and on-demand media.[6]

d. The advantage of this approach is that it could lead to the development of more relevant, context-specific, and actionable industry codes.

e. Having a separate Commissioner specifically for news media and journalism will also potentially help the Independent Regulator meet the stated goal to uphold the principle of freedom of the press (p. 19, para. 19).

i. As the discussion document notes, freedom of the press is "important in holding Government and those who exercise public power to account" and also "supports democracy by keeping the public informed on important issues" (p. 19, para. 19).

ii. Those important roles often require journalists and news organisations to have some latitude in the kind of content they share, as well as some extra considerations concerning the protections they are afforded, such as the need to keep some sources confidential.

iii. However, the ability for nearly anyone with an Internet connected device to create and share content has expanded and, arguably, complicated what can be classified as "journalistic content".

iv. A specific set of industry codes for news media and journalism, and a dedicated Commissioner to oversee those codes, would provide needed protections for journalists while also ensuring that only those who follow these industry codes are afforded journalistic advantages or exemptions.

[6] More information on the Irish regulatory framework can be found in section 3.4 of *International Regulatory Frameworks for Online Content*, available at https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/International-Regulatory-Frameworks-for-Online-Content-Report.pdf.

R3 The Independent Regulator should initially consist of three "Commissioners", one each for traditional or linear media, news media and journalism, and online platforms.

Proposal for an independent Advisory Board

24. The Independent Regulator would need to be flexible enough to respond rapidly to emerging harms and threats, and be subject to significant oversight and transparency requirements.

25. The discussion document mentions a Board that "oversees the regulator's activities and sets its strategic direction and priorities" (p. 67, para. 145). This Board would be appointed by the Government (p. 24, para. 29).

26. We propose that this Board not only guides the Regulator's strategic direction and priorities, but also acts as an advisory body. This Board would be tasked with staying up-to-date on emerging technologies, harms, and threats, and engaging with the needs of Māori and developments in the other communities most affected by online harms and regulatory frameworks intended to reduce those harms. These communities include, but are not limited to, women, ethnic and religious minorities, and the LGBTTQIA+ community.

27. To be effective, this Advisory Board should comprise Māori as Treaty partners, Māori and other representatives of communities most targeted and affected by harmful content and attempts to regulate it, and technology, legal, and subject matter experts. Members should be actively engaged in their fields of expertise or communities (or both), depending on their role.

    a. The "significant Māori presence on the Board of the regulator" mentioned in the discussion document (p. 67, para. 145) would provide an avenue for "a role for Māori in the governance and decision-making of the regulator" (p. 24, para. 29).

    b. All Advisory Board members should be compensated and resourced for their work.

28. The information and knowledge collected by this Board, including through proactive research, would be provided to the Independent Regulator to inform decisions about harmful content, inform the Regulator's educational and awareness efforts, connect the development of industry codes to international efforts and practice, and provide Aotearoa-specific, Te Ao Māori informed context to the Regulator's understanding of online harms.

29. In line with the discussion document's emphasis on transparency, the information and advice provided by this Board to the Independent Regulator should be made publicly available for review.

30. This open, multistakeholder, evidence-based approach to regulation is in line with UNESCO guidelines to developing regulatory frameworks that are consistent with international human rights standards.[7]

31. Moreover, this Advisory Board could (and should) be established early in the process and inform the development of the rest of the structure of the Independent Regulator, to better ensure it is responsive to people and communities in Aotearoa.

R4 Government should establish an Advisory Board made up of Māori as Treaty partners, Māori and other representatives of communities most targeted and affected by harmful content, and technology, legal, and subject matter experts to advise and inform the Independent Regulator.

Proposal for an independent Recourse Council

32. The discussion document notes: "Appeals from platforms' complaint processes would also go to a quasi-judicial body associated with or approved by the regulator. This would be incorporated into the codes that apply to those platforms" (p. 44, para. 71).

33. We support the establishment of a formal body to address complaints and disagreements about actions taken by platforms.

34. However, while the discussion document stipulates that complaints that cannot be resolved directly with a platform can be laid with the Regulator (p. 61, para. 124), we believe that these complaints should be resolved with a separate and independent body we refer to as the **Recourse Council**.

    a. We believe that this Recourse Council should also provide an added level of oversight by making it possible for both platforms and users by addressing appeals of decisions made by the Independent Regulator.

35. The ability to challenge both the Independent Regulator's and platforms' decisions can act as a safeguard to prevent misuse or weaponisation of the "simplified complaints processes" (p. 25), particularly against the most affected individuals and communities that this framework is meant to protect.

    a. This model is based on a similar proposal for a Digital Recourse Council included in the Canadian Government's proposed Online Harms Legislation.

        i. Some submissions did raise concerns with that proposal, however, including that the proposed Canadian Digital

[7] For more, see "Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms", available at https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en.

Recourse Council "would not be able to judge the variety of content it would be asked to evaluate."[8]

ii. We agree that expecting individual adjudicators to have the breadth of knowledge required is unreasonable.

iii. We therefore propose that the Advisory Board discussed above also provide information and advice to this Recourse Council so that it can make better-informed and contextual decisions.

b. This model is also in line with UNESCO recommendations for complaints processes, which can help align regulatory frameworks to international human rights standards.[9]

R5 Establish a Recourse Council to examine and respond to complaints about both platforms' and the Independent Regulator's decisions regarding content, with advice and support from the proposed Advisory Board.

Proposal to maintain the censorship function within the current Office of Film and Literature Classification

36. The discussion document outlines several options concerning how a censorship function that determines 'objectionable' or illegal content could operate (pp. 52–53, paras. 93–94).

a. That function is currently the remit of the Office of Film and Literature Classification led by the Chief Censor.

b. The three suggestions in the document regarding this function are to:

i. Have the Chief Executive of the Independent Regulator assume this function

ii. Develop a separate 'statutory officer' within the Independent Regulator to assume that role, or

iii. Establish a "tribunal or panel with legal expertise that can be rapidly convened" to make decisions.

c. Each of these options presents issues, however.

i. The first two options conflict with the stated intention that the Independent Regulator "would not be involved in moderating individual pieces of legal content" (p. 22, para. 26),

[8] A summary of feedback collected by Canadian Heritage in response to the government's online harms proposal is available at: https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/what-we-heard.html

[9] See "Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms", p. 16, para. 49g, available at https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en.

and does present the potential for too much authority to be ascribed to the Regulator.

ii. The third option introduces logistical uncertainty that a standing or permanent entity, i.e., one that does not need to be convened, mitigates.

d. Having the Independent Regulator assume the functions of the Classification Office would represent the potential for concentration of too much authority within the Independent Regulator.

i. Having this power remain outside of the Independent Regulator would introduce a system of checks and balances between the Classification Office and the Independent Regulator, which would potentially help ensure fairness and limit overreach.

