Loyola Law School
Loyola Marymount University
Los Angeles

919 Albany Street
Los Angeles, California
90015

T 213.736.1028
C 818.648.8137
F 213.736.1001
www.lls.edu

SUBMISSION OF PROFESSOR REBECCA DELFINO ON
CRIMINAL CODE AMENDMENT (DEEPFAKE SEXUAL MATERIAL) BILL 2024,
To
Legal And Constitutional Affairs Legislation Committee
July 18, 2024

I am Rebecca A. Delfino, a Law Professor at LMU Loyola Law School, Los Angeles, California. I have completed a substantial legal scholarship about artificial intelligence and its impact on society. The following perspective on the proposed legislation, Criminal Code Amendment (Deepfake Sexual Material) Bill 2024, is offered at the request of the Legal and Constitutional Affairs Legislation Committee. It is based on personal knowledge and/or information and belief derived from my five years of academic and legal research and my legal expertise on deepfakes and nonconsensual pornographic deepfakes. I present this testimony under the penalty under the laws of California and the United States:

I applaud the Australian parliament for its efforts to criminalize non-consensual pornographic deepfakes. Through the work of the Australian Communications and Media Authority and the Broadcasting Services Act, Australia is a world leader in protecting the most vulnerable from the harms unleashed by the internet. With that in mind, I offer the following commentary on the proposed amendments to the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024.

A. Preliminary Considerations on the impact of deepfakes on the victims and the need for specific criminalization:

More than 96% of deepfakes target women. Sensity AI, a cybersecurity and research company that has tracked online deepfake videos since December 2018, has consistently found that between 90% and 95% of deepfakes are nonconsensual porn. Not surprisingly, the primary victims' images and likenesses are used without their consent and often without their knowledge. According to Sensity, deepfakes are growing exponentially, doubling every six months.

Pornographic deepfakes have the potential to affect all women, not only celebrities, politicians, or others who have an intentional presence online. In October 2020, further research from Sensity found that a new AI bot created and shared fake nude images of

more than 680,000 women without their knowledge, often requiring only one photo of the person to generate the fake image. In addition, this trend is even more worrying in the context of the pandemic. A deepfake helpline in the UK's caseload has nearly doubled since the pandemic started. Existing abusive relationships have worsened, and digital abuse has seen an uptick as people have grown increasingly isolated and spend more time online.

According to a report published in Wired Magazine in the fall of 2023: At least 244,625 videos have been uploaded to the top 35 websites set up either exclusively or partially to host deepfake porn videos in the past seven years. Some videos have been watched millions of times. Research also identified an additional 300 general pornography websites, including "leak" websites and websites that exist to repost people's social media pictures incorporating deepfake images. One website dealing in photographs claims it has "undressed" people in 350,000 photos.

These deepfake videos may be fake, but their emotional impacts are real. Victims are left with multiple unknowns: who made them? Who has seen them? How can they be contained? Because once something is online, it can reappear at any moment. In its consequences, this type of violation can be as devastating as revenge porn—real intimate photos released without consent. This takes a well-documented and real toll on victims' careers and personal lives. For example, one victim reported that after she was alerted to a series of deepfakes on a porn site that appeared to show her engaging in extreme acts of sexual violence, the images replayed themselves over and over in nightmares, and an all-consuming feeling of dread gripped her; she suffered panic attacks and was afraid to leave her home because a feeling of shame and fear permeated her life. Other victims have equated their deepfake pornography victimization to their prior real-life experiences with sexual assaults and violence.

In addition, women have suffered economic, reputation, mental health, and emotional injuries. In some cases, they've had to change their names. Others have had to remove themselves from the internet completely. Thus, the abuse creates "the silencing effect," where women feel discouraged from participating online and in other public spaces. Moreover, victims constantly fear being retraumatized because, at any moment, the images could resurface and once again ruin their lives. Finally, like other female victims of sexual violence, many victims of pornographic deepfakes and cybercrimes do not come forward, owing to the shame and harassment that can follow. Consequently, these harms are underreported, which makes it difficult to gauge the full scope of the problem and its devastating consequences.

As deepfake technology becomes more refined and easier to create, the inadequacy of the law to protect potential victims has also emerged. Because deepfake technology could be used to create realistic pornographic videos without the consent of the individuals depicted, these deepfakes exist in the realm of other sexually exploitative cybercrimes

such as "revenge porn." Pornographic deepfakes and revenge porn have troubling commonalities. Like revenge porn, pornographic deepfakes predominately affect women, and they both violate the individuals' expectation that sexual activity should be founded on consent. Both revenge porn and pornographic deepfakes can also cause similar harm to the victim's reputation and emotional well-being.

Despite the similarities, pornographic deepfakes and revenge porn differ. Revenge porn typically involves images of private individuals engaged in intimate acts intended to remain private. Thus, criminal liability has been imposed in part because revenge porn violates the victim's right to sexual privacy. Pornographic deepfakes, in contrast, do not necessarily raise the same sexual privacy concerns. Because deepfakes do not depict an actual person, no one has a privacy concern at stake in a deepfake. Nonetheless, in the case of a private individual whose face was used to create a pornographic deepfake, viewers may not realize the video depiction is fake, and the audience may assume that the video is genuine.

