POLICY OPTIONS ON NON-CONSENSUAL DEEPNUDES AND SEXUAL DEEPFAKES

Advancements in technology create the potential for new forms of gender-based violence that Canada needs to recognize, address, and monitor. Two of these new forms of violence are known as non-consensual deepnudes and sexual deepfakes. This Brief shares information on non-consensual deepnudes and sexual deepfakes as gendered forms of sexual violence. It explores the impacts this violence has on survivors, and what Canadian policy can do to address it.

Non-consensual deepnudes and sexual deepfakes are media that use artificial intelligence to digitally insert an individual's image into sexual videos (deepfakes) and photos (deepnudes) without their consent.

Creating a non-consensual deepnude or sexual deepfake requires only that an individual have access to another person's photos and videos which may be found through videos from professional speaking opportunities, posts on social media platforms, photos taken by an intimate partner, and more.
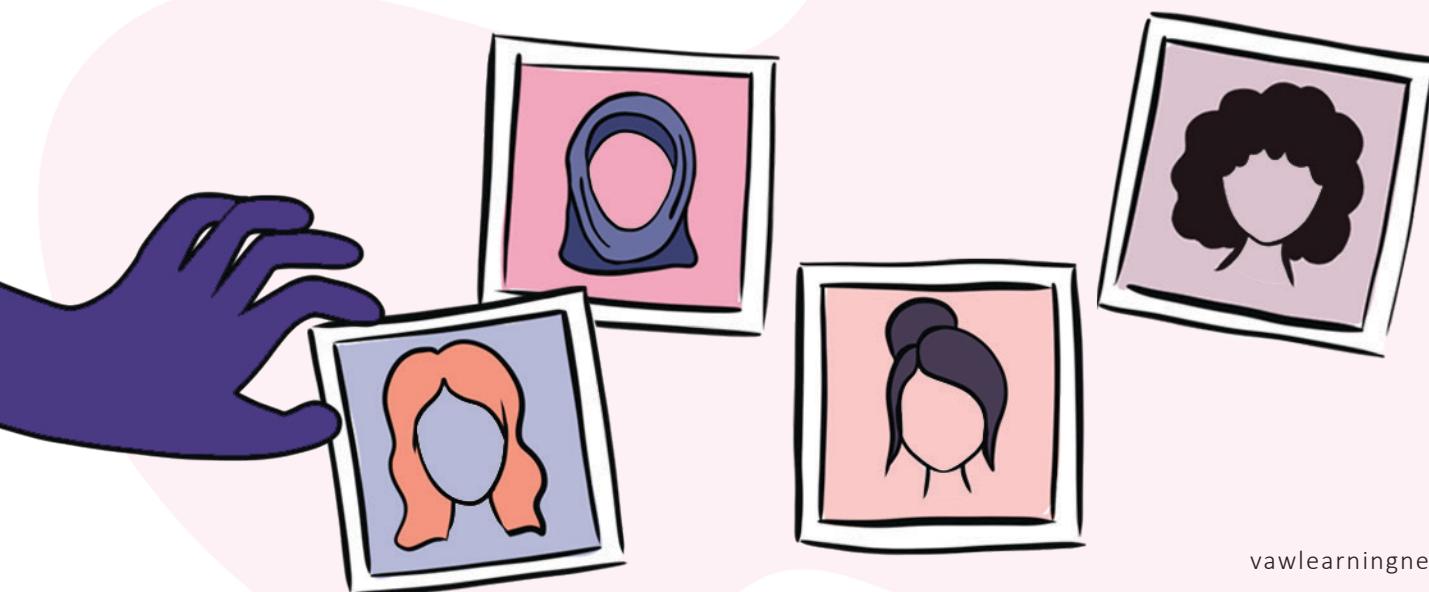
Deepnudes and sexual deepfakes emerge from past forms of image manipulation with early examples being the use of Photoshop or photo editing technology.[1] The tools to manipulate images has since advanced, making these forgeries more convincing. The technology has also become more widely accessible, often only requiring a phone or computer.