37. The current system, in which the Chief Censor or delegate can make rapid, interim decisions in emergency situations, and permanent classifications after deliberation, operates sufficiently well in the current environment and provides flexibility and oversight.

38. Finally, the Classification Office has proven in the past to be judicious and restrained when classifying content as objectionable, doing so only in cases where content represented a clear danger. This experience is needed to protect freedom of expression.

39. We therefore propose that the three options presented for the censorship role be set aside in favour of maintaining the current Office of Film and Literature Classification as a separate entity outside of the Independent Regulator.

a. The Classification Office has established procedures in place to review and classify potentially objectionable content, and has historically only classified content as such in limited and extraordinary circumstances.

b. However, If that option is not desired, we would prefer to see the censorship role assumed by a separate, independent statutory officer within the Independent Regulator.

R6 Maintain the current Office of Film and Literature Classification (Classification Office) as a separate and independent entity to investigate reported content and make rulings on 'objectionable' content.

Overview of the proposed structure for an Aotearoa Media Commission

40. As stated above, we envision the Independent Regulator a primary element of an **Aotearoa Media Commission**, alongside the proposed Advisory Board and Recourse Council.

41. Among the other duties outlined in the discussion document, such as overseeing code development and developing educational and awareness programmes, the Independent Regulator would be responsible for coordinating and communicating with other related entities such as the Classification Office.

42. A diagram of our vision for the Aotearoa Media Commission is below on page 19.

Structural review

43. Because online harms are a contentious topic and are constantly evolving and emerging, we anticipate that this or any proposed structure will require rapid, regular, and ongoing review and adjustment.

44. We recommend that this structure be internally reviewed annually for the first three years of its existence, and then every three years thereafter.

    a. Our preference would be for each review to be led by the Advisory Board outlined above, with consultation and input from community groups and civil society, academics, legal advisors, technology experts and industry.

45. A summary of collected input, any changes resulting from the review, and the anticipated outcomes of those changes should be made publicly available before implementation. (See more on oversight and transparency in the section below.)

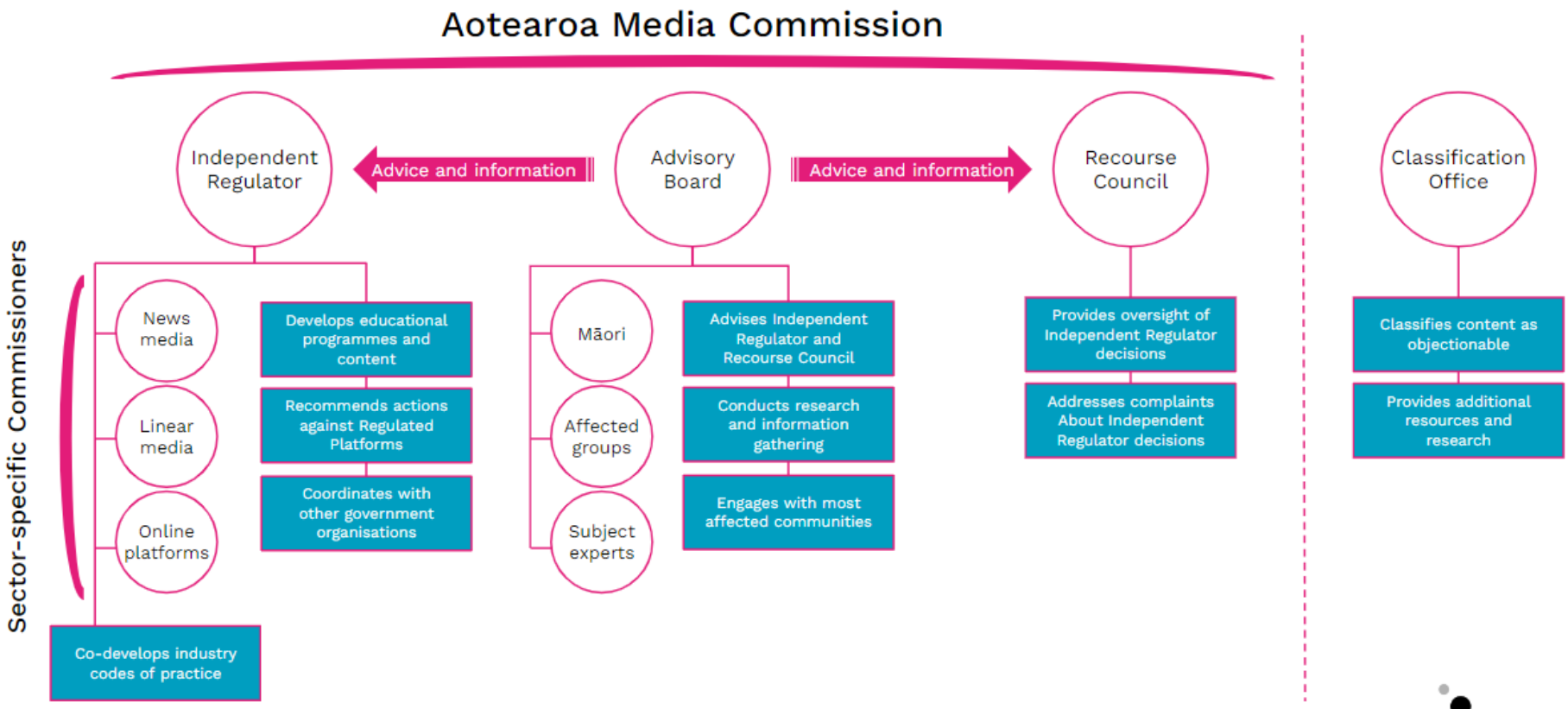
R7 Review and revise the structure of the proposed Aotearoa Media Commission annually for the first three years of its existence, then every three years thereafter so that it can remain responsive to emerging technologies, issues, and threats; ensure it does not limit human rights, freedom of expression, or freedom of the press; and protects and supports everyone in Aotearoa, especially those from the communities most targeted and affected by online harms.

Oversight and transparency

46. The emphasis on platform transparency and reporting throughout the document is encouraging.

    a. For example, periodic transparency reporting that includes information on how the impact of harm from content is reduced and how platforms perform against industry codes (p. 36, para. 53) is a welcome step to accountability, proportionality, and efficacy.

47. One element that requires more clarity is the expectation that Regulated Platforms "use transparent and regular reporting on their algorithmic controls and settings" (p. 36, para. 54).