In addition, deepfakes also differ from revenge porn because of the challenge of identifying the victims. Revenge porn usually involves easily identifiable victims. Each deepfake video, however, depicts at minimum two people — the person whose body is shown, who may be challenging to identify, and whose face has been added. Problems also exist in locating and bringing the perpetrators to justice and removing the video once published online.

To address the mounting dangers of pornographic deepfakes, legislative action is needed. The Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 (the "Bill") is a good start. However, as I explain below, certain aspects of the Bill appear to be underinclusive while others are potentially over-inclusive and could benefit from additional revision.

B. Suggested Revisions

1) **Definition of "Transmit."** The Bill states that the offense is committed when a person "transmits [s]" the material. Transmit is defined elsewhere as "make available, publish, distribute, advertise, and promote." Notably absent from this definition is the act of "creating" the deepfake content; this omission is problematic. The content creator's criminal culpability is equal to, if not greater than, those who transmit the harmful content. The creator should be subject to criminal sanction under this law, along with those who make the content available. I recommend the Bill be revised to include in the definition of "transmit" the term "create" in section 474.17A.

   I recognize that those who "create" these deepfakes may be subject to the aggravated penalty in 474.17AA. Those who create *and* transmit these harmful images should receive a severe sanction. However, there is a gap in

the architecture between sections 474.17A and 474.17AA. Criminal liability for a creator is conditioned on their committing the 474.17A offense *and* that they have been subjected to civil penalty orders before committing the offense. Unless the creator has already been subject to the requisite civil penalty order, the creator of deepfake might escape criminal punishment for their crime. This omission can be corrected by revising section 474.17AA to include "create" in the definition of transmit.

2) **Definition of "Other Person."** The Bill states that the victim is the "other person," but it does not define what the "other person" means (except that the other person appears to be 18 years or older). Because deepfakes are often created using digital audio images of more than one person or an actual person and AI-generated body parts, the law needs to clarify who the law should apply. A definition of "other person" should encompass all the necessary possible iterations, including multiple persons.

3) **Knowledge Requirement.** Regarding the section scienter element in the Bill, I recommend expanding the standard for knowledge of consent to the transmission to include constructive knowledge based on the totality of the circumstances. Excluding those who should know but have remained intentionally blind to this conduct is troublesome, yet willful blindness is not included in the knowledge element.

4) **Application to Minors/Children.** The class of perpetrators also needs further refinement. I agree with the commentators who are concerned that the Bill could criminalize the conduct of those under the age of 18. Although imposing adult criminal liability might be appropriate for older minors (over 16 years old) who otherwise satisfy the elements of the offense, applying the Bill to young minors is a troubling possibility in the Bill's current form. Determining which cases involving minors might be appropriate for a section 474.17A charge requires a case-by-case consideration, which might better be addressed to other subsections, specifically including the age of the perpetrator as among the express considerations under exception (3)(d) in 474.17A.

5) **Reasonable Person Exception.** Finally, as to the "reasonable person" exception in section 474.17A (3)(d), I have two concerns. First, I am concerned about using the term "acceptable." Even with the qualifications that follow it, the word "acceptable" is ambiguous and overbroad; it might include many unanticipated meanings and applications. If the purpose of this exception is to embed community standards and common sense into the application of this law, that can be more directly achieved by revising this exception accordingly: "(d) a reasonable person would conclude transmitting

material does not violate subsection (1), having regard to the following…". Deleting the word "acceptable" removes an additional layer of ambiguity.

Another part of the exception in section 474.17A (3)(d), which should be revisited, is the use of the word "privacy" in (vi): "the degree to which the transmission of the material affects the privacy of the person depicted, or appearing to be depicted, in the material." As I observed elsewhere, because deepfakes do not depict an actual person, no one person has a privacy concern at stake in a deepfake. That said, I would recommend a minor revision to (vi) to account for the possibility that it could affect the privacy of individuals: "(vi) the degree to which the transmission of the material *has or could affect* the privacy *the person or persons* depicted, or appearing to be depicted, in the material." This revision allows for consideration of the privacy interests of persons whose likenesses have been used in the deepfake but cannot be identified.

C. Final thoughts

In addition to amending the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 to address the harms posed by pornographic deepfakes, I would also recommend that Australia's Online Safety Act and Sharing of Abhorrent Violent Material Act also be reviewed, and if necessary, amended to include within their relevant definition sections provisions to address deepfakes.

Australia has developed a robust framework of laws to regulate the Internet and social media companies over the last 15 years. Victims of pornographic deepfakes need concrete solutions for holding content creators liable, policing distributors, and the removal of offensive content from the internet. Criminalizing deepfakes under the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 is a step in the right direction.

Professor Rebecca A. Delfino
LMU Loyola Law School, Los Angeles, California

July 18, 2024