

Saturday Seminar | Technology | Aug 14, 2021

Responding to Deepfakes and Disinformation

Soojin Jeong, Margaret Sturtevant, and Karis Stephen



Experts assess ways that regulation might respond to the problems of deepfakes and disinformation.

Did Tom Cruise really [impersonate](#) a snapping turtle on TikTok? Did former President Barack Obama actually [reference](#) the movie Black Panther and say the supervillain was right? In both cases, the answer is “no.” These videos are deepfakes.

“Deepfakes”—a hybrid of the terms “deep learning” and “fake”—[are](#) videos, photos, or audio recordings that have been modified to make false content appear real. The technology can [replace](#) faces and speech to make it appear as if someone said or did something that never happened.

The most sophisticated deepfake videos [require](#) thousands of images to train algorithms to recognize and then manipulate a face. In March 2021, realistic deepfake videos of Tom Cruise on TikTok [fooled](#) deepfake detection software, because the video algorithm used over 13,000 images of Tom Cruise that captured him from almost every angle.

Although only sophisticated experts can make the most realistic deepfakes, anyone with an iPhone and a single photo can now [use](#) free apps to create simple deepfakes. These amateur deepfakes are stilted and noticeably computer-generated, but as technology to create realistic deepfakes becomes more accessible, soon anyone could have the power to wield deepfakes and interfere with other people’s identities and reputations.

Deepfakes have profound potential to cause harm. The first deepfakes [appeared](#) around 2017 and comprised pornographic videos that swapped female celebrities’ faces onto other persons’ bodies. By 2019, a startup [reported](#) that approximately 96 percent of deepfakes are pornographic and disproportionately victimize women, casting them in humiliating, violent situations and [exposing](#) them to rape and death threats. Bad actors can also [exploit](#) others by using deepfakes to commit identity theft, blackmail, and fraud.

Deepfakes are uniquely effective at spreading disinformation, which [raises](#) critical concerns for democracy and national security. Effective democratic discourse requires that voters start from the same foundation of facts, but deepfakes can [lead](#) individuals to live in their own subjective realities and [exacerbate](#) social divisions.

Deepfake videos depicting public figures making incendiary comments or behaving inappropriately could also [alter](#) election outcomes. As deepfakes become more well

known, public officials caught on camera can [exploit](#) a “liar’s dividend” and claim that a real video is a deepfake. Without clear methods to distinguish what is real from what is not real, the public may [lose](#) trust in media and other public institutions.

In an initial effort to respond to deepfakes, the U.S. Congress adopted [legislation](#) in 2019 requiring the Director of National Intelligence to prepare annual reports about the national security impacts of deepfakes and foreign weaponization of deepfakes.

Members of Congress have introduced other legislation to address deepfakes. In July, a bipartisan group of Senators introduced the [Deepfake Task Force Act](#), which would establish a task force in the U.S. Department of Homeland Security to research and develop a government-wide plan to counter deepfakes.

Earlier proposals included the [Identifying Outputs of Generative Adversarial Networks Act](#), which would authorize the National Science Foundation to support research on deepfakes and the National Institute of Standards and Technology to develop measurements and standards for technologies to detect deepfakes. The [Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act](#) would mandate that deepfakes identify themselves as altered media by containing embedded digital watermarks and including verbal or written disclosures that describe the alterations.

In addition to these federal proposals, various states have [passed](#) deepfake legislation. And social media platforms have changed their terms of service in efforts to reduce harms to individuals and to protect the integrity of elections.

In this week’s Saturday Seminar, scholars explore the problems created by deepfakes and the challenges for regulators seeking to address these problems.

- Deepfake videos raise serious concerns, because videos are inherently credible, interact with cognitive biases, and travel quickly on social media platforms, argue [Danielle K. Citron](#) of [Boston University School of Law](#) and [Robert Chesney](#) of the [University of Texas School of Law](#) in an [article](#) in the [California Law Review](#). Analyzing statutory authority for three federal agencies, Citron and Chesney [conclude](#) that the Federal Trade Commission, Federal Communications Commission, and Federal Elections Commission have only limited abilities to address deepfakes. Citron and Chesney [recommend](#) other legal remedies and more creative thinking to respond to the problem of deepfakes.