Aotearoa Media Commission
Independent Regulator
Advice and information
Advisory Board
Advice and information
Recourse Council
Classification Office
Sector-specific Commissioners
News media
Linear media
Online platforms
Develops educational programmes and content
Recommends actions against Regulated Platforms
Coordinates with other government organisations
Co-develops industry codes of practice
Māori
Affected groups
Subject experts
Advises Independent Regulator and Recourse Council
Conducts research and information gathering
Engages with most affected communities
Provides oversight of Independent Regulator decisions
Addresses complaints About Independent Regulator decisions
Classifies content as objectionable
Provides additional resources and research

internetnz

48. We strongly agree that there are significant concerns about “the transparency and performance of the algorithms and other systems that social media companies use to direct content to users” (p. 80, para. 15).

49. However, there are significant barriers to achieving transparency of algorithmic systems.

    a. For commercial platforms, algorithmic processes are proprietary and often carefully guarded.

    b. In addition, algorithms are complex, requiring specialist knowledge to understand and analyse. This complexity raises the question of who will be able to adequately and independently assess algorithmic controls and settings included in transparency reports.

50. To assist with this process, data on the outcomes of algorithmic controls and amplification, and related data on human and algorithmic or AI interactions on platforms should be made available to independent researchers or assessors, but in a way that protects both the privacy of users and platforms’ proprietary information.

    a. For example, through our involvement in the Christchurch Call Advisory Network, we understand the company OpenMined, which is an open-source software community, has developed a product that allows platforms to upload de-identified data of this kind to a server, against which an independent researcher can run queries.

    b. These queries can help demonstrate the lengths to which platforms are adjusting their algorithms to address harmful content without accessing underlying (e.g., sensitive proprietary) data.

    c. Some platforms are testing this system to minimise terrorist and violent extremist content (TVEC) on their platforms in relation to the Christchurch Call.

    d. While we are not endorsing OpenMined, we do support the use of a similar process for assessing platforms’ adherence to industry codes and standards.

51. Finally, this data should be made available for general research that contributes to public understanding and enables greater participation and accountability.

R8 Ensure that the Independent Regulator has the authority to develop a system that provides independent auditors and researchers with access to de-identified data to enable independent assessment of the results of algorithmic controls and platforms' adherence to industry codes, and to support general research that contributes to public knowledge.

52. As stated above, we believe all elements of our proposed Aotearoa Media Commissioner should also be subject to oversight and rigorous transparency requirements to ensure the Independent Regulator is not infringing upon freedom of expression, freedom of the press, and human rights.

    a. We would like this framework to include a requirement that the Independent Regulator publicly make available an annual transparency report that includes information about:

        i. New codes developed, including the goals of these codes;

        ii. Amendments to code, including the reasons for the changes;

        iii. Retired or rejected codes, and the reasons for these decisions;

        iv. Actions taken or recommended to be taken against platforms that have violated the code, and referrals made to other entities, if any;

        v. The development, implementation, and efficacy of education and awareness-raising initiatives.

    b. Similarly, the Advisory Board should be required to provide a public annual report that summarises:

        i. Advice and information provided to the Independent Regulator;

        ii. Advice given to the Recourse Council;

        iii. Engagements with the public, civil society, communities, and other relevant stakeholders.

    c. Finally, the Recourse Council should provide a publicly accessible annual report that provides detailed information on:

        i. Appeals made by individuals and platforms, while ensuring that privacy and proprietary information are protected where necessary;

        ii. Decisions on those appeals and the reasons for those decisions.

53. Adding this level of transparency will ensure that all three entities—the Advisory Board, the Independent Regulator, and the Recourse Council—remain accountable both to the public and that their actions are

transparent to everyone affected by their actions, including the Regulated Platforms.

54. These transparency recommendations are also in line with UNESCO recommendations for government transparency designed to protect and respect human rights.[10]

R9 All entities within our proposed Aotearoa Media Commissioner model should, like the platforms it regulates, be subject to rigorous oversight and be required to publish regular transparency reports.

Industry codes of practice

55. We are concerned that co-development of codes of practices with industry is seemingly given priority in the SOSMP discussion document.

    a. The "Aotearoa New Zealand Code of Practice for Online Safety and Harms" demonstrates that industry-developed codes are insufficient.

        i. In our submission on what was then referred to as the Code of Practice for Online Safety and Harms, we argued that there were "foundational legitimacy gaps in the process, funding model, and operation of the proposed Code" (p. 5, para. 29).

        ii. Some of these concerns have been shared by the Government as well as other civil society organisations such as Tohatoha and Inclusive Aotearoa Collective Tāhono.

        iii. Of particular concern is the lack of engagement with the most affected communities and civil society groups both now and during the development of that Code.

        iv. Another common critique heard during our consultations is that the Oversight Board for the Code is not independent and does not include any representatives of the most affected communities.

56. However, the discussion document does include mentions of Te Ao Māori (p. 35) informing codes of practice, and a role for "community groups, and civil society to develop the codes" (p. 61, para. 122).

    a. At present, it is unclear how partnership with these groups in the development of industry codes would work.

    b. While the Independent Regulator would have the power to institute industry codes (or not), this decision would be better informed via a structure that formally includes Māori as Treaty partners, as well as

[10] See "Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms", p. 10, para. 29g, available at https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en.

Māori and other communities disproportionately affected by online harms and attempts to regulate them.

c. We believe the Advisory Board outlined above would be an important, if initial, step in formalising the involvement of communities disproportionately affected by online harms in the development of industry codes and be a way to ensure that Te Ao Māori is indeed integral to those codes.

57. In cases where industry groups are unable or unwilling to develop adequate codes in cooperation with the Independent Regulator (and, preferably, with input from the proposed Advisory Board), we would favour the Independent Regulator having the option of instituting an **industry standard** that Regulated Platforms would be required to follow.

    a. This suggestion is based on a similar function in Australia's *Online Safety Act 2021*, which empowers their eSafety Commissioner to establish such standards in cases where the Commissioner finds industry codes to be ineffective or insufficient.[11]

    b. This option should be used as sparingly as possible and only after every other option for cooperative development of industry codes is exhausted.

    c. In Aotearoa, we would like these industry standards to be developed in consultation with the Advisory Board so that they are context specific to Aotearoa.

    d. In addition, we believe that these industry standards should be reviewed every three months by the Advisory Board to ensure that they are fit for purpose and that they do not allow for overreach by the Independent Regulator.

    e. Finally, the Independent Regulator should continue to attempt to co-develop an industry code with industry representatives and the Advisory Board even after instituting an industry standard, and replace any imposed industry standards with these co-developed industry codes once they have been agreed to.

