AELJ Blog

Journal Blogs

1-19-2020

Deepfakes May be in Deep Trouble: How the Law Has and Should Respond to the Rise of the AIAssisted Technology of Deepfake Videos

Amanda Inglesh
Cardozo Arts & Entertainment Law Journal

Follow this and additional works at: https://larc.cardozo.yu.edu/aelj-blog

Part of the Law Commons
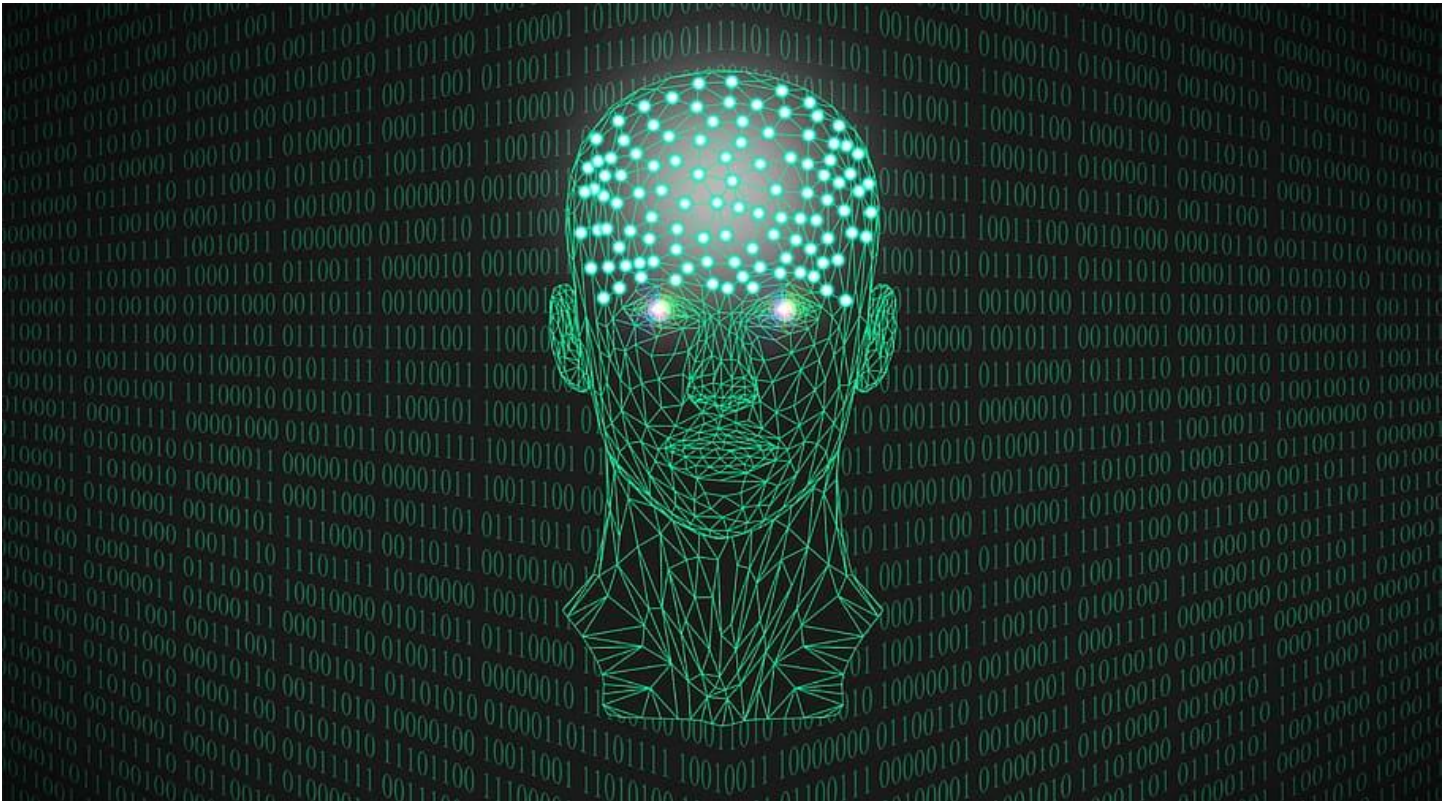
Deepfakes May be in Deep Trouble: How the Law Has and Should Respond to the Rise of the AI-Assisted Technology of Deepfake Videos

BY AMANDA INGLESH / ON JANUARY 19, 2020

Advanced technology and the rise of social media networks have led to a vast increase in video sharing, whether it be for entertainment or news. Recent innovations in technology have also allowed people to edit videos, creating falsified or fabricated content. The popular term millennials have come to associate with this trend is "fake news." Fake news is defined as information that is invented by people or governments that are "fictions deliberately fabricated and presented as nonfiction with the intent to mislead recipients into treating fiction as fact or into doubting verifiable fact."[1]

New artificial intelligence ("AI") technology has recently been introduced to the world of "fake news," now allowing producers to superimpose human faces onto those of others and manipulate sound. As a result, content creators can make it appear as though some person is saying and doing things they have never done.[2] These are called "deepfake" videos and they have the potential to threaten our democracy.[3] Just imagine if one of these videos featured a political figure taking a bribe or declaring war. Such images could incite incredible violence between and within countries. Moreover, some agree that by the 2020 election, deepfake

videos will be more prevalent and problematic.[4] Deepfakes also raise legal questions about defamation, privacy, and other IP-related issues.

