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

TESTIMONY OF PROFESSOR REBECCA DELFINO
ON DEEPFAKES
To Maryland General Assembly
February 13, 2023

The following testimony is true based on my 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 as a law professor at LMU Loyola Law School, Los Angeles, California. I present this testimony under the penalty under the laws of California and the United States:

The impact of deepfakes on the victims:

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's start. Existing abusive relationships have worsened, and digital abuse has seen an uptick as people have grown increasingly isolated and spend more time online.

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 pervaded 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. In others, they've 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.

Why is it important to provide a legislative solution:

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 difficult to identify, and the person whose face has been added. Problems also exist in locating and bringing the perpetrators to justice and removing the video once it has been published to the internet in light of the Communications Decency Act of 1996, 47 U.S.C. Section 230, which shields websites and content distributors from liability for third-party content.

Furthermore, neither the creation nor distribution of a pornographic deepfake violates federal law. And although 46 states criminalize revenge porn, only two—Virginia and California have laws explicitly addressing deepfake pornography. Prosecutors might attempt to prosecute pornographic deepfakes under existing criminal laws, including cyberstalking, criminal threats, unauthorized access to digital files, or a state’s revenge porn statute. These crimes, however, contain elements that are often absent in deepfakes, including the fear of bodily injury or death, a communication of a threat, or a pattern of conduct. In addition, a revenge porn statute may require proof that the perpetrator distributed images of an identifiable person, that the victim had a reasonable expectation of privacy in the depiction, and suffered emotional distress.

This leaves only a smattering of existing civil and criminal laws that may apply in specific situations. If a victim’s face is pulled from a copyrighted photo, it’s possible to use IP law. And if the victim can prove the perpetrator’s intent to harm, it’s possible to use harassment law. But gathering such evidence is often impossible, leaving no legal remedies for most cases. Thus, most pornographic deepfake victims have no recourse when their image is used on the internet without their consent.

Victims of pornographic deepfakes like revenge porn need concrete solutions for holding content creators liable, policing distributors, and the removal of offensive content from the internet. As in the case of revenge porn, criminalizing deepfakes under state law would be a good start.

Why is the issue complex requiring a task force of diverse experts:

Regulating deepfakes presents complicated issues requiring experts’ feedback and input from various constituencies. First, unlike other conduct subject to governmental regulation, especially the subject of criminal laws, deepfakes are not inherently harmful. Like any technology, there are also positive uses for deepfakes. They can be used for satire and entertainment. Using deepfakes, historical figures may be brought back to give a virtual tour around a museum or teach a history lesson in high school. Deepfakes are also increasingly being used in the film industry, either by having deceased actors finish a movie or by having an actor’s deepfake perform stunts. Deepfakes foster business and enable people to communicate more effectively. Deepfake influencers exist and get sponsorship contracts. Deepfake technology can be used when people speak in a different language over Zoom or Skype for their words to be translated live and have their lips synchronized to match the translation. Deepfakes can also be used to safeguard the privacy of individuals. They can be used for medical applications, forensic research, and education systems by creating learning tools and can also be used as an accessibility feature within technology. Further, deepfake tech has been used in marketing to help small businesses promote their products; for example, it can be used to create a virtual fitting room, allowing customers to try on clothes based on data about their gender, length, and weight. Charities use deepfakes, for example, by having a celebrity seemingly call out in all the languages of the world to support their cause, or by giving people a general idea of what a destroyed city looks like, by giving famous cities a 'remake,' to increase support for taking in refugees. Thus, many stakeholders and communities should be considered when considering regulatory responses to deepfakes.

Second, because deepfakes involve conduct considered "expressive," regulation of any deepfakes raises constitutional concerns. Currently, the First Amendment is a significant challenge to the government's ability to regulate deepfakes because of restrictions on limiting free speech. Relatedly, deepfake creators often have a First Amendment defense in civil claims against them. Thus, crafting a law to criminalize pornographic deepfakes must balance first amendment concerns with the legitimate need to protect victims. Drafting legislation that will withstand First Amendment scrutiny will require a clause-by-clause, word-by-word—analysis to test the breadth and scope of any prohibition which might inhibit expressive conduct to ensure that the law will protect victims and not chill protected speech.

Finally, almost all deepfakes appear and are spread on the internet, which is an inherently challenging space to regulate. The internet does not adhere to jurisdictional boundaries and is partially regulated by federal law. Thus, any state regulation of deepfakes will require a careful, multi-view perspective and coordination with other laws.

Professor Rebecca Delfino
LMU Loyola Law School

February 13, 2013