[11] More information on the Australian *Online Safety Act 2021* can be found in section 1 of *International Regulatory Frameworks for Online Content*, available at https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/International-Regulatory-Frameworks-for-Online-Content-Report.pdf.

R10 The new framework should give the Independent Regulator the authority to institute industry standards, in cases where co-development of industry codes proves ineffective, subject to regular review and oversight.

Supportive vs prescriptive approach

58. As this recommendation suggests, we are more in favour of a somewhat more **prescriptive approach** in which the Independent Regulator is more directive to industry, particularly when it comes to the development of industry codes, over a supportive approach, in which the industry would have more latitude to self-regulate (see p. 33, p. 49).

59. We also believe the Independent Regulator should have the power to issue **takedown notices for illegal or objectionable content**.

    a. Having the Regulator in charge of takedown notices for all Regulated Platforms would ensure consistency and a single point of contact for platforms and industry bodies seeking clarification.

    b. These takedown notices could then be appealed through the Governance Council in cases where platforms or individuals believe their content was taken down in error.

    c. However, as outlined above, we believe the power to decide what content is objectionable should remain with the Classification Office and therefore separate from the Regulator.

    d. Finally, we believe that the Independent Regulator should have the power to **recommend criminal prosecutions against Regulated Platforms on the grounds of significant non-compliance with industry codes**, but that existing law enforcement agencies should carry out investigations of and prosecutions for illegal material.

60. More information can be found in the “Monitoring and enforcement” section below.

Key concepts and definitions

Harmful content

61. The objective of the SOSMP stated in the discussion document is to “enhance protection for New Zealanders by reducing their exposure to harmful content, regardless of delivery method” (p. 19).

    a. We support this goal, while acknowledging the difficulty in defining and identifying harmful content and reducing exposure to content in a way that does not violate other rights.

62. The discussion document defines content as harmful when “the experience of content causes loss or damage to rights, property, or physical, social, emotional, and mental wellbeing” (p. 18, para. 10).

    a. We appreciate that the document also clarified that “[b]eing harmed is **distinct from feeling offended** (although content that is harmful will often also cause offence)” (p. 18, para. 10), which is a distinction that is necessary to protect freedom of expression.

63. Our conversations with various groups revealed that there was concern that this definition puts the onus on victims of harmful content to demonstrate the loss or damage they suffered.

    a. There was also concern that the proposed definition overlooks that people can be harmed by harmful content online, even if they do not themselves consume that content.

    b. For example, members of the LGBTTQIA+ community often experience harm when others engage with content that incites hate and violence against that community.

64. We were therefore pleased to hear during discussions with DIA that the **risk of harm**, rather than a demonstration of harm from victims, will determine whether content can be considered harmful.

    a. Similarly, we were encouraged to hear from DIA that the Independent Regulator will be expected to commission research into ways to assess potential harm such as the work currently being conducted by the Mental Health Foundation, so that the definition of “risk of harm” is evidence-based rather than arbitrary.

65. Finally, there are context-dependent cases where content may cause harm but only under certain conditions.

    a. For example, it is possible for a particular domain name to be innocuous, and website content to be similarly innocuous, but when combined they could represent harmful content.

    b. In addition, some forms of content experienced once may not cause harm, but repeated exposure could result in cumulative effects that cause damage to emotional or mental wellbeing.

66. We believe it is important for the Independent Regulator to consider these forms of harm when co-developing industry codes, and believe the proposed Advisory Board outlined above will play an integral role in informing the Independent Regulator of these contextual cases.

R11 Revise the definition of "harmful content" so that it is clear that an evidence-based evaluation of risk of harm, rather than the experience of harm, is central to the definition or classification of harmful content.

R12 Require the Independent Regulator to seek out and incorporate information about context-dependent risks of harm as a part of its decision-making processes.

Regulated Platforms

67. Another key concept in the SOSMP discussion document is a "Regulated Platform".

    a. The document makes a distinction between **platforms**, which it defines as "providers of content and services – for example, social media companies or broadcasters" (p. 4), and Regulated Platforms, which are those that would be subject to industry codes.

68. In the discussion document there is a clear emphasis on "larger or riskier platforms" (p. 6).

    a. The document indicates that a Regulated Platform is likely to have an expected annual audience of 100,000 or more, or 25,000 account holders per year in Aotearoa New Zealand (p. 28).

69. The SOSMP discussion document states that "there would be little consumer safety value in placing a high compliance burden on very small platforms" (p. 28, para. 31)

70. However, while the scale and influence of larger platforms does make them an important focus of regulation, smaller platforms can sometimes have the most harmful content, in part because of indifference to that content, and in part because smaller platforms lack the personnel and resources necessary to moderate harmful content effectively.

71. In addition, larger platforms can share or link to content hosted on smaller platforms, broadening the reach of these smaller platforms. While that may mean that these smaller platforms meet the threshold for becoming a Regulated Platform based on audience size, that is not always the case.

72. Considering these issues, we are encouraged to see that the Independent Regulator may designate smaller platforms that demonstrate significant influence, repeatedly share objectionable content, or repeatedly ignore requests to remove objectionable content as a Regulated Platform (p. 28).

73. However, because of the potential for insufficient staff or resources for effective content moderation, these smaller platforms may need assistance with identification of, and proactive monitoring for, harmful and objectionable content.

74. Smaller platforms may also benefit from resources, such as monitoring and response models and baseline requirements.

R13 The Independent Regulator should be provided with budget and other resources necessary to develop informational resources to help smaller platforms address harmful content and, if necessary, assists smaller platforms with monitoring for and removing harmful and objectionable content.

75. The discussion document currently does not offer a lot of clarity about possible Regulated Platforms beyond (large) social media platforms.

    a. For example, the discussion document clearly states: "The major change, in practice, would be in the way social media platforms are regulated" (p. 25).

76. Other 'platforms', including static websites, message boards, and blogs, as well as multimedia content such as podcasts and emerging, de-centralised systems such as the Fediverse or Bluesky can also be a source of harmful and objectionable content.

    a. For example, a recent report from Stanford University details how inconsistent moderation policies on the Fediverese system Mastodon have led to a significant increase in child sexual abuse material (CSAM) there.[12]

    b. How these various media would be subject to Industry Codes and what industry body would represent them is unclear.

77. Finally, we are concerned about potential loopholes in the definition of "Regulated Platform".

    a. The discussion document makes clear that the listed definition "intends to exclude platforms and services that exist primarily to enable other services and products – for example, the websites of retailers, professional services, clubs, and charities would not be considered Regulated Platforms" (p. 28, para. 32).

    b. However, an individual or group intent on spreading harmful content could theoretically register as a charity or a club and use their website to spread that content.

    c. In addition, retailer and professional services websites could potentially sell goods that could be classified as harmful or objectionable, and many feature customer comments and reviews that could also contain harmful content.