- Tom Dobber of the [University of Amsterdam](#) and several coauthors in a report in the *International Journal of Press/Politics* find that deepfakes can substantially impact viewers' beliefs about a political candidate and influence their views of the candidate's political party. Dobber and his coauthors warn that, as technology advances, the precision and effectiveness of deepfakes are likely to increase. They argue that deepfakes are a "potential new frontier of disinformation warfare" that requires policy action.
- Deepfakes create major challenges to the integrity of the democratic process in the United States, argues Richard L. Hasen of [University of California, Irvine School of Law](#) in an article in the *Saint Louis University Law Journal*. But because governmental efforts to control deepfakes could implicate free speech rights, Hasen warns that such efforts must be justified as serving a compelling state interest and then designed in the most narrow fashion possible. Hasen explains that combating election misinformation will constitute a compelling interest. But he worries that it may be more difficult for governmental entities to show that they have employed a narrowly tailored approach to addressing deepfakes. Hasen suggests that a better alternative, which might avoid First Amendment concerns altogether, might be to mandate truth-in-labeling requirements for all campaign communications and require more robust campaign financial disclosures.
- In an article for the *Catholic University Journal of Law and Technology*, Holly Kathleen Hall of [Arkansas State University](#) explains that government authorities will find it difficult to take action against those individuals who create deepfakes. Deepfake videos often spread anonymously in public forums, Hall notes. In addition, any speech regulations in public forums must be content-neutral to comport with the First Amendment. Hall recommends a multilevel approach to discourage deepfake videos, including elevating media literacy, increasing fact-checking efforts, and encouraging internet companies to create new policies.
- In an article in the *Journal of Intellectual Property Law & Practice*, Edvinas Meskys of [Vilnius University School of Law](#) and three coauthors develop a taxonomy of deepfakes, classifying them into four categories: revenge porn; political campaigns; commercial uses; and creative uses. They conclude that market-driven solutions may work best to address the varied problems created by these different types of deepfakes and suggest the need for developers to create sophisticated technologies to detect deepfakes.
- Although some lies may receive free speech protection under the First Amendment, a deepfake spreading false information may lose First Amendment protection if it is "not only a falsity, but a forgery as well," suggests Marc Jonathan Blitz of [Oklahoma](#)

City University School of Law in an [article](#) published in *Oklahoma Law Review*. Blitz [argues](#) that deepfakes presented as news can create a “war of all against all” where every source of information could be false and impact individual decisions or collective democracy negatively. Blitz [argues](#) that, like individuals imitating government officials, deepfakes that appear in the form of reliable news should be subject to more government regulation.

The Saturday Seminar is a weekly feature that aims to put into written form the kind of content that would be conveyed in a live seminar involving regulatory experts. Each week, *The Regulatory Review* publishes a brief overview of a selected regulatory topic and then distills recent research and scholarly writing on that topic.

Tagged: [Artificial Intelligence](#), [Deep Fake News](#), [Disinformation](#), [facial recognition](#)

Related Essays

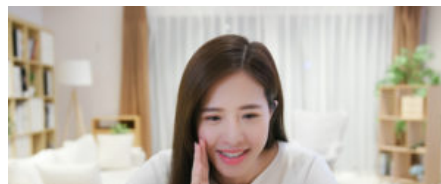


SYNOPSIS | TECHNOLOGY
MAR 9, 2023

Creating a Safe Testing Space for High-Risk AI

Joe Katz

Scholars argue for the creation of a regulatory safe space to supplement AI regulations.



SATURDAY SEMINAR |
TECHNOLOGY
FEB 18, 2023

Hey Siri, Are You Regulated?

Elizabeth Yin, Mary Moynihan, and Alexandra Walsh

In this week's Saturday Seminar, experts propose ways to regulate voice-activated technology and protect consumer privacy.



ANALYSIS | TECHNOLOGY
JAN 24, 2023

AI Art Is in Legal Greyscale

Elizabeth Penava

The legal ambiguity of art created by artificial intelligence adds confusion to controversy.

Subscribe to Updates

Your Email

SIGN UP »



Contact Info

Penn Program on Regulation
University of Pennsylvania Carey Law School
3501 Sansom Street
Philadelphia, Pennsylvania 19104

Editor-in-Chief
Soojin Jeong
Editor@TheRegReview.org

Faculty Advisor
Cary Coglianese
+1 215.898.6867

Topics

Business

Education

Environment

Health

Infrastructure

International

Process

Rights

Technology

Resources

Contributors

Submissions

Subscribe

Series

Saturday Seminar

Week in Review

E-Rulemaking

RuleFinder

Links



©2023 University of Pennsylvania Law School

[Privacy Policy](#)

Site by
Capital Technology Services