17 June 2026

Submission on “Deepfake Digital Harm and Exploitation Bill”

Social Services and Community Select Committee

internetnz

Who we are and why we are submitting

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. As the managers of the .nz domain, InternetNZ is committed to ensuring that our digital environment remains open, secure, and resilient for all.

We are a not-for-profit organisation, and the money we receive from .nz domain names is reinvested in the community. We provide grants, help to fund initiatives, and advocate for an accessible Internet that benefits everyone in Aotearoa.

You can read more about our work here and in our latest Annual Report. Our submissions on other public policy reviews are available here.

InternetNZ conducts research and works in partnership with others - both here and globally - to better understand Internet-related issues, including how New Zealanders are experiencing the online environment, .nz consumer research, and technical issues with the Internet.

Alongside our technical expertise, the research and funding role that we play within the information ecosystem in New Zealand provides us with unique insights into the issues raised by the Deep Fake Digital Harm and Exploitation Bill.

Executive Summary

This submission responds to changes proposed in the Deepfake Digital Harm and Exploitation Bill, currently under review by the Social Services and Community Select Committee.

Our submission supports the specific changes proposed to update the definition of ‘intimate visual recording’ in both the Crimes Act 1961 and the Harmful Digital Communications Act 2015. The proposed changes respond to technological changes that have expanded the types of intimate recordings that can be generated and will better account for the current digital environment that includes intimate deepfake imagery and audio recordings on a large scale.

To account for the full range of Artificial Intelligence-generated multimedia that can be considered intimate and cause harm, audio deepfakes should be included alongside the expanded definitions of ‘intimate visual recording’ as further changes to the Crimes Act 1961 and the Harmful Digital Communications Act 2015. This is in line with international categorisation of deepfakes including the definition used by the Australian eSafety Commissioner.

While we agree the harm caused by non-consensual sexually explicit deepfakes that this bill focuses on should be addressed, we highlight the complexity of the drivers and forms of online harm that New Zealanders experience daily. Effective solutions require systems-level changes to New Zealand's digital policy environment. We

recommend that work is progressed to conduct a review of New Zealand's digital regulation settings and an independent digital regulator is established.

We recommend the independent digital regulator is mandated to build public digital literacy and online harm resilience through a comprehensive digital education programme. We know that public understanding of how to interact safely with online content, including how to prevent harm and respond effectively if they encounter potentially harmful content, is essential to support online safety.

We can provide further advice on any of the points raised in this submission upon request.

Recommendations

We recommend:

1. Progressing the suggested amendments to definitions of 'intimate visual recording' in the Crimes Act 1961 and Harmful Digital Communications Act 2015 to align with technological changes, like generative Artificial Intelligence, that have expanded the types of intimate visual recordings that can be created
2. Expanding the proposed changes to include deepfake audio recordings to account for the full range of Artificial Intelligence-generated multimedia that can be considered intimate and cause harm
3. Establishing an independent digital regulator responsible for mitigating online harm
4. Mandating the independent digital regulator to lead a comprehensive digital education programme to support online safety.

We support expanding definitions of “intimate visual recording” to reflect technological changes

The digital environment is changing rapidly. The emergence of generative Artificial Intelligence (AI) has led to the widespread accessibility of tools that can generate digital imagery. This includes the production of deepfakes, generally understood to be: a digital photo, video or sound file of a real person created with AI to make a realistic but false depiction of them doing or saying something that they did not do.[1]

Technological changes have amplified the speed at which a deepfake can be created and contributed to significant growth in the number of deepfake images. Cybersecurity firm DeepStrike estimated a global increase from roughly 500,000 online deepfakes in 2023 to about 8 million in 2025, citing annual growth of deepfake imagery at nearly 900%.[2]

We know that deepfake use is not inherently exploitative or harmful and can be used for purposes that are beneficial or neutral such as to synthesise new pharmaceutical compounds,[3] protect wildlife from poachers[4] and for entertainment purposes such as the creation of satire videos.

