

Scams and abuse

Key points

- AI is making existing crimes easier to commit and harder to detect.
- Criminals are using cloned voices and deepfaked public figures to defraud New Zealanders.
- AI is used to generate sexual images of children and adults, and to groom and harass them.
- Our laws and support services have not kept up with the rapid pace of change.

AI has lowered the cost and skill needed to commit fraud and abuse. Producing a convincing fake image, video or voice recording once took expertise and time; it is now quick, cheap and widely available. This brief sets out two categories of harm already affecting New Zealanders and what the government can do.

Scams

AI is being used to impersonate people and organisations that victims trust, and to run more scams with fewer mistakes.

- A person's voice can be copied from a few seconds of audio and used to [impersonate a relative in distress](#) and defraud the person who answers.
- The Financial Markets Authority has [warned of fake news articles](#) using deepfaked images of New Zealand public figures directing people to fraudulent trading platforms. It found 110 such ads on Meta sites in one 24-hour period.
- In 2024, a Taranaki woman [lost \\$224,000](#) after a deepfake video of the Prime Minister urged superannuitants to invest in cryptocurrency.
- Phishing that was once given away by poor spelling now arrives clean, and some campaigns [target New Zealand businesses in te reo Māori](#).
- The National Cyber Security Centre recorded [\\$5.6 million in direct financial losses](#) in the first quarter of 2026, up 76 percent on the previous quarter.
- Public agencies are impersonated as well. Fake letters, calls and websites presented as coming from a government department are cheap to produce.
- A [deepfake image of Westpac's chief executive appeared in scam advertising](#). The bank contacted Meta and received no reply, while altered versions continued appearing for weeks.

Abuse

AI is being used to fabricate sexual images of people who did not consent, and to groom and harass them through chatbots, with harms falling mainly on women and girls.

- One widely cited study found [98 percent of deepfake videos online](#) were pornographic and mostly targeted women.
- AI tools are used to create child sexual abuse material and to "nudify" photos of real children. The Internet Watch Foundation [identified over 8,000 AI-generated images and videos](#) as showing realistic child sexual abuse in 2025.

- A study across 11 countries found [more than 1.2 million children](#) reported having their images turned into sexually explicit deepfakes in a single year, overwhelmingly girls.
- Netsafe received [667 reports of sextortion threats](#) in the first three months of 2025, up 68 percent on the year before. The youngest victim was nine.
- Australia's online safety regulator has [warned that AI companions](#) expose children to sexualised conversation, which undermines young people's ability to identify unsafe interactions and age-inappropriate behaviour, increasing the risk of sexual abuse.
- Researchers report chatbots [initiating unwanted sexual conversations](#) with users, including children, and giving personalised guidance to people stalking or harassing others.

The policy gap

New Zealand has [no criminal offence covering sexualised deepfakes](#). The offence of sharing an intimate visual recording without consent applies only to genuine recordings. Prosecutors can use the offence of "causing harm by posting a digital communication", but this requires proving the offender intended serious emotional distress.¹

The Deepfake Digital Harm and Exploitation Bill, now before Parliament, would address this by widening the definition of "intimate visual recording" in the Crimes Act 1961 and the Harmful Digital Communications Act 2015 to include synthetic images. However, as drafted, the Bill leaves many related issues unaddressed. For example:

- The Bill [contains no mechanism](#) to regulate the spread of synthetic image generation technology or hold to account the platforms on which content is generated or shared.
- Deepfake scams fall outside the Bill's scope.
- Netsafe, the approved agency under the Harmful Digital Communications Act, receives over 20,000 reports of online harm a year, but it negotiates removal rather than investigating and cannot compel a platform to act.
- Most offending is never reported. The Classification Office found that [two-thirds of New Zealand adults](#) have encountered extreme or potentially illegal content online and that it is significantly under-reported.
- Even when harmful media is reported, tracing whoever made it can be impossible.

The Bill's sponsor Laura McClure has said the Bill is only [a first step](#) and that MPs should also consider a full review of relevant laws, including the Harmful Digital Communications Act.

What New Zealand can do

The government should act now. Priorities include:

- Passing the [Deepfake Digital Harm and Exploitation Bill](#), which would make creating and sharing altered or synthetic intimate visual recordings a criminal offence.
- A [full review of relevant laws](#), including the Harmful Digital Communications Act. The scope should cover deepfake scams, regulating access to synthetic media tools, and holding social media companies accountable for criminal content shared on their platforms.
- Increasing funding for Netsafe, Police and the National Cyber Security Centre to match the volume of reports they now receive.

Our accompanying [policy papers](#) set out these and other steps in detail.

¹ New Zealand's [first known deepfake conviction](#), in 2025, succeeded only because the victim's family traced the offender and he confessed.