In terms of privacy, pornography is a major area most affected by deepfake videos.[5] A recent trend was to create fake sex videos featuring public figures and celebrities, such as Emma Watson.[6] Private individuals are also affected.[7] Victims of these videos may claim defamation, which is an oral or written communication of a false statement about another that unjustly harms their reputation.[8] Sexual deepfake videos implicate false statements about a person, while placing them in offending situations. Normally, since the First Amendment of United States Constitution protects American citizens' freedom of speech,[9] any law that interferes with Americans' ability to express themselves is problematic. However, neither false speech nor obscenity are afforded First Amendment protection.[10] Moreover, untrue and damaging statements may lead to claims of defamation. In most cases, the only effective target of a defamation suit would be the online platform (i.e. Facebook or Twitter). However, Congress immunized platforms from such tortious liability in 1996 with Section 230(c)(1) of the Communications Decency Act.[11] In other words, when a Facebook user shares defamatory fake news, such as a pornographic deepfake video, the one who posted the video will be treated as the publisher and not Facebook.

Section 230 includes an intellectual property exception.[12] Thus, instead of bringing a tort claim for defamation, a victim may instead claim copyright infringement of his or her photographs used to create the deepfake. However, this claim is likely to fail because the deepfake producer is making fair use of the collated copyrighted images used in the videos.[13] Moreover, deepfakes are largely transformative and often not for commercial profit. Alternatively, a victim may claim a violation of his or her right of publicity. Celebrities commonly make right of publicity claims, often seeking to prevent businesses from unauthorized use of their images or identity in an advertisement.[14] However, a right of publicity claim has not yet been tested in the courts concerning deepfakes and could be difficult to prove.[15]

Some argue that, despite the unlikelihood of success for these claims, because there are remedies available through both tort law (i.e. defamation claims) and IP law (i.e. copyright and right of publicity claims), we don't need new laws addressing deepfake videos at all.[16] However, with the increased development of AI-technology and social media, deepfakes may only get more sophisticated and dangerous. In 2018, Congress recognized this danger and introduced the Malicious Deep Fake Prohibition Act ("MDFPA"), which would establish a new criminal offense related to the creation or distribution of fake electronic media records that appear realistic.[17] The DEEPFAKES Accountability Act ("DAA") was also introduced in 2019, requiring that watermarks be placed on all deepfakes so that people will know what is real versus fake news.[18] Unfortunately, the MDFPA expired at the end of 2018 and the DAA watermark requirement is likely to be ignored by those whose sole purpose is to weaponize a deepfake as real information.[19] Moreover, the DAA legislation broadly defines

"deepfake" as any AV work that "appears to authentically depict any speech or conduct of a person who did not in fact engage in such speech or conduct."[20] Although federal legislation may be necessary, Congress has yet to draft a promising law that appropriately solves the issues of deepfakes.

In the meantime, California has already legislated. Effective in 2020, two new California laws will regulate the distribution of deepfakes. AB 730 prohibits the use of deepfakes to influence political campaigns within sixty days of an election,[21] while AB 602 provides that a depicted individual in a pornographic deepfake video has a cause of action against the person who creates and discloses sexually explicit material of this kind.[22] For victims of deepfake videos, these California laws seem reasonable and well-tailored to the issues at hand.

Nonetheless, the biggest concern with these laws—whether state or federal—is free speech, especially political speech. The predominant purpose of the First Amendment right of free speech is to allow Americans the ability to say or publish their opinions without government interference.[23] The proposed federal bills concerning deepfakes do not exactly balance the threat of deepfakes with First Amendment interests, since one seeks to create harsher criminal liability and the other broadly defines terms that could implicate protected speech. However, California's AB 730 is more narrowly directed at the upcoming 2020 election by simply banning political deepfakes within sixty days of an election rather than implicating all political speech that is a protected opinion or a well-supported statement. The law also provides exceptions for satire and parody.[24] Therefore, content producers can express themselves with respect to the country's political climate without fear of liability.