The number of sexual deepfake videos is doubling every six months and estimates suggest that by summer 2022 there could be 720,000 sexual deepfakes online.[2]

LANGUAGE MATTERS

Colloquially, non-consensual deepnudes and sexual deepfakes are sometimes referred to as "revenge porn." However, "revenge porn" is an inaccurate and harmful term.
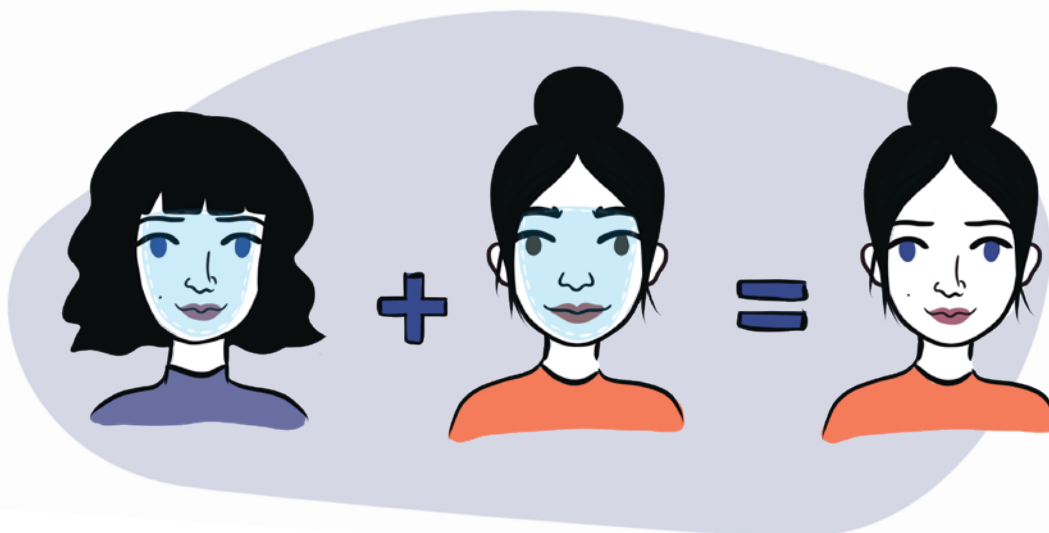Use of the term "revenge" is harmful because it implies that the person did something wrong that warrants or justifies this violence, but no one deserves violence. Non-consensual deepnudes and sexual deepfakes are also made for a variety of motivations other than revenge, including maintaining power and control in an intimate relationship. Likewise, the use of porn in this context conflates the difference between consensual pornography and the non-consensual creation of sexual images and videos.

vawlearningnetwork.ca

WHAT ARE DEEPNUDES AND SEXUAL DEEPFAKES?

Deepnudes and sexual deepfakes use a form of machine learning called deep learning that has been developing since the 1950s and has benefited from significant technological advancements in the last few years. Deep learning can "grind through gigantic volumes of data, teaching itself how to detect and classify patterns or anomalies and make predictions and recommendations."[3]

To illustrate how this process works, consider teaching a machine to recognize a person. Before deep learning, it would be necessary to tell the machine what features are characteristic of the person, such as green eyes and brown hair. With deep learning, the machine need only be shown images of the person and it can extract from those images the features of the person without the need for input. Deep learning is already common in everyday life, for instance it is used by Google in facial recognition (identifying the same person in multiple photos without input). In the context of deepnudes and sexual deepfakes, deep learning is used to face swap individuals from original content into sexual images or videos. What deep learning does is assess if the new manipulated image looks realistic by comparing the fake to the real material and getting as close as possible to the person's likeness.

Sometimes images and videos used for non-consensual deepnudes and sexual deepfakes may be publicly available through, for instance, social media accounts. While the COVID-19 pandemic has forced more people to go online with increased use of image and videos, typically it has been younger Canadians aged 15-19 who are most likely to share social media content publicly.[4] Other times non-consensual deepnudes and sexual deepfakes may be created by those who hold non-publicly available images (e.g. family, partner).

Some non-consensual deepnudes and sexual deepfakes are made by individuals acting alone, while others involve an individual paying either another individual or an application to create the content.

