DEEPFAKES: their threat to identity, national, security, and democracy.

By Grace Sa

Figure 1 is a cover photo from the video “Deepfake Videos Are Getting Real and That’s a Problem” from the Wall Street Journal website.

What are Deepfakes?

Deepfakes are videos produced using artificial intelligence where someone can look and sound like anyone they would like. Simply just having the software, camera, and microphone, a person can make a video of their idol or enemy say or do whatever they please. According to Cole (2017), they surfaced around December of 2017 and have been getting more advanced ever since.

The Problem with Deepfakes:

Imagine the former President of the United States calling the current President a “d*psh*t.” That is exactly happened in Boylan’s article (2018) when Jordan Peele created a deepfake using Barack Obama’s identity, and no one could tell a difference until he revealed himself. The problem with deepfakes is that anyone has the power to make influential people say and do whatever they want. It becomes an issue because when these videos are spread, many people cannot tell the difference and might believe that these actions were that of the real person. It harms celebrities, identities, politicians, democracy, minority groups and sex workers.

Myths and Misconceptions:

Many people believe that because major social media platforms and major pornography websites such as Reddit and pornography websites such as Reddit and Pornhub swore to have taken deepfakes videos down, people cannot access deepfakes. However, that is not true. According to Askham (2018) they are still easily found on those websites and other popular websites.

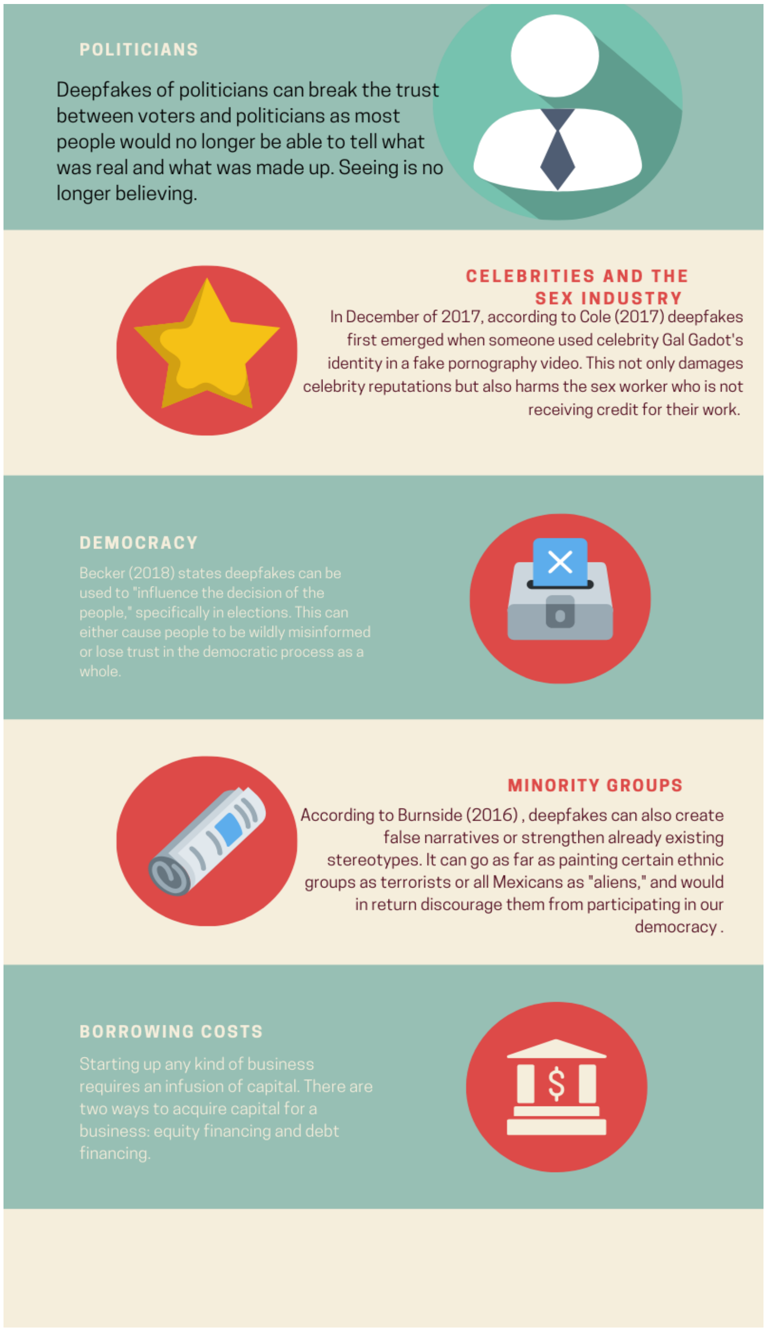
Perhaps the biggest misconception is that if we ban deepfakes, that would mean apps such as SnapChat would have to be deleted because they have face swapping filters. However, the difference between the two is that snapchat filters “face swap” using a still picture and do not record your voice, therefore it will never be technologically advanced to be a deepfake.