[12] See the report "Child Safety on Federated Social Media" at https://purl.stanford.edu/vb515nd6874 for more information.

d. In order to be equitable and effective, we believe these exemptions should be removed from the definition of Regulated Platforms.

e. Any issues arising from the lack of these exemptions can be addressed in the development of industry codes, with support as needed for smaller platforms that lack necessary resources, or through redress processes available through the proposed Recourse Council outlined above.

R14 Amend the definition of "Regulated Platform" so that there are no exclusions for charities, clubs, retailers, or professional services.

Monitoring and enforcement

Proactive monitoring requirements

78. An element of the SOSMP discussion document that would benefit from more discussion and development, and perhaps some caution, is the inclusion of proactive monitoring and removal requirements.

79. The discussion document states that Regulated Platforms would need to engage in "**proactive and consistent moderation** of both content and harmful conduct in relation to content" (p. 28, para. 30).

    a. A proactive approach can be beneficial in its potential to address systemic harms in advance, rather than responding to harmful content after it has caused harm.

    b. This approach is currently often used by platforms to identify terrorist violent extremist content (TVEC) and child sexual abuse material (CSAM).

    c. This proactive monitoring is carried out both using automated systems, as well as human monitoring, depending on the platform and resources available.

    d. Organisations such as the Global Internet Forum to Counter Terrorism (GIFCT) also support platforms proactively monitoring for objectionable content (in this case, terrorist and violent extremist content [TVEC]) by providing information, tools, and databases of identified TVEC content to assist with automatic detection.

80. We have some questions and concerns about how proactive monitoring will work in practice.

    a. Example 2 under, "High-level safety objectives/outcomes" (p. 38), suggests that a code provision for a platform with user-generated content (UGC) might be that a "machine learning process will be used to flag potentially explicit user-uploaded content; a warning will be displayed".

i. However, research shows that these automatic detection processes can be inconsistent, inaccurate (resulting in false positives), and easy to 'fool' (resulting in harmful content going unnoticed), and replicate human biases such as misogyny or racism.[13]

b. It is also unclear what "authenticity markers" (p. 36) will be used to identify harmful content, how they will be determined, and by whom.

i. The determination of these authenticity markers is another opportunity for the introduction of bias, and therefore needs significant input from members of communities most affected by online harms, and oversight from community advocates and technology, legal, and subject matter experts.

c. Finally, the procurement, development, and deployment of proactive measures can, as outlined above, overburden smaller platforms that lack sufficient staff or resources.

81. In addition, both civil society organisations and tech companies have expressed concerns about requirements to remove such content quickly, which the SOSMP discussion document suggests the Independent Regulator would be able to impose (p. 7).

a. For example, civil society groups in other countries have argued that limited time requirements can overburden smaller platforms or result in over-censorship of content such as that from journalists, researchers, activists, and those sharing counter-messaging to content such as hate speech or TVEC.

1. In Australia, both Google and the Australian Industry Group (Ai) argued that a proposal in the *Online Safety Act 2021* to shorten the takedown time for illegal content from 48 to 24 hours might disadvantage smaller service providers, possibly because they are less likely to have the staff or resources needed to respond quickly.[14]

2. In Canada, several responses to the Technical Paper released in relation to the proposed online harm legislation there express concern that heavy penalties for not removing content flagged as harmful within 24

[13] See, for example, *The impact of algorithms for online content filtering or moderation*, available at https://www.europarl.europa.eu/RegData/etudes/STUD/2020/657101/IPOL_STU(2020)657101_EN.pdf, pp. 45-47.

[14] More information on responsese to *Australia's Online Safety Act 2021* can be found in section 1.13 of *International Regulatory Frameworks for Online Content*, available at https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/International-Regulatory-Frameworks-for-Online-Content-Report.pdf.

hours might result in the wide scale removal of legal and non-harmful content.

3. Google also expressed concern that proactive monitoring could result in the suppression of legal expression in a manner inconsistent with international norms.

4. Finally, critics in both the technology sector and civil society noted that a 24-hour time limit could be weaponised against marginalised communities or taken advantage of to harass others and limit expression.[15]

82. Therefore, we agree with the proposal that the Independent Regulator adopt a "**risk-based approach** to ensuring a platform's compliance with relevant codes of practice" (p. 43, para. 69).

    a. Our discussions as a part of the Christchurch Call Advisory Network have highlighted how risk-based approaches can move beyond addressing individual pieces of content and instead address broader issues such as platform design, processes, and policies.

    b. In other words, a risk-based approach expands possibility for intervention beyond content blocking or removal, and looks for opportunities to limit harm by investigating elements such as algorithmic amplification or recommendation systems, and user interactions both the platform and other users.

Expansion to material illegal under other New Zealand regimes

83. The option to expand takedown powers to "material that has been found to be illegal under other New Zealand regimes" (p. 55, para. 106) is one way to provide flexibility to this regulatory framework.

84. We are generally supportive of this approach, but this power would need some practical limitations to prevent overreach and limit scope. We would be disappointed to see this framework be used as a way to tighten copyright controls, for example.

85. As a starting point, our recommendation is that this expansion of takedown powers be limited to other Aotearoa regimes that deal with activities that can result in risk of harm to physical, social, emotional, and mental wellbeing, in line with the definition of harmful content outlined in the discussion document.

[15] See *International Regulatory Frameworks for Online Content*, pp. 85-86, available at https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/International-Regulatory-Frameworks-for-Online-Content-Report.pdf.

86. Doing so would ensure that the focus of this framework remains on limiting the spread of harmful and objectionable content while also ensuring the remit of the Independent Regulator does not become too broad.

R15 The Independent Regulator should have takedown powers for material that is illegal under other New Zealand regimes, but only material that can result in harm to physical, social, emotional, and mental wellbeing.

Penalties

87. As noted in the discussion document, "online content hosts that fail to comply with takedown notices are subject to a civil penalty of up to $200,000 for each incident of non-compliance" (p. 56, para. 109)

88. While we agree that there should be statutory civil penalties for non-compliance, particularly for failing to comply with takedown notices relating to objectionable content, these penalties need to be significant enough to encourage compliance and also adaptable enough so that penalties are proportional to the level of non-compliance (e.g., higher for repeated offences) and the financial resources of the offending platform.

