What to do if you or someone you know is targeted with deepfake porn or AI nudes

June 12 2024, by Nicola Henry

Credit: Pixabay/CC0 Public Domain

This week, about 50 female students from Victoria's Bacchus Marsh Grammar School found fake, sexually explicit images of themselves shared without their consent on Instagram and Snapchat. Images of their faces, purportedly obtained from social media, were stitched onto pornographic images using artificial intelligence (AI).

Deepfake porn, or what our team calls "AI-generated image-based sexual abuse," involves the use of AI to create a nude and/or sexual image of a person doing or saying things they haven't said or done.

Celebrities and public figures, predominantly women, have experienced such abuse for nearly a decade, with various deepfake porn sites and "nudify apps" readily available online.

But as these technologies become more accessible and sophisticated, we're starting to see this problem creep into our homes and schools. Teens—and even children—are now being targeted.

How widespread is deepfake abuse?

In 2023, my colleagues and I surveyed more than 16,000 adults in ten countries and found that, despite widespread media coverage (particularly in Western countries), the concept of deepfake porn isn't well known. When informed about it, however, most respondents indicated it should be criminalized.

Among respondents from Australia, 3.7% had been a victim of deepfake porn as an adult. This was the highest rate reported from the countries we surveyed.

At the same time, 2.4% of Australian respondents said they had created,

shared or threatened to share a deepfake photo or video of another person without their consent. This too was a higher figure than every other country we surveyed except the United States.

Men were more likely to report being a victim of deepfake abuse, and more likely to report being a perpetrator. Men were also less likely to find the viewing, creating and/or sharing of deepfake pornography to be problematic.

What can you do if you're targeted?

Image-based abuse can be a distressing experience. But victims should know they're not alone, it isn't their fault and there is plenty of help out there. Here are some steps they can take.

1. Report it

Creating or sharing deepfake sexual images of minors is a criminal offense under Australia's federal child sexual abuse material ("child pornography") laws. It's also a criminal offense to share non-consensual deepfake porn of an adult (and a crime to create it if you're in Victoria).

Whether you're the victim, or someone you know is, you can report deepfake abuse to digital platforms, to the Australian Centre to Counter Child Exploitation (if the person depicted is a minor) and to the eSafety Commissioner.

If you're in danger, contact the police or ambulance on triple zero (000). If it's not an emergency, you can call the Police Assistance Line (131 444) or your local police station. The same steps apply if you're a bystander who has come across non-consensual deepfake pornography of someone else online.

The eSafety commissioner can take action against image-based abuse under the federal Online Safety Act, and can work with victims and their supporters to get the content taken down within 24 hours. They can also issue formal warnings, take-down orders and civil penalties to individuals and technology companies that fail to take action.

Unfortunately, the deepfake content may continue to circulate even after it is taken down from the initial platform.

2. Seek help

If you've been targeted, it's a good idea to talk to someone you trust, such as a friend, family member, teacher, counselor or psychologist.

Our website has a list of relevant support services for victim-survivors of image-based abuse, including specialist services for Aboriginal and Torres Strait Islander people, migrants and refugees, young people, people with disabilities, people from LGBTQI+ communities and sex workers.

Even if you're not ready to talk about the experience, you can still find useful information about image-based abuse online, including on the eSafety commissioner's website.

We've also developed a chatbot called Umibot, which provides free confidential advice and support to people who have experienced image-based abuse, including deepfake abuse. Umibot also has information for bystanders and perpetrators.

If you're Aboriginal or Torres Strait Islander, you can check out WellMob. This is an online resource made by Indigenous Australians to provide information on social and emotional well-being.

Resources for young people are also available from ReachOut, Beyond Blue, Youth Law Australia and Kids Helpline.

3. Create a digital hash to stop the spread

The United Kingdom's Revenge Porn Helpline and Meta have developed two digital hashing tools for victim-survivors. These are Stop NCII for adults, and Take It Down for minors.

Anyone in the world can use these tools to generate an anonymous digital hash (a unique numerical code) by scanning the image from their device. This hash is then shared with the companies participating in the scheme (including Facebook, Instagram, Pornhub, TikTok and OnlyFans) so they may detect and block any matches on their platform. You aren't required to upload the image, which means no one else sees it, nor does it leave your device.

It's important to note this tool won't block the image from appearing on platforms that aren't part of the scheme. You also need to have access to the images in the first place to use the tool.

4. Block, report and distance yourself from the perpetrator (if it's safe to do so)

You can block the perpetrator(s) through your mobile and on social media, and report them to the relevant platforms and authorities. In the case of platforms, it's not always clear what will be done once a report is lodged, so it's a good idea to ask about this.

If the perpetrator is someone you know, such as a classmate or student, authorities can take action to ensure you don't interact with that person anymore.

Last week, a boy was expelled from Melbourne's Salesian College after he used AI to create sexually explicit images of a female teacher.

5. Boost your online safety

The eSafety commissioner has step-by-step video guides on a range of online safety topics, from how to change your privacy settings on social media, to how to choose strong passwords.

For women experiencing family or domestic violence, the following resources may also be helpful:

- WESNET's comprehensive safety and privacy toolkit for women
- The Australian eSafety Commissioner's online safety checklist
- 1800 RESPECT's Device Safety page.

Provided by The Conversation