CONSIDER THESE POSSIBLE USES:

Sue leaves her abusive relationship and goes to a shelter. After, her partner uses personal videos of Sue to make a sexual deepfake and threatens to send the deepfake to Sue's family if she does not return.

Donovan receives an email that a deepnude of him exists. The original picture was taken from his Facebook. The emailer threatens to send the deepnude to Donovan's employers unless Donovan sends $1,000.

Jalen is an outspoken activist who posts videos on Tik Tok. She starts to receive aggressive messages on her phone, and one has a link. A sexual deepfake of her has been shared online with her address and phone number.

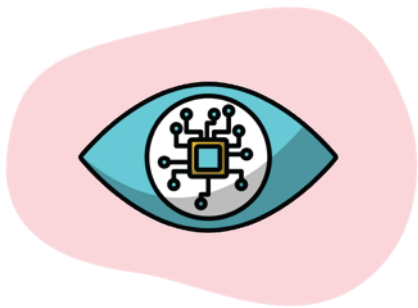# DEEPNUDES AND SEXUAL DEEPFAKES ARE GENDERED FORMS OF VIOLENCE.

Similar to other forms of sexual violence, non-consensual deepnudes and sexual deepfakes involve a violation of consent and autonomy. Specifically, deepnudes and sexual deepfakes deny individuals' sexual privacy.

Non-consensual deepnudes and sexual deepfakes are unique from non-digital forms of sexual violence in that they are distributed through public channels on the Internet and private networks (like WhatsApp). This online sharing makes the content extremely difficult to detect and remove as sharing becomes crowdsourced so that many people are involved in continually spreading the sexually objectifying content.

Sexual privacy is "the behaviors, expectations, and choices that manage access to and information about the human body, sex, sexuality, gender, and intimate activities."[5]

DEEPNUDES

Deepnudes are named after an app entitled DeepNude that enabled users to upload images of clothed women that would then be 'stripped' of their clothing as the app would match a woman's face to a nude body. While the original application was removed, many versions of it subsist online, including a popular version on messaging application Telegram.
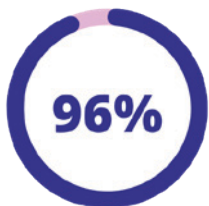

Currently, deepnude applications only work on images of women.

Approximately 104,852 women have had deepnudes created and posted on Telegram as of the end of July 2020.[6]

SEXUAL DEEPFAKES

The first sexual deepfake was created in 2016 and involved the non-consensual placement of a popular actress into a sexual video. The very name "deepfakes" comes from a Reddit user that made this original video and the majority of deepfakes continue to involve placing celebrities into sexual videos.7 The software used to produce deepfakes has been openly shared on the Internet.

A study of 14,678 deepfake videos between January 2019 to July 2019 found that:[8]

96% of deepfakes featured sexual content

100% of those sexual deepfakes featured women

The sexual images and videos used to create non-consensual deepnudes and sexual deepfakes often come from pornography, specifically from women adult performers. Adult performers describe the experience of having their videos used for sexual deepfakes as "violating" and "eviscerating."[9]

SURVIVORS FACE ONGOING IMPACTS FROM THIS VIOLENCE.

Research on the harms of non-consensual deepnudes and sexual deepfakes is limited at this time due to its emerging nature but we can learn from survivors and from the impacts of similar forms of violence like online harassment and the non-consensual distribution of intimate images.

TRAUMA

An investigative journalist in India, Rana Ayyub, was targeted in a sexual deepfake after she criticized the response of a political party to the gang rape of an 8-year-old girl.[10] Consistent with the practice of doxing, the posting of the sexual deepfake targeting Ayyub included personal information about her (e.g. phone number). Ayyub details how the experience led to continuous sexualized comments through social media and a trip to the hospital as a result of a violent reaction to the stress. She deleted her social media and was isolated from her friends and family. Ayyub's experience is consistent with the experiences and trauma of survivors who share their distress, anxiety, suicidal ideation, shame, and social isolation after their intimate images were non-consensually shared online.[11]

ECONOMIC HARM

Non-consensual deepnudes and sexual deepfakes pose a threat for the job security and economic opportunities of those targeted. For instance, one prominent YouTuber lost their brand partnership after a sexual deepfake of her was posed on a pornography website.[12] Economic harms are exacerbated by oppressive social scripts related to gender and sexuality which increase the risk for economic harm when sexual content of the individual is spread online. For instance, women are seen as immoral for sexual activity but the same stigma does not exist for men.[13] LBGTQI2S individuals may also be targeted with non-consensual deepnudes and sexual deepfakes that not only spread sexual content but threaten to "out" the individual.