89. One reason for the need for this flexible approach is that there is evidence that what seem like significant financial penalties are actually ineffective against the largest online platforms in terms of user base and revenue.

    a. For example, after the company then known as Facebook was issued a record US$5 billion fine by the US Federal Trade Commission (FTC) in 2019 for misuse of users' personal information, the company's stock price increased, and the net worth of the company's founder, chairman and CEO, Mark Zuckerberg, actually *increased*.

    b. At the time, some analysts attributed the rise in the company's stock price to the relatively insignificant amount of the fine, in proportion to the company's revenue, which was US$15 billion in the quarter before the fine was issued alone.

90. At the same time, a comparatively small fine for a large platform can be beyond the means of and disproportionately debilitating for a smaller platform that does not have the resources to pay it.

91. We are therefore in favour of exploring options for fines similar to those introduced in the EU's Digital Services Act, which stipulates that platforms can be fined up to 6% of their global turnover in cases of non-compliance.[16]

    a. A similar approach in Aotearoa can provide the Independent Regulator with the flexibility to impose fines that are both contextual to the level of non-compliance and the financial capabilities of individual platforms.

[16] The full text of the EU Digital Services Act can be found at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R2065.

Last resort remedies

92. The SOSMP discussion document raises the specific question of "what the right 'last resort' remedy" might be for persistent and serious non-compliance by platforms that host the most harmful content" (p. 48, para. 81).

93. Possible options outlined in the document (p. 48, para. 82) include:

    a. Additional, larger financial penalties.

    b. Prosecution for breaching New Zealand law.

    c. Asking a judge to impose service restrictions or service disruptions.

94. Of these, our view is that financial penalties are the best and most enforceable option of those presented.

    a. As outlined in the section "Penalties" above, however, the system for determining financial penalties needs to be flexible, contextual, and proportionate, both to the resources of the platform in question and the severity of the offence.

    b. Prosecutions or criminal sanctions, for people associated with platforms that are noncompliant, would be difficult to enforce considering the size of Aotearoa and the fact that executives or directors of large online platforms, such as social media companies, are generally outside of the reach of our law enforcement agencies.

        i. People within the Aotearoa jurisdiction could be held to account but, for large online platforms, any employees who might be in Aotearoa are unlikely to have a significant role in determining platform policies or whether that company will comply with takedown orders.

    c. We are also not in favour of imposing service restrictions or service disruptions, particularly at the domain name level and on online platforms on which people can contribute content.

    d. Historically, InternetNZ | Ipurangi Aotearoa has opposed filtering or blocking of content for reasons including those listed below, which is what a service disruption essentially represents.[17]

        i. As we have previously outlined, service disruptions such as blocking or filtering are relatively easy to circumvent, for example, via tools such as Virtual Private Networks (VPNs).

[17] See, for example, our position paper "*To block or not to block: Technical and policy considerations of Internet filtering*", available at https://internetnz.nz/assets/Archives/Content_Blocking_InternetNZ.pdf.

ii. In addition, the replicability of digital content, especially online, makes it difficult to effectively target blocks or filters.

iii. Finally, blocking or filtering can inadvertently block legal content and affect activists, minority voices, academics, or journalists looking to analyse or provide counter-narratives for harmful and objectionable content.

iv. In short, service disruptions can pose significant challenges to human rights and freedom of expression.

95. Despite our opposition to imposed service disruptions or similar remedies such as filtering or blocking, we are in favour of "safety measures including voluntary filters" (p. 54, para. 101), but only if there is transparency, accountability, and proper consultation.[18]

    a. In particular, we support new voluntary filters if they would "apply only to material that's already illegal, or where the material can confidently be deemed illegal" (p. 55, para. 105), as is specified in the discussion document.

    b. However, we would encourage that these voluntary filters be coupled with educational and support information (see the "Education and support" section below for more on this proposed approach).

    c. DIA has used a similar approach in conjunction with the Digital Child Exploitation Filtering System (DCEFS). The operation of the DCEFS is overseen by an Independent Reference Group (IRG) to ensure the filter's scope does not expand beyond its remit.

    d. Any voluntary filters developed as a part of this framework could be subject to oversight by the Advisory Board outlined above.

Interaction with other agencies and laws

96. Since this framework is intended to regulate platforms, we believe it makes sense, and are supportive of the fact that individuals would not be "directly subject to the regulator" unless they are also platform operators (p. 30, para. 36).

97. One reason we support the exclusion of individuals from this framework is that we believe existing legal regimes exist that address the actions of individuals.

[18] For more on our position on voluntary filters, see our previous submission on the *Films, Videos, and Publications Classification (Urgent Interim Classification of Publications and Prevention of Online Harm) Amendment Bill*, available at https://internetnz.nz/assets/Archives/InternetNZ-Final-submission-on-FVPCA-Bill-and-Internet-filtering.pdf.

98. However, a question that often arose in our consultations with communities and stakeholders was how this proposed framework will interact with, or affect, existing laws and regimes.

99. One commonly mentioned law is the **Harmful Digital Communications Act (HDCA)**.

    a. For example, several people we talked to asked why this new framework was necessary, considering the fact that the HDCA exists.

    b. The SOSMP discussion document does clarify that "offences under the Harmful Digital Communications Act, such as online bullying and harassment, would likely not meet the current threshold for a takedown notice issued by the Department of Internal Affairs" (p. 55, para. 107).

    c. This difference is one of the reasons we are supportive of expanding the Independent Regulator's takedown powers to expand "material that has been found to be illegal under other New Zealand regimes" (p. 55, para. 106) in cases where that material can result in harm to a person's physical, social, emotional, or mental wellbeing (see "Expansion to material illegal under other New Zealand regimes" section above).

100. Another common question raised during our consultations is why online scams and cybersecurity breaches are outside the scope of this framework (p. 62, para. 127).

    a. According to the definition of harmful content outlined in the discussion document, which includes "damage to rights, property", a scam resulting in loss of money could potentially be classified as damage to property.

    b. We recognise that scams and other forms of fraud are already illegal under other legal regimes and that other bodies, such as CERT NZ are currently tasked with responding to such events.

101. However, we encourage future iterations of this framework to better explicate how this framework will work in conjunction with existing laws and organisations.

    a. In addition, we recommend that a formal process for "funnelling" complaints to other bodies and agencies, as needed, be developed, so that the goal of a "simplified complaints process" is realised.

R16 Develop a process to channel complaints handled by other agencies to those organisations to maintain a simplified complaints process.

Education and support

102. We also welcome the inclusion of **education and awareness initiatives** in the Independent Regulator's responsibilities.

    a. UNESCO guidelines that outline a multistakeholder approach to regulating online platforms specify that governments should couple educational programmes alongside regulation in order to empower users. The programmes should address a range of topics including information literacy, online safety, and rights to freedom of expression and privacy. UNESCO also stresses the importing of drawing on the expertise of academics, information literacy experts, libraries, and civil society.[19]

103. In particular, we support the plan for the Independent Regulator to fund "education initiatives that would be **by Māori**" (p. 67, para. 147), and "**encourage, or fund, industry or community-led awareness initiatives**, as well as consumer advisories about certain types of content" (p. 57, para. 112).

    a. Our preference would be for the Government to prioritise Māori and other community-led initiatives over industry initiatives. See para. 11 on p. 9 for more information.

