



Netsafe submission on the Deepfake Digital Harm and Exploitation Bill

Summary

Netsafe strongly supports the Deepfake Digital Harm and Exploitation Bill (**the Bill**).

The Bill addresses a gap in New Zealand's legal framework. It clarifies that existing intimate-image protections apply equally to synthetic and AI-generated content. Its strength lies in building on an established framework rather than creating a separate regulatory regime for deepfakes.¹ The Bill helps ensure that protections against non-consensual intimate imagery remain effective as technology evolves.²

In summary, Netsafe submits:

- People should receive the same protection whether intimate content is authentic, altered, or entirely synthetic.
- Netsafe advocated for this reform during Parliament's consideration of the Harmful Digital Communications (Unauthorised Posting of Intimate Visual Recording) Amendment Bill in 2021 and welcomes Parliament now progressing legislation consistent with that position.
- Netsafe's frontline experience as the Approved Agency under the Harmful Digital Communications Act 2015 (HDCA) demonstrates that the harms associated with intimate visual recordings are not limited to authentic images. Victims experience the same humiliation, distress and reputational harm regardless of how the content was created.
- Netsafe supports the overall intent and approach of the Bill but recommends one targeted amendment to ensure the definition of "intimate visual recording" remains consistent with the existing HDCA framework. Specifically, Netsafe recommends that the proposed synthetic-media provision applies **with or without the knowledge or consent** of the person who is the subject of the recording, consistent with the existing HDCA definition of "intimate visual recording".

About Netsafe

Netsafe is New Zealand's independent, non-profit online safety organisation. Embracing a technology-positive approach, we aim to help New Zealanders take advantage of the digital opportunities available

¹ Esther Wong, "Deepfakes and New Zealand's Law: An Evaluation of the Deepfake Digital Harm and Exploitation Bill," *Equal Justice Project* (Amicus Curiae), January 27, 2026, <https://www.equaljusticeproject.co.nz/articles/deepfakes-and-new-zealands-law-the-deepfake-digital-harm-and-exploitation-bill>

² Diana Clement, "Legislative Gap Leaves New Zealand Exposed to Deepfakes," McVeagh Fleming, March 18, 2024, <https://www.mcveaghflaming.co.nz/articles/legislative-gap-leaves-new-zealand-exposed-to-deepfakes>

while managing online challenges. We provide free support, advice and education through a national helpline, dispute resolution services and online resources.

We have been appointed as the Approved Agency under the HDCA since 2016 and deliver the statutory functions of receiving, assessing and seeking to resolve complaints, maintaining relationships with online content hosts and relevant agencies, and providing education and advice on online safety.

Reform is Overdue

The HDCA has never been limited to authentic content. Netsafe has long dealt with complaints involving false, manipulated and misleading digital communications. This Bill provides greater certainty that existing intimate-image protections apply equally to synthetic intimate imagery.

Victims experience the same level of distress and loss of dignity regardless of whether intimate content is authentic, altered or completely synthetic.³ Generative AI has not created a new category of harm; rather, it has made existing forms of image-based abuse easier, cheaper and faster to distribute on a large scale.⁴

This is not a new issue. Netsafe and YouthLaw Aotearoa raised concerns during the Justice Committee's consideration of the Harmful Digital Communications (Unauthorised Posting of Intimate Visual Recording) Amendment Bill 2021. YouthLaw submitted that victims of pornographic deepfakes should receive the same protection as victims of intimate image-based abuse. They suggested amending the definition of "intimate visual recording" in section 4 of the HDCA to capture manipulated intimate imagery.⁵ Netsafe similarly submitted that the definition should be clarified to ensure it was capable of capturing synthetic media, including deepfakes.⁶ We therefore welcome Parliament now progressing a legislative solution that addresses a previously identified gap in the HDCA framework.

³ VINE, "New Evidence Brief on the Escalation of Online Violence Against Women in the Public Sphere," VINE, January 26, 2026, <https://vine.org.nz/news/new-evidence-brief-on-the-escalation-of-online-violence-against-women-in-the-public-sphere>

⁴ Katrine Evans, "Deepfakes and Privacy: It's Time to Respond," Privacy Foundation NZ Inc, January 23, 2026, <https://privacyfoundation.nz/deepfakes-and-privacy-its-time-to-respond/>

⁵ YouthLaw Aotearoa, Submission on the Harmful Digital Communications (Unauthorised Posting of Intimate Visual Recording) Amendment Bill (submission to the Justice Committee, New Zealand Parliament, 23 April 2021), https://youthlaw.co.nz/wp-content/uploads/2016/07/20210423_FINAL_YouthLaw_Harmful-Digital-Communications-Unauthorised-Posting-of-Intimate-Visual-Recording-Amendment-Bill.pdf