However, we also know the vast majority of deepfake videos involve pornography and deepfake sexual images are being generated at an unprecedented rate. A Bloomberg article from January 2026 reports Grok being used to generate upwards of 6,700 sexual images per hour[5] and a 2023 study from Security Hero[6] found that:

- deepfake pornography makes up 98% of all deepfake videos online, amassing over 300 million video views in 2023 across the top 10 dedicated deepfake pornography websites,
- 99% of deepfake pornography features women as the primary subject,
- one in every three deepfake tools allow users to create deepfake pornography, and;

---

[1] Australian eSafety Commissioner: https://www.esafety.gov.au/industry/tech-trends-and-challenges/deepfakes#:~:text=A%20deepfake%20is%20a%20digital,not%20actually%20do%20or%20say.

[2] Khalil, M. (2025, September 8). *Deepfake statistics 2025: AI fraud data & trends*. DeepStrike. https://deepstrike.io/blog/deepfake-statistics-2025

[3] See example: https://theconversation.com/generative-ai-and-deepfakes-are-fuelling-health-misinformation-heres-what-to-look-out-for-so-you-dont-get-scammed-246149

[4] See example: https://www.aspi.org.au/report/weaponised-deep-fakes.

[5] See example: https://www.bloomberg.com/news/articles/2026-01-07/musk-s-grok-ai-generated-thousands-of-undressed-images-per-hour-on-x

[6] Security Hero. (2023). *2023 state of deepfakes: Realities, threats, and impact*. https://www.securityhero.io/state-of-deepfakes/#key-findings

- it takes less than 25 minutes and costs $0 to create a 60-second deepfake pornographic video of a person using one clear face image.

Recent reporting similarly highlights the continued proliferation of non-consensual, AI-generated pornography. For example, analysis of Internet Watch Foundation data showed a 260-fold increase in AI-generated child sexual abuse material (AI-CSAM) between 2024 and 2025.[7]

Changes to the definition of 'intimate visual recording' in the Crimes Act 1961 and the Harmful Digital Communications Act 2015 (HDCA), as proposed by this bill, supports these legal regimes to more accurately reflect these changes in the digital environment and the potential for harm to be caused by deepfake digital recordings and multimedia.

Deepfake audio recordings should be considered alongside other deepfake multimedia

Audio deepfakes are similar to video or photo deepfake imagery where AI generated or manipulated audio is created to convincingly mimic a person's voice. Audio deepfakes can be used for exploitative purposes including in an intimate context where they have the ability to cause reputational, psychological, relationship and economic damage to a person.

Advanced audio synthesis is a core element of contemporary AI-generated multimedia creation with a 2024 review citing over 1000 separate online tools capable of AI voice generation and cloning.[8] These tools carry the potential for considerable harm and AI-generated child sexual abuse material with an audio component has been identified as a growing area of concern by global online harm prevention groups like Internet Watch Foundation.[9]

We recommend proposed changes progressed under this bill be expanded to include audio recordings to cover all forms of content manipulation that can be used in the generation of non-consensual, intimate deepfakes.

This inclusion aligns with regulatory approaches in comparable jurisdictions such as Australia, where a 2025 amendment to New South Wales State legislation criminalised the alteration of image and audio material used in the production of "wholly digitally generated sexually explicit intimate image and audio material."[10]

---

[7] Internet Watch Foundation. (2026). IWF AI CSAM report 2026. https://www.iwf.org.uk/media/hl1nvdti/iwf-ai-csam-report-2026.pdf

[8] Cavalli, F. (2024). The state of deepfakes 2024. Sensity AI. https://5865987.fs1.hubspotusercontent-na1.net/hubfs/5865987/SODF%202024.pdf

[9] (Internet Watch Foundation, 2026).

[10] Crimes Amendment (Intimate Image and Audio Material) Act 2025 (NSW). https://www.parliament.nsw.gov.au/bills/Pages/bill-details.aspx?pk=18792

Online harm is complex – system level digital regulation and public education is essential

Online harm encompasses all content and activities on the Internet that can cause harm to individuals and society. In addition to the exploitation and harm caused by non-consensual deepfake imagery that this bill focuses on, online harm includes violent extremist content, child sexual abuse material, mis- and disinformation, dangerous and discriminatory speech (including hate speech and online harassment), encouragement of self-harm, image-based abuse and cyberbullying.