SILENCING

Research has shown that online abuse on social media platforms like Twitter is disproportionally experienced by women, especially racialized women, and serves to discourage women from participating online.[14] Likewise, non-consensual deepnudes and sexual deepfakes have a silencing effect whereby survivors may stop engaging online, and in general, women may feel less safe and able to post online. To date, non-consensual deepnudes and sexual deepfakes have targeted women activists and journalists who espouse feminist and anti-oppressive views.[15]

SEXUAL OBJECTIFICATION

Sexual objectification reduces individuals into physical objects. Non-consensual deepnudes and sexual deepfakes are a form of sexual objectification and part of a "broader environment in which women's images are understood as consumable, malleable, and brought into being for the enjoyment and gratification of men."[16] Sexual objectification leads to diminished views regarding women's competence, morality, and humanity or personhood.[17] It has also been linked to greater tolerance of sexual violence towards women.[18]

> *Deepfakes are "used as a weapon to silence women, degrade women, show power over women, reducing us to sex objects. This isn't just a fun-and-games thing. This can destroy lives."*
> *— Anita Sarkeesian, Canadian-American media critic who was targeted with sexual deepfakes[19]*

LEARNING FROM LIVED EXPERIENCE

Click here for a video by Noelle Martin where she shares her own experience of having images of her taken and used in the production of fake sexual materials. Noelle used her experience to help change Australian law so that it better supported survivors.

5 WAYS TO USE POLICY TO PREVENT AND RESPOND TO NON-CONSENSUAL DEEPNUDES AND SEXUAL DEEPFAKES IN CANADA

1. Criminalize the production and distribution of non-consensual deepnudes and sexual deepfakes.

Currently, Canada has no law criminalizing non-consensual deepnudes and sexual deepfakes. There are other legal responses that individuals may be able to utilize like defamation, appropriation of personality, and Canadian Elections Act laws, depending on the context.[20] Non-consensual distribution of intimate images is also criminalized in Canadian law but the definition of intimate images does not explicitly include deepnudes and sexual deepfakes. In order to simplify the legal process for survivors and affirm that non-consensual deepnudes and sexual deepfakes are a form of violence, the creation and distribution of non-consensual deepnudes and sexual deepfakes should be criminalized in Canadian law. Criminalization of non-consensual deepnudes and sexual deepfakes will better enable the attribution of liability since law enforcement agencies have greater investigative capacities than private individuals or lawyers.[21] Such capacity is especially important since non-consensual deepnudes and sexual deepfakes are often posted anonymously. In addition, criminalization would serve as a stronger deterrent. Danielle Citron, a University of Maryland law professor and expert on sexual deepfakes, argues for comprehensive criminal statues on deepfakes by noting that:

"We need real deterrents, otherwise, it's just a game of whack-a-mole."[22]

2. Conduct Intersectional and gender-based research on non-consensual deepnudes and sexual deepfakes.

Non-consensual deepnudes and sexual deepfakes are an emerging and evolving issue that require further monitoring and examination. Currently, concerns about deepfakes primarily focus on their potential use as forms of political misinformation and disinformation, as opposed to their most common form as sexual violence.[23] There is a need for quantitative, qualitative, and theoretical research on the gendered and intersectional dimensions of non-consensual deepnudes and sexual deepfakes.

