R Street

Free markets. Real solutions.

R SHEET ON DEEP FAKES

March 2019

BACKGROUND

The term "deep fake" refers to AI-generated audiovisual media that can convincingly resemble authentic content. Although realistic animated video has been a Hollywood mainstay for decades, the concept has attracted attention recently due to advances in AI technology that significantly simplify the ability to generate realistic representations. Specifically, "deep learning" algorithms (hence the name "deep fakes") are designed to teach computers how to create fake videos that are virtually identical to real ones (at least to the naked eye).

CURRENT DEBATE

While such technology has some important benefits (for example, investigators can use it to realistically age a photo of a missing child and teachers can bring historical figures to life in the classroom), many have raised concerns about its potential for harm. In particular, the role of social media in the 2016 election has led to grave concerns about the possible impact of this technology going forward. For example, many worry that a well-placed deep fake that falsely maligns a politician or causes other kinds of unnecessary national panic could traverse the web at breakneck speeds, allowing foreign actors to have a potentially catastrophic impact on the electoral process or national security.

Such fears are exacerbated by the supposedly unique sophistication of today's fake content and the ubiquity of the technology that creates it.

ACTION ITEMS

Narrowly Tailor Responses

While deep fakes may appear to be unprecedented, many of the concerns they raise merely mimic those associated with the rise of digital photography and manipulation tools such as Photoshop in the late 1990s and early 2000s. However, as that software made it easier to convincingly manipulate images, society reacted and adapted to their existence. In fact, nearly everyone is now aware of digital photo manipulation and indeed has come to expect it in certain contexts, such as advertising. It follows that society will likewise adapt on its own to the proliferation of deep fakes.

SUMMARY

- Deep fakes are a class of AI-generated audiovisual material designed to appear realistic.
- Because of how closely they mimic reality and the ease with which a bad actor can utilize the technology, some have begun to worry about the harms that the technology can do to society.
- However, deep fakes are merely the most recent in a long line of deceptive technologies.
- Regulators should not use the fear of deep fakes as an independent justification for regulating the more general problem of disinformation.

Moreover, fears associated with this new technology are being conflated with a more general concern about fake news. However, deep fakes are a specific issue—part of but separate from—the larger, more general issue of fake news dissemination.

For these reasons, any attempt to combat them must be narrowly tailored to focus on the specific harms associated with deep fake-related disinformation. And, such an approach should only target the person who generates and shares the content, rather than the deep fakes themselves—or the technology used to create them.

Resist the Urge to Codify Liability

Another countermeasure that has been suggested is to regulate the spread of deep fakes by amending Section 230 of the Communications Decency Act, a legal protection that currently renders online service providers not liable for third-party content. However, amending

Section 230 to put more liability on platforms will have other unintended consequences. For example, it will almost certainly limit the spread of legitimate online content. Moreover, while some providers may choose to take less editorial discretion and simply treat their platform as a "dumb pipe" to avoid liability, others may overregulate content to ensure that they can still avail themselves of the section's protections. This, of course, would arbitrarily restrict content available to users.

And finally, it is also important to understand existing market responses. Industry and academia are already taking steps to limit the spread of harmful deep fake videos. Many platforms and content generators alike face pressure from their users to limit the spread of blatantly deceptive video. In fact, some platforms have already chosen to invest in and employ deep fake detection technology that can identify fake videos and target suspected deep fakes for review. Likewise, working with the Defense Advanced Research Projects Agency (DARPA), academics have been using research on biological functions such as eye blinking to distinguish fake videos from real ones. Additional techniques remain undisclosed to allow detection to stay in front of generation. And, while it is true that deep fake technology will soon adapt to these detection techniques, so too will the techniques continue to adapt to the fakes. In light of this, policymakers should be instructed by the response to digital editing and Photoshop, and should therefore focus on educating the public about deep fakes and how to recognize them. After all, as society becomes more aware of the videos, their harms will be substantially diminished.

In any event, rushing to force liability onto companies who already face pressure to stop the spread of such content will do little to solve the problems unique to deep fakes (as opposed to disinformation generally), and will drastically hurt an internet ecosystem that relies heavily on the protection of internet service providers. Lawmakers should therefore exercise restraint when addressing deep fake regulation.

CONTACT US

For more information on this subject, contact the R Street Institute: 1212 New York Ave NW Suite 900, Washington D.C. 20005, 202.525.5717.

Jeffrey Westling
Technology and Innovation Policy Associate
jwestling@rstreet.org
602-284-0553

Charles Duan
Director of Technology and Innovation Policy
cduan@rstreet.org
202-900-8247