Potential Solutions:

Sullivan (2017) argues that dishonest misuse of someone’s digital transaction identity should be considered identity theft. A possible solution would be to also treat deepfakes as a form of identity theft if it were created without the individual’s consent.

Abbasi (2010) critiques how statistical learning theory (SLT) programs are used to detect fake websites but can often be ineffective because they are too simple. However, I believe the SLT programs could provide a good framework for a system much more advanced than the current SLT programs to detect deepfakes.

Another step towards a potential solution would be to make the host sites are held liable for the content they share. Many sites who claim to have gotten rid of deepfake pornography still have them uploaded. For the sake of victims, these websites need to regulate their content.



POLITICIANS

Deepfakes of politicians can break the trust between voters and politicians as most people would no longer be able to tell what was real and what was made up. Seeing is no longer believing.

CELEBRITIES AND THE SEX INDUSTRY

In December of 2017, according to Cole (2017) deepfakes first emerged when someone used celebrity Gal Gadot's identity in a fake pornography video. This not only damages celebrity reputations but also harms the sex worker who is not receiving credit for their work.

DEMOCRACY

Becker (2018) states deepfakes can be used to "influence the decision of the people," specifically in elections. This can either cause people to be wildly misinformed or lose trust in the democratic process as a whole.

MINORITY GROUPS

According to Burnside (2016) , deepfakes can also create false narratives or strengthen already existing stereotypes. It can go as far as painting certain ethnic groups as terrorists or all Mexicans as "aliens," and would in return discourage them from participating in our democracy .

BORROWING COSTS

Starting up any kind of business requires an infusion of capital. There are two ways to acquire capital for a business: equity financing and debt financing.

Works Cited

Abbasi, Ahmed, Zhu Zhang, David Zimbra, Hsinchun Chen, and Jay F. Nunamaker. "Detecting Fake Websites: The Contribution of Statistical Learning Theory." *MIS Quarterly* 34, no. 3 (2010): 435-61. doi:10.2307/25750686.

Askham, Gemma. “Are Deepfakes the New Revenge Porn? - BBC Three.” *BBC*, BBC, 25 Apr. 2018, www.bbc.co.uk/bbcthree/article/779c940c-c6c3-4d6b-9104-bef9459cc8bd.

Becker, Daniel. "Desiring Fakes: AI, Avatars, and the Body of Fake Information in Digital Art." In *Faking, Forging, Counterfeiting: Discredited Practices at the Margins of Mimesis*, edited by Becker Daniel, Fischer Annalisa, and Schmitz Yola, 199-222. Bielefeld: Transcript Verlag, 2018. http://www.jstor.org/stable/j.ctv1wxr9t.15.

Boylan, Jennifer Finney. "Will Deep-Fake Technology Destroy Democracy?" The New York Times. October 17, 2018.

Burnside, Julian. "Indefinite Disinformation: The Political Capital of Fear." *AQ: Australian Quarterly* 87, no. 1

Burnside, Julian. "Indefinite Disinformation: The Political Capital of Fear." *AQ: Australian Quarterly* 87, no. 1 (2016): 9-40. http://www.jstor.org/stable/24877807.

Chesney, Robert, and Danielle K. Citron. "Disinformation on Steroids: The Threat of Deep Fakes." *Council on Foreign Relations*, Council on Foreign Relations, 16 Oct. 2018, www.cfr.org/report/deep-fake-disinformation-steroids.

Cole, Samantha. "AI-Assisted Fake Porn Is Here and We're All Fucked." *Motherboard*, VICE, 11 Dec. 2017, motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn.

Franklin, Zak. "Justice for Revenge Porn Victims: Legal Theories to Overcome Claims of Civil Immunity by Operators of Revenge Porn Websites." *California Law Review* 102, no. 5 (2014): 1303-335. http://www.jstor.org/stable/24758167.

Sullivan, Clare. "Digital Identity — Protection." In *Digital Identity: An Emergent Legal Concept*, 107-36. South Australia: University of Adelaide Press, 2011. http://www.jstor.org/stable/10.20851/j.ctt1sq5wqb.12.