At this point in time, there is no well-established caselaw on deepfakes. However, California is leading the way with legislation on the matter and Congress should follow suit by finding ways to narrowly tailor such laws to the purpose of protecting people from damaging, false content. We have seen behavior similar to deepfakes criminalized[25] and there are some civil remedies available.[26] Alternatively, Congress could approach the issue under a narrower IP-lens, addressing the ways in which deepfake technology violates individuals' IP rights. Ultimately, to think that we do not need any new laws regarding deepfakes would be deeply incorrect.

Amanda Inglesh is a second-year law student at the Benjamin N. Cardozo School of Law and a Staff Editor at the Cardozo Arts & Entertainment Law Journal. Amanda is interested in music law and copyright litigation. Amanda is also the current co-music chair on the executive board of Cardozo's Entertainment Law Society and is planning a spring event called Women in Entertainment. She has previously interned at Serling Rooks Hunter McKoy Worob & Averill LLP and volunteered at Volunteer Lawyers for the Arts.

[1] Eric Emanuelson, Jr., *Fake Left, Fake Right: Promoting an Informed Public in the Era of Alternative Facts*, 70 Admin. L. Rev. 209, 218 (Winter 2018).

[2] Kevin Roose, *Here Come the Fake Videos*, *Too,* N.Y TIMES (Mar. 4, 2018), https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html.

[3] *See* Juan Carlos Escudero de Jesus, *Fake News and the Systemic Lie in the Marketplace of Ideas: a Judicial Problem?*, 87 Rev. Jur. U.P.R. 1394, 1406, 1415 (2018); *see also* Ishaan Tharoor, *'Fake news' and the Trumpian threat to democracy*, The Wash. Post (Feb. 7, 2018), https://www.washingtonpost.com/news/worldviews/wp/2018/02/07/fake-news-and-the-trumpian-threat-to-democracy/.

[4] *See* Holly Kathleen Hall, *Deepfake Videos: When Seeing Isn't Believing*, 27 Cath. U. J. L. & Tech. 51, 56(2018).

[5] *Id.*; *see also* Douglas Harris, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, 17 Duke L. & Tech. Rev. 99 (2019).

[6] *See* Robert Chesney & Danielle Citron, *Deep Fakes: A Looming Crisis for National Security, Democracy and Privacy?* LAWFARE (Feb. 21, 2018), https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy#, https://perma.cc/L6B5-DGNR.

[7] *See* Harris, *supra* note 5.

[8] *Defamation*, Merriam Webster Dictionary, https://www.merriam-webster.com/dictionary/defamation.

[9] U.S. Const. amend. I ("Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.").

[10] Harris, *supra* note 5 at 104.

[11] 47 U.S.C. § 230(c)(1) ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.").

[12] 47 U.S.C. § 230(e)(2) (Nothing in the section "shall be construed to limit or expand any law pertaining to intellectual property").

[13] Harris, *supra* note 7 at 107.

[14] Jesse Lempel, *Combatting Deep Fakes through the Right of Publicity*, LAWFARE (March 30, 2018), https://www.lawfareblog.com/combatting-deep-fakes-through-right-publicity.

[15] *Id.*

[16] David Greene, *We Don't Need New Laws for Faked Videos, We Already Have Them*, Electronic Frontier Foundation (Feb. 13, 2018), https://www.eff.org/deeplinks/2018/02/we-dont-need-new-laws-faked-videos-we-already-have-them.

[17] S.3805 – Malicious Deep Fake Prohibition Act (2017-2018), 115th Congress, https://www.congress.gov/bill/115th-congress/senate-bill/3805.

[18] H. R. 3230 – 116th Congress (2019-2020), https://www.congress.gov/bill/116th-congress/house-bill/3230/text.

[19] Nina Iacono Brown, *Congress Wants to Solve Deepfakes by 2020*, Slate.com, https://slate.com/technology/2019/07/congress-deepfake-regulation-230-2020.html (July 15, 2019).

[20] *Supra* note 18.

[21] AB 730 (2019-2020), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB730.

[22] AB 602 (2019-2020), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB602.

[23] Tala Esmaili, *First Amendment: An Overview*, https://www.law.cornell.edu/wex/first_amendment (Last Updated June 2017) ("The Supreme Court requires the government to provide substantial justification for the interference with the right of free speech where it attempts to regulate the content of the speech. Generally, a person cannot be held liable, either criminally or civilly for anything written or spoken about a person or topic, so long as it is truthful or based on an honest opinion, and such statements.").

[24] *Supra* note 18.

[25] Harris, *supra* note 5 at 120 (examining how recent state criminal statutes concerning "revenge porn" fail to prohibit this conduct because producers likely do not intend to harm the featured person, which is a requirement in most of these statutes).

[26] *Supra* note 11.