104. For these educational initiatives to be effective, they would **require ample resourcing**, and input from communities, particularly those most affected by harmful content, while being developed.

    a. The proposed Advisory Board, outlined above, could provide the necessary context.

R17 Prioritise Māori and other community-led educational initiatives (and provide monetary support and other resources to them) to develop Aotearoa-specific educational content on how to report and minimise harmful content on media platforms.

105. In addition to developing educational programmes and content in cooperation with Māori and other community groups, we would prefer to see any voluntary filtering system (see the "Last resort remedies" section above) incorporate interstitials that explain why the content is being filtered and provide links to educational material and support services.

[19] See "*Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms*", p. 10, para. 29k, available at https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en.

.

a. This approach can help prevent people from accidently accessing objectionably or potentially harmful content, and help inform those who are merely curious about why the material they are attempting to access may cause harm.

b. DIA has used a similar approach in conjunction with the Digital Child Exploitation Filtering System (DCEFS). The specific content of these interstitials could be developed with input from the Advisory Board outlined above.

R18 Any voluntary online filters should incorporate the use of interstitials that provide links with education and support information to help minimise the effects of harmful and objectionable content.

Conclusion

106. This proposed framework marks the first opportunity in over 30 years to address harmful content on media platforms, and the first time to do so for online platforms.

107. Recent experiences and events clearly demonstrate that the current framework is not responsive to the unique challenges presented by modern platforms.

108. Moreover, the speed at which technologies can emerge and be taken up by users is far faster than developing legislative responses, and the scale at which those technologies can disseminate harmful content is amplified by the affordances of modern platforms.

109. Because of those challenges, we recognise how difficult this endeavour is, and that whatever framework comes out of this consultation will not work perfectly, or even well, from the start.

110. In short, the modern media environment underscores the importance of effective regulation. But, in order to be effective, it needs to be flexible enough to adjust where things are not working as intended and to respond to emerging technologies, issues, and threats; transparent and subject to oversight so that it does not limit human rights, freedom of expression, or freedom of the press; and responsive enough to both protect and support everyone in New Zealand, especially those from the communities most targeted and affected by online harms.

111. We appreciate that the DIA have provided this discussion document at a relatively early stage in the process and have indicated areas where they would like more feedback.

112. We encourage the DIA to continue their discussions, and to provide a future discussion document that provides more details and clarity based on feedback received during this process.

Want more detail? Get in touch.

Thank you again to DIA for the opportunity to comment on the Safer Online Systems and Media Platforms discussion document. We welcome the opportunity for further dialogue on how best to realise the outcomes online, both in that document and our submission above.

Please contact us at policy@internetnz.nz.

Dr Michael S Daubs

Senior Policy Advisor

InternetNZ

Ana McAllister

Senior Advisor, Māori Outcomes

InternetNZ

Responses to consultation questions

Definitions in the proposals

1. What do you think about the way we have defined unsafe and harmful content?

We have some suggestions on how to modify the definition of harmful content to reflect that the risk of harm, rather than the experience of harm, is central to the definition. Please see the section "Harmful content" starting on p. 24, including Recommendation 11, for more.

2. Does the way we have defined unsafe and harmful content accurately reflect your concerns and/or experiences relating to harmful content?

The definition does not clearly address cumulative or systemic harm experienced by some individuals and groups with whom we consulted while preparing this submission. Please see the section "Harmful content" starting on p. 24, para. 17 on p. 11, and Recommendation 2 for more.

About our proposed new framework to regulate platforms

3. Have we got the right breakdown of roles and responsibilities between legislation, the regulator and industry?

We have proposed a structure for a new Aotearoa Media Commission that includes additional structures, an Advisory Board and a Recourse Council, that includes more oversight and ensures the inclusion of Māori voices and those from the communities most targeted and affected by harmful content, a component missing from the roles and responsibilities breakdown in this question. We feel this structure will be better equipped to minimise these harms while also protecting human rights. The details of our proposed model are found in the section "Structure of the regulatory entity" starting on p. 11, and in Recommendations 3–6.

4. Do you agree that Government should set high-level safety objectives and minimum expectations that industry must meet through codes of practice?

Yes.

5. Do you agree with how we have defined 'platforms'? Do you think our definition is too narrow, or too broad? If so, why?

We have no specific feedback on how platforms are defined.

6. We are trying to focus on platforms with the greatest reach and potential to cause harm. Have we got the criteria for 'Regulated Platforms' right?

We understand the need for a starting point for classifying what is, or is not, a Regulated Platform. However, smaller platforms can also have content that causes significant harm disproportionate to their audience size or user base. We support the Independent Regulator having the ability to designate these

platforms as Regulated Platforms and argue for the need to support smaller platforms with resources to help them minimise objectionable and harmful content. In addition, we are concerned about exclusions for charities, clubs, retailers, or professional services. Please see paras. 70-77, starting on p. 26, and Recommendations 13 and 14 for more information.

7. Do you think we have covered all core requirements needed for codes of practice?

We believe that there should be significant and formalised input from Māori as treaty partners, Māori and other communities most affected and targeted by harmful content, and subject matter experts in law, media, and technology. This belief is one of the reasons behind our proposal for a robust Advisory Board. In addition, we propose that the Independent Regulator should have the power to impose industry standards in cases where co-development of industry codes proves ineffective. Please see the section "Industry codes of practice" starting on p. 22 and Recommendation 10 for more detail.

8. What types of codes and industry groupings do you think should be grouped together?

Our submission includes a proposal for sector-specific Commissioners within the Independent Regulator as a starting point for how codes and industries should be grouped. Please see the section "Proposal for sector-specific Commissioners within the Independent Regulator" starting on p. 12 and Recommendation 3 for more information.

9. Do you think some types of platforms should be looked at more closely, depending on the type of content they have?

Yes. In particular, we have flagged static websites, message boards, and blogs, as well as multimedia content such as podcasts and emerging, de-centralised systems such as the Fediverse or Bluesky as platforms that might require more consideration. Please see para. 76 starting on p. 27 for more detail.

10. Do you think the proposed code development process would be flexible enough to respond to different types of content and harm in the future? Is there something we're not thinking about?

Yes, but only if there is regular review of codes and development or revision of codes is context specific, informed by the knowledge of subject matter experts and the experiences and communities most affected by harmful content. We believe our proposed Advisory Board would be an important inclusion.