Online harm is increasing as the Internet's uses and technical enablers rapidly change. Some communities, particularly minority and marginalised groups, are particularly exposed to online harm. The severity of online harm is now widely documented, with impacts that encompass all aspects of civic participation, well-being, and personal safety.

New Zealand's digital regulation landscape remains fragmented and ineffective

Internet use has evolved quickly. New Zealanders rely on the Internet to participate in modern society, including for essential daily activities such as accessing government and financial services, and communications. The Internet is now a converged space where social media and other digital platforms are widely used, and anyone can create, share and amplify content.

Much of the existing legislation was designed to regulate traditional media with distinct broadcasters and a passive audience. As a result, New Zealand's digital regulation is highly fragmented and ill-equipped to handle the speed, scale, and architectural complexity of today's Internet.

Current digital regulation traverses a range of legal regimes and is managed by multiple government agencies. Responsibilities are scattered across entities, including the NZ Police, NetSafe, the Department of Internal Affairs, National Cyber Security Centre (NCSC), the Human Rights Commission, and media bodies. This creates legislative gaps, enforcement vacuums, and public confusion about where to report issues.

Due to the global nature of online harm, it is also important to note that New Zealand is becoming increasingly out of step with our international counterparts who are updating their digital regulation to respond to online harm. See Appendix A (page 7) where we have set out an overview of online harm regulation in comparable international jurisdictions.

Establish an independent digital regulator to promote online safety

There is a strong need for a comprehensive review of existing regulatory settings and intentional design of system-level digital regulatory architecture. Future regulation must be dynamic, technologically precise and enable future proofing.

Effective digital regulation would be supported by an independent digital regulator, Māori advisory representation, and real-time audit powers to enforce dynamic compliance, drive "safety by design", and ensure localised, accessible and responsive complaint systems.

Enabling the independent regulator to take a prescriptive approach to the development and enforcement of industry codes of practice, and to institute industry standards when industry representatives cannot or will not develop satisfactory codes of practice would also support effectiveness.

Any further development of these online regulation proposals should be co-designed and co-governed with Māori and further engagement with Māori should be undertaken, with a particular focus on wāhine Māori who have been identified as targets of harmful content.

We can provide you with more detailed advice, including comparisons with other international jurisdictions who have implemented similar regulatory structures upon request.

Mandate an independent digital regulator to lead public education on online safety

Due to the rapidly evolving enablers of online harm, public understanding, including how to prevent, identify and report harmful online content, is an essential part of steps required to achieve greater online safety.

Netsafe, who are mandated to provide online safety education under the Harmful Digital Communications Act 2015, are not sufficiently resourced to provide the digital education required to keep pace with the online harm environment. Netsafe have noted that their scams helpline is not sufficiently funded to keep up with demand and flagged barriers in engaging with young people in their outreach efforts.[11]

We recommend an independent digital regulator is mandated and resourced to build digital literacy and online harm resilience through a comprehensive digital education programme. This work would complement Netsafe's existing work and fill critical gaps. It would be important that the independent regulator works in partnership with Māori and other community groups to develop Aotearoa-specific educational content on how to report and minimise harmful online content.

[11] Netsafe. (2025). *Netsafe annual review*. https://resource.netsafe.org.nz/Netsafe-2025-Annual-Review.pdf

Appendix A

Overview of International Online Harm Regulation

Jurisdiction	Criminalisation of deepfakes (non-consensual, explicit)	Online safety legislation	Digital regulator	Image-based abuse reporting portal	Online platform reporting requirements
New Zealand	*proposed*	×	×	✓	×
Australia	✓*	✓	✓	✓	✓
Canada	*proposed*	*proposed*	*proposed*	*proposed*	✓
United Kingdom	✓	✓	✓	✓	✓
European Union	✓*	✓	✓	×	✓
United States	✓	×	×	✓**	✓*

** Some states/by directive (EU)*

*** For those under 18 (US).*