⁶ Netsafe, Submission on Harmful Digital Communications (Unauthorised Posting of Intimate Visual Recording) Amendment Bill (submission to the Justice Committee, New Zealand Parliament), https://www3.parliament.nz/resource/en-NZ/53S CJU_EVI_99360_JU1443/8bd4e6c2568b0d98c15ebd433beb27e3fbfcd1c2

Frontline Evidence of Harm

Netsafe's support for the Bill is informed by our frontline experience. During FY25, Netsafe's helplines, reporting services and online resources recorded more than 400,000 interactions, including more than 26,000 reports of scams and online harm incidents requiring intervention and over 375,000 visits to Netsafe's online safety resources. In FY25, Netsafe received a record 6,404 HDCA complaints. With only a few days remaining in FY26, that figure has already increased by more than 22 percent. Complaints involving deepfakes and other forms of synthetic media have more than doubled in the last year.⁷

While reports to Netsafe represent only a small proportion of the harms occurring online, they reflect broader global trends. The number of deepfakes shared online worldwide has increased from around 500,000 in 2023 to approximately eight million last year.⁸

Netsafe increasingly encounters harms facilitated by generative AI and image manipulation technologies among the complaints it receives. Examples from complaints we received in FY25 illustrate how synthetic media is amplifying existing forms of online harm:

- **Non-consensual intimate imagery:** Netsafe increasingly receives complaints involving the creation and distribution of AI-generated intimate imagery without consent. In one case, a student used AI tools to create and share intimate images of their classmates. This caused significant distress to those affected and disrupted both the school community and the students' learning.
- **Harassment:** Netsafe is seeing growing numbers of complaints involving the use of AI-generated and manipulated intimate imagery to harass, intimidate and exert pressure on others. These harms occur across a range of contexts, including intimate relationships, schools, workplaces and interactions between people who have never met.
- **Sextortion:** Netsafe receives reports of offenders using synthetic intimate imagery to facilitate sextortion. In one recent case, a minor was targeted using a sexually explicit image generated from a non-intimate photograph that was publicly available online. The offender threatened to distribute the photo to the young person's social media contacts unless the victim complied with their demands for money.

⁷ Hanna McCallum, "Netsafe Sees Surge in Reports of Sexual Deepfake Abuse," Newsroom, June 10, 2026, <https://newsroom.co.nz/2026/06/10/netsafe-sees-surge-in-reports-of-sexual-deepfake-abuse>

⁸ Amrith Ramkumar, "AI Deepfakes Are Getting Weirder and Harder to Spot in the Midterms," *The Wall Street Journal*, June 15, 2026, <https://www.wsj.com/politics/elections/ai-deepfakes-are-getting-weirder-and-harder-to-spot-in-the-midterms-88b4f7ad>

- **Impersonation and reputational harm:** Netsafe is encountering more reports involving AI-enabled impersonation. In one recent case, Netsafe assisted a school principal after an offender created a fake social media profile using AI-generated and manipulated images that falsely depicted and impersonated them, causing concern for their professional reputation and the wider school community.

These examples demonstrate why legislative clarification is necessary. From the victim's perspective, the harm does not depend on whether an image is authentic or synthetic. Netsafe has observed these harms emerging for several years and expects them to become increasingly common as generative AI tools become more accessible.

Deepfake intimate abuse disproportionately harms women and girls

We urge the Committee to recognise that sexually explicit deepfakes are not a novelty or fringe phenomenon. They are a scalable form of image-based sexual abuse. International evidence consistently shows that the dominant use of deepfake technology is the creation of pornographic or sexualised content and that women and girls are disproportionately targeted.⁹ Research commonly estimates that around 90 to 95 percent of identified deepfake content is non-consensual pornography and that the overwhelming majority depicts women.¹⁰

This form of online abuse represents a technology-enabled extension of violence against women and girls. It can have serious negative consequences for dignity, safety, participation in public life, employment opportunities, relationships and mental wellbeing.¹¹ This Bill is consistent with international approaches that treat sexually explicit deepfakes as a form of image-based sexual abuse.¹² The focus should remain on the harm and victim impact rather than the technology used to create the content.

Targeted drafting improvement

We support the Bill's intent and approach but recommend a targeted amendment to avoid an unintended narrowing of scope.

⁹ Cassandra Mudgway, "NZ Is Criminalising Sexualised Deepfakes – Banning Apps That Make Them Should Be Next," The Conversation, May 19, 2026, <https://theconversation.com/nz-is-criminalising-sexualised-deepfakes-banning-apps-that-make-them-should-be-next-282127>

¹⁰ Laura Barrón-López, "Women Face New Sexual Harassment with Deepfake Pornography," *PBS NewsHour*, June 27, 2023, <https://www.pbs.org/newshour/show/women-face-new-sexual-harassment-with-deepfake-pornography>

¹¹ VINE, "New Evidence Brief on the Escalation of Online Violence Against Women in the Public Sphere."