11. What do you think about the different approaches we could take, including the supportive and prescriptive alternatives?

We are more in favour of a somewhat more prescriptive approach, but believe the power to decide what content is objectionable should remain with the Classification Office and therefore separate from the Regulator. Please see

the section “Proposal to maintain the censorship function within the current Office of Film and Literature Classification” starting on p. 16, the section “Supportive vs prescriptive approach” starting on p. 24, and Recommendation 6 for more detail.

12. Do you think that the proposed model of enforcing codes of practice would work?

We have included a detailed discussion of our views on enforcement and potential penalties in the section “Monitoring and enforcement”, starting on p. 28. The sections “Expansion to material illegal under other New Zealand regimes” starting on p. 30, “Penalties” starting on p. 31, and “Last resort remedies” starting on p. 32 are particularly relevant to this question.

13. Do you think the regulator would have sufficient powers to effectively oversee the framework? Why/why not?

We have proposed a structure for a new Aotearoa Media Commission that includes two additional entities, an Advisory Board and a Recourse Council, to provide more oversight and ensure the inclusion of voices from the communities most targeted and affected by harmful content. These details are found in the section “Structure of the regulatory entity” starting on p. 11, and in Recommendations 3–6.

14. Do you agree that the regulator’s enforcement powers should be limited to civil liability actions? (For example, issuing formal warnings and seeking civil penalties for non-compliance)

Yes. Please see para. 88 on p. 31 for more discussion.

15. How do you think the system should respond to persistent non-compliance?

We believe that financial penalties are the most enforceable penalty for persistence non-compliance. However, these penalties need to be significant and adaptable, i.e., proportional to the level of non-compliance (e.g., higher for repeated offences) and the financial resources of the offending platform. Please see the section “Penalties” starting on p. 31 for more detail.

16. What are your views on transferring the current approach of determining illegal material into the new framework?

Our view is that the authority to determine what material is illegal or objectionable remains with the current Office of Film and Literature Classification as a separate entity outside of the Independent Regulator. This structure would prevent the concentration of too much authority within the Independent Regulator and also incorporate a level of oversight. Please see the section “Proposal to maintain the censorship function within the current Office of Film and Literature Classification” starting on p. 16 and Recommendation 6 for more information.

About the regulator

17. Should the regulator have powers to undertake criminal prosecutions?

Our view is that the Independent Regulator should have the power to *recommend* criminal prosecutions against Regulated Platforms on the grounds of significant non-compliance with industry codes, but that existing law enforcement agencies should carry out investigations of and prosecutions for illegal material. Please see para. 59 on p. 24 for more detail.

18. Is the regulator the appropriate body to exercise take-down powers?

Yes, with the caveat that the determination of objectionable material remains with the current Office of Film and Literature Classification.

19. Should takedown powers be extended to content that is illegal under other New Zealand laws? If so, how wide should this power be?

Our recommendation is that this expansion of takedown powers be limited to other New Zealand regimes that deal with activities that can result in harm to physical, social, emotional, and mental wellbeing, in line with the definition of harmful content outlined in the discussion document. Please see the section "Expansion to material illegal under other New Zealand regimes" starting on p. 30 and Recommendation 15 for more discussion.

20. If takedown powers are available for content that is illegal under other New Zealand laws, should an interim takedown be available in advance of a conviction, like an injunction?

Yes, but subject to rapid review in order to prevent Regulator overreach, similar to the emergency classification authority of the Chief Censor.

Summary of potential roles and responsibilities under the proposed framework

21. What do you think about the proposed roles that different players would have in the new framework?

We have proposed a structure for a new Aotearoa Media Commission that includes additional structures, an Advisory Board, and a Recourse Council, which has more oversight and ensures the inclusion of voices from communities most affected by harmful content. These details are found in the section "Structure of the regulatory entity" starting on p. 11, and in Recommendations 3-6.

22. Have we identified all key actors with responsibilities within the framework? Are there any additional entities that should be included?

We would like to see more, and formalised, inclusion of voices from the communities most targeted and affected by harmful content as part of this framework. Our proposed structure for a new Aotearoa Media Commission,

including an Advisory Board and a Recourse Council, reflect this desire. The details of our proposed model are also found in the section “Structure of the regulatory entity” starting on p. 11, and in Recommendations 3-6.

What would the proposed model achieve?

23. What do you think about how we’re proposing to provide for Te Tiriti o Waitangi through this mahi? Can you think of a more effective way of doing so?

The primary concern InternetNZ | Ipurangi Aotearoa has heard from our Māori partners is that engagement with Māori up to this point has not been adequate. In addition, the processes and mechanisms by which the government would meet its obligations under Te Tiriti o Waitangi need clarification. Please see the section “Partnership with Māori” starting on p. 8, and Recommendations 1 and 2, for a detailed discussion.

24. Do you think that our proposals will sufficiently address harms experienced by Māori?

The harms experienced by Māori, particularly wāhine Māori, are systemic, cumulative, and generational. We are not confident at present that this proposal will effectively address those harms. We recommend that more targeted engagement is undertaken by DIA on this issue, with a particular focus on wāhine Māori who have been victims of harmful content through orchestrated and deliberate campaigns of hate. “Partnership with Māori” starting on p. 8, and Recommendation 2, for more information.

25. What do you think about how rights and press freedoms are upheld under the proposed framework?

We support the stated commitment to upholding human rights, freedom of expression, and press freedom in the discussion document, and recognition that freedom of expression needs to be balanced with other human rights. We acknowledge that this is a difficult balance. Our proposed model for an Aotearoa Media Commission incorporates rigorous transparent and oversight requirements for both platforms and the Independent Regulator in an effort to achieve that balance. In addition, we propose a “Commissioner” within the Independent Regulator specific to news and journalism platforms as an additional layer of protection for press freedom. Please see the sections “Human rights and freedom of expression” starting on p. 7, “Proposal for sector-specific Commissioners within the Independent Regulator” starting on p. 12, and “Oversight and transparency” starting on p. 18, as well as Recommendations 3 and 9 for more information.

26. Do you think that our proposals sufficiently ensure a flexible approach? Can you think of other ways to balance certainty, consistency, and flexibility in the framework?

We have made several recommendations on ways to ensure consistency and flexibility, as well as transparency, in the framework. The sections “Structural review” and “Oversight and transparency” starting on p. 18, and Recommendations 7 and 9 are most relevant to this question. However, our overall proposed structure, outlined in the section “Structure of the regulatory entity” starting on p. 11, as well as recommendations 3-5, 7, 9, 10, 12, and 15 are relevant as well.