3. Work with media companies to detect and remove non-consensual deepnudes and sexual deepfakes.

Government should work with website and social media companies to collaboratively find ways to prevent, detect, and remove non-consensual deepnudes and sexual deepfakes. The United States, for instance, introduced a bill entitled the "Deep Fake Detection Prize Competition Act" directing the National Science Foundation to carry out prize competitions to detect deepfakes.[24] Facebook held its own Deep Detection Challenge as well and shared the results.[25] In addition to detection, companies could collaborate to make transparent, shared policies prohibiting non-consensual deepnudes and sexual deepfakes on their platforms in their terms-of-service agreements.

Canadian law does have the option to potentially hold online providers of interactive websites and applications liable for acts of their users (such as posting deepnudes and sexual deepfakes). However, this option is complicated by the fact that many companies hosting non-consensual deepnudes and sexual deepfakes may not be Canadian and other countries, such as the United States, do not hold companies liable for user content.[26]

4. Fund awareness and prevention campaigns on non-consensual deepnudes and sexual deepfakes as forms of violence.

The conversation on non-consensual deepnudes and sexual deepfakes is at times focused on the necessity of not sharing images and videos online. This perpetuates the harmful narrative that this violence is the fault of the individual for posting online. Rather than blaming survivors, we need to hold individuals perpetrating such violence accountable and to challenge the normalization of technology-facilitated gender-based violence.

Innovative and evidence-based prevention campaigns should be funded focused on raising awareness about non-consensual deepnudes and sexual deepfakes as forms of violence. These campaigns should be developed and implemented by diverse communities.

5. Create national initiatives with survivors to support survivors.

Other states have already implemented national initiatives to support survivors of non-consensual deepnudes and sexual deepfakes like a Helpline in the United Kingdom, the Cyber Civil Rights Initiative Helpline in the United States, and the Office of eSafety in Australia. These initiatives provide survivors and allies with information on cyber rights and how to report violence, in addition to more general chat and phone support. In Canada, a similar initiative called Cybertip is available but only for those under 18.

Canada needs an equivalent, national initiative available to all ages that is funded and staffed with trained personnel to provide support and information to those seeking assistance and to assume a role in research and education on deepnudes and sexual deepfakes in this country. Such an initiative could assist the important work of gender-based violence organizations in Canada that continue to spread awareness and support survivors.

If you need support, please reach out.

Let us know what you think. Your input is important to us.
Please click here to complete this short survey on your thoughts of this Brief.

Suggested Citation: Lalonde, D. (2021). Policy Options on Non-Consensual Deepnudes and Sexual Deepfakes. *Learning Network Brief 39*. London, Ontario: Learning Network, Centre for Research & Education on Violence Against Women & Children. ISBN: 978-1-988412-49-8.

Graphic design by: Emily Kumpf

REFERENCES

1   Paris, B. & Donovan, J. (2019). Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence. *Data & Society*. https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal.pdf

2   Schick, N. (2020, Dec 20). Deepfakes are jumping from porn to politics. It's time to fight back. *Wired.* https://www.wired.co.uk/article/deepfakes-porn-politics

3   Tibbetts, J. (2018). The frontiers of artificial intelligence: Deep learning brings speed, accuracy to the life sciences. *Bioscience*, 68(1), 5–10, p. 5. https://doi.org/10.1093/biosci/bix136

4   Schimmele, C., Fonberg, J., & Schellenberg, G. (2021, March 24). Canadians' assessments of social media in their lives. *Statistics Canada*. https://www150.statcan.gc.ca/n1/pub/36-28-0001/2021003/article/00004-eng.htm

5   Citron, D. K. (2019). Sexual privacy. *The Yale Law Journal*, 128(7), 1870–1960, p. 1870. https://scholarship.law.bu.edu/faculty_scholarship/620

6   Ajder, H., Patrini, G., & Cavalli, F. (2020). Automating image abuse: Deepfake bots on Telegram. *Sensity.* https://sensity.ai/reports/

7   Ajder, H., Patrini, G., Cavalli, F., & Cullen, L. (2019). The state of deepfakes: Landscape, threats, and impact. *Sensity.* https://sensity.ai/reports/

8   Ajder, H., Patrini, G., Cavalli, F., & Cullen, L. (2019). The state of deepfakes: Landscape, threats, and impact. *Sensity.* https://sensity.ai/reports/