¹² Barrón-López, "Women Face New Sexual Harassment with Deepfake Pornography."

The Bill inserts a new limb into the definition of "intimate visual recording" that applies where a visual recording has been created, synthesised, or altered "without the knowledge or consent" of the person who is the subject of the recording. Netsafe is concerned that this departs from the existing drafting approach and may create uncertainty as to whether Parliament intended synthetic intimate imagery to be treated differently from other forms of intimate visual recording.

The proposed wording may invite arguments that synthetic intimate imagery falls outside the definition where a person had knowledge of, or previously consented to, an underlying image or recording, but did not consent to its subsequent synthetic alteration, intimate depiction, use, or dissemination.

This outcome would sit uneasily with the broader structure of the HDCA. Section 22A focuses on whether the victim consented to the posting of an intimate visual recording, not whether they knew of, or consented to, the existence of an image at an earlier point in time. Netsafe sees no policy reason for synthetic intimate imagery to be subject to a narrower threshold than other intimate visual recordings captured by the definition.

Current position under the HDCA:

Section 4 of the HDCA defines an "intimate visual recording" as a visual recording made "with or without the knowledge or consent" of the individual who is the subject of the recording. Parliament therefore adopted a deliberately broad definition that does not depend on whether the person knew of, or consented to, the creation of the recording.

Recommendation:

To maintain consistency with the existing definition and avoid unnecessary technical arguments about the scope of the provision, Netsafe recommends that the synthetic-media limb adopt the same "with or without the knowledge or consent" formulation used elsewhere in the definition of "intimate visual recording".

Suggested drafting approach:

In section 4, definition of intimate visual recording, after paragraph (a) insert:

*(ab) "includes a visual recording that has been created, synthesised, or altered (**whether with or without the knowledge or consent of the person who is the subject of the recording**) and appears to show the person ..."*

Future State Observations and Next Steps

This Bill is an important example of technology-neutral law reform.

The Bill focuses on the correct question: the harm experienced by victims, not the technology used to cause it. That approach reflects the broader purpose of the HDCA. It also raises a wider question for Parliament about whether New Zealand's online safety framework remains fit for purpose when harmful content can be generated, modified, replicated and redistributed at near-zero cost.¹³

Future reform should consider the growing importance of identity-based harms. Much of the HDCA focuses on the content of a communication and the harm it causes. Synthetic media often causes harm because someone appropriates or manipulates a person's likeness, voice, reputation or identity. Deepfake pornography, cloned voices, fabricated screenshots and AI-enabled impersonation all share that feature.¹⁴

Future reform should also consider whether existing remedies under the HDCA actually remain effective for synthetic media. The HDCA was developed in a context where harmful content could often be identified and removed as a discrete communication. Synthetic media operates differently. Once created, harmful content can be edited, regenerated and redistributed across multiple services and platforms. This may require remedies that can target substantially similar harmful content at a widespread scale.

Parliament may also wish to consider whether future reforms should focus on whether content appears to depict an identifiable person, rather than how it was created.¹⁵ For victims, the harm often arises because others believe the content is real and depicts them. As synthetic media becomes more sophisticated, this may become more important than the technical process used to generate the content.

None of these issues require amendment through the present Bill. The Bill is strongest as a targeted clarification that intimate-image protections apply equally to authentic and synthetic content. It also gives

¹³ Science Media Centre, "Deepfake and the Law – Expert Reaction," *Science Media Centre*, May 21, 2019, <https://www.sciencemediacentre.co.nz/2019/05/21/deepfake-and-the-law-expert-reaction/>.

¹⁴ Evans, "Deepfakes and Privacy: It's Time to Respond"

¹⁵ Cassandra Mudgway and Andrew Lensen, "Sexualised Deepfakes on X Are a Sign of Things to Come. NZ Law Is Already Way Behind," *The Conversation*, January 20, 2026, <https://theconversation.com/sexualised-deepfakes-on-x-are-a-sign-of-things-to-come-nz-law-is-already-way-behind-273562>

Parliament an opportunity to recognise that synthetic media will place increasing pressure on other parts of the HDCA. Future reform should continue to focus on consent, identity, harm and effective redress rather than any particular generation of technology.

Conclusion

Netsafe strongly supports the Bill and encourages the Committee to recommend its passage.

The Bill addresses a genuine and growing form of harm that Netsafe has encountered for many years and is seeing with increasing frequency through its frontline services. It provides important legislative certainty that New Zealand's intimate-image protections apply equally to synthetic and AI-generated content.

We also recommend one targeted drafting amendment to ensure the proposed synthetic-media provision does not inadvertently narrow the existing definition of "intimate visual recording" and remains consistent with the wider structure and purpose of the HDCA.

Netsafe

17 June 2026