9   Alptraum, L. (2020, Jan 15). Deepfake porn harms adult performers, too. *Wired*. https://www.wired.com/story/deepfake-porn-harms-adult-performers-too/

10   Ayyub, R. (2018, Nov 23). I was the victim of a deepfake porn plot intended to silence me. *Huffpost*. https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316

11   Eaton, A.A., Jacobs, H., & Ruvalcaba, Y. (2017). 2017 Nationwide Online Study of Nonconsensual Porn Victimization and Perpetration: A Summary Report. *Cyber Civil Rights Initiative, Inc.* https://www.cybercivilrights.org/wp-content/uploads/2017/06/CCRI-2017-Research-Report.pdf

12   Dickson, E.J. (2020, Oct 26). Tiktok stars are being turned into deepfake porn without their consent. *Rolling Stone*. https://www.rollingstone.com/culture/culture-features/tiktok-creators-deepfake-pornography-discord-pornhub-1078859/

13   Citron, D., & Franks, M. (2014). Criminalizing revenge porn. *Wake Forest Law Review,* 49(2), 345–391. https://digitalcommons.law.umaryland.edu/fac_pubs/1420/

14   Amnesty International. (2018). #ToxicTwitter: Violence and abuse against women online. *Amnesty International*. https://www.amnesty.ca/sites/amnesty/files/%23TOXICTWITTER%20report%20EMBARGOED.pdf;

15   Harwell, D. (2018, Dec 30). Fake-porn videos are being weaponized to harass and humiliate women. *The Washington Post*. https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/?utm_term=.f0f347a45ba3

16   van der Nagel, E. (2020). Verifying images: deepfakes, control, and consent. *Porn Studies (Abingdon, UK)*, 7(4), 424–429, p. 426. https://doi.org/10.1080/23268743.2020.1741434

17   Ward, L. (2016). Media and sexualization: State of empirical research, 1995-2015. *The Journal of Sex Research*, 53(4-5), 560–577. https://doi.org/10.1080/00224499.2016.1142496; Loughnan, S., Haslam, N., Murnane, T., Vaes, J., Reynolds, C., & Suitner, C. (2010). The denial of mind and moral concern to objectified others. *European Journal of Social Psychology*, 40(5), 709–717. https://doi.org/10.1002/ejsp.755

18   Seabrook, R., Ward, L., & Giaccardi, S. (2019). Less than human? Media use, objectification of women, and men's acceptance of sexual aggression. *Psychology of Violence*, 9(5), 536–545. https://doi.org/10.1037/vio0000198

19   Harwell, D. (2018, Dec 30). Fake-porn videos are being weaponized to harass and humiliate women. *The Washington Post.* https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/?utm_term=.f0f347a45ba3

20   Siekierski, B. J. (2019, April 8). Deep fakes: What can be done about synthetic audio and video? *Library of Parliament*, Publication No. 2019-11-E. https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/201911E

21   Khodayari, A. (2020, Sept 14). Regulation of deepfakes in Canada: Is criminal liability the right answer? *Tech Law McGill Blog.* https://www.techlawmcgill.com/post/regulation-of-deepfakes-in-canada-is-criminal-liability-the-right-answer

22   Harwell, D. (2018, Dec 30). Fake-porn videos are being weaponized to harass and humiliate women. *The Washington Post.* https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/?utm_term=.f0f347a45ba3

23   Gosse, C., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication*, *37*(5), 497–511. https://doi.org/10.1080/15295036.2020.1832697

24   United States Congress. (2020). H.R.5532 - Deep fake detection prize competition act. https://www.congress.gov/bill/116th-congress/house-bill/5532/text?format=txt

25   Facebook AI. (2020, June 12). Deepfake detection challenge results: An open initiative to advance AI. *Facebook AI Blog.* https://ai.facebook.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/

26   Tseng, P. (2018, March). What can the law do about 'deepfake'? *McMillan Litigation and Intellectual Property Bulletin*. https://mcmillan.ca/insights/what-can-the-law-do-about-deepfake/