Should AI Deepfakes be Regulated?

Tyson Hallin

Department of English, Anoka Ramsey Community College

Engl 1121: College Writing and Critical Reading

Prof. Chris McCarthy

4/5/2023

I saw a video the other day of former presidents Barrack Obama and Donald Trump playing a videogame with current president Joe Biden. They were talking about hanging out, going to McDonalds, and then going down by a river to go fishing. If it weren't for the visuals just being still images of the men, I would have been completely sold on this being real. It turns out, however, that this video was an AI deepfake created by one person and a text to speech algorithm. Artificial intelligence, AI, can be explained as a computer-generated program that can learn entirely on its own. It teaches itself to help us out with everyday stuff, such as the autocorrect in our smart devices, or even the backup cameras in modern cars. It is no secret that AI has made leaps and bounds worth of improvements in recent years. Things like computer-generated images, chat algorithms, face tracking, voice fabrication, and many others that can help us make life much easier.

Computers and AI have been programmed to learn all sorts of unique and powerful abilities that are still partly inconceivable in the modern day. Do you want your art project done in three minutes? Ask AI to make a painting in the style of Van Gogh. Maybe you aren't worried about the morality of generated deepfakes, what can AI do for you? Start by writing a script and gathering information, then with enough substance, a robot can impersonate an important figure with near perfect accuracy. Are you feeling lazy, and you want to make an AI do chores for you? Get a Roomba, which uses a small amount of AI and a lot of sensors to scan your floors and rooms to map its way around your house. Is all this improvement really such a good thing? Don't get me wrong, I love playing games with unique non-playable characters, NPCs, that can perform actions and be dynamic throughout conversations and environments thanks to their AI systems. I enjoy watching videos where famous people are scripted to sound like they say things you would not otherwise hear come out of their mouths, but AI deepfakes, as they are known to

the public, can be dangerous to someone's character and to the community that surrounds them. AI is becoming a problem, and without government intervention we risk letting it get too far, taking over our jobs and lives.

Many people share the claim that AI is potent, just not very dominant yet. Lily Hay Newman, senior writer for *Wired,* titles her work "AI-Generated Voice Deepfakes Aren't Scary Good—Yet." In the article, Newman (2023) repeatedly argues that while AI deepfakes are a point for concern, scientists and programmers have not advanced them enough to be a global issue. She puts forward the idea that AI programs could require dozens of hours of someone's voice in order to perfectly recreate it. Also, early in the essay via *Wired*, Newman (2023) draws connections between scammers, fraudsters, criminals, and documentarians. Now the question becomes, "How does a documentary relate to crime in terms of AI algorithms?" She notes that a documentary crew made a film starring the late world-renowned chef Anthony Bourdain, using old footage of him and an AI deepfake program to recreate his voice. The author hints that without Bourdain's voice, the documentary would not have been as profitable, and whatever profits it generated should not be theirs due to Bourdain's likeness. What kinds of ethics does this call into question? Newman (2023) argues that even though Bourdain's voice wasn't a major contribution to the documentary, it still brought eyes to the otherwise unremarkable presentation that wouldn't have seen it otherwise.

A fair counterargument from Matt Binder, journalist and occasional guest on *The Majority Report*, is that people do not need to take hours to teach AI about someone, because AI already knows them. Binder (2023) puts it simply: most people that someone thinks about making a deepfake of already have an almost infinite supply of high-quality videos and images accessible anywhere over the internet at any given time. Take, for example, your favorite

YouTube content creator, or President Joe Biden, or even Binder himself. They all have hundreds of hours of "them looking directly at a 4k camera giving a perfect view of their face" (Binder 2023) across the internet. He also notes that most deepfakes, if not all, refuse to take into consideration the consent required for them to be technically legal. Due to the easily spreadable nature of deepfakes and fake news as a whole, millions of people have been influenced by something that isn't actually real. Binder (2023) says that there is an invisible line between acceptable and unacceptable, and deepfakes go way beyond unacceptable and straight to dangerous when used by the wrong people. Our older relatives are the most likely to be taken advantage of, in one way or another, with deepfake technology being the direct cause. Anyone can be influenced by deepfakes, for example: a president to skew your political views, a celebrity to alter how you view them or a product they may or may not support, or characters making a cameo in your favorite TV shows, who have either long grown out of their roles or since passed on. I'm talking about you, Luke Skywalker at the end of *The Mandalorian* season two.

As for my take, the inclusion of the word "yet" stood out to me because I find it hard to believe that "yet" is going to be a long time. While Newman (2023) doesn't hold back, I have a tough time taking her side when presented with the ideas of Seder and his colleagues. As mentioned previously, generating deepfakes can take a few hours, not a few decades. If this were a more futuristic issue, I would have no problem backing up Newman and her arguments, but she doesn't make it obvious that she knows how immediate the creation of deepfakes is. The point of a scam artist using chat AI to impersonate voices is not to make them perfect, because that takes time and money. Instead, they focus on minimizing expenses and maximizing revenue. Newman has a tainted view of the subject, where she thinks deepfakes need to be perfect to be usable.

Binder (2023) points out the simplicity of creating a poor, yet believable video of someone's likeness without their knowledge.

With the recreational use of AI, scammers also take advantage of deepfakes. "Technology is making it easier and cheaper for bad actors to mimic voices, convincing people, often the elderly, that their loved ones are in distress. In 2022, impostor scams were the second most popular racket in America, with over 36,000 reports of people being swindled by those pretending to be friends and family" (Verma, 2023). The likelihood that someone we know personally being attacked by a cybercrime assisted by deepfakes is on the rise, and there doesn't appear to be an end in sight. This "36,000" number is only the amount of people who reported the scams, so think about the possibility of how many people did not report their losses. An article by Katyanna Quach (2023) from *The Register* furthers this idea, except she chooses to focus on people who will attempt to swindle with images. Anyone can do this with a little bit of time, money, and Photoshop experience, but Quach (2023) debates that Generative AI can forge an entire experiment and its results in just seconds. Someone looking to do misdeeds can easily spread their false information through Google, Instagram, TikTok, and other wildly popular social media sites. To summarize the author's piece; someone looking to scam using statistics or science is likely to turn to AI and its limitless possibilities to look convincing, even if the evidence means nothing or doesn't really exist in the first place. The morally corrupt will find any way to earn a quick buck. This can also be expanded to a disconnect between teachers and students. Instructors, particularly those at the college level, have the added difficulty of deciphering the difference between an article written by a student, and one written by an AI algorithm. This is on the same level as a more traditional scam because in both cases someone is

willing to take the easy route to achieve results, in this case a better grade without having to put in any of the work.

The biggest scam of the past few years has been impersonating bank officials or customer service representatives to convince people to give up their important passwords and accounts. Moving into the future, with the large sums of money certain scammers bring in, they will be easily able to use AI to impersonate and reassure innocent people. A video from *Crash Course Computer Science* back in 2017 lays out the basics of the damage caused by scammers and hackers who use AI. When talking about damage that cyber-attacks cause due to AI, Carrie Anne says, "They cost the global economy roughly half-a-trillion dollars annually […]" (Crash Course, 2017). A sizable portion of this money comes from people who fall victim to phishing frauds at the hands of AI deepfakes.

AI deepfakes have also changed elections in the past. In 2019, there was a similar struggle with artificial intelligence leading up to the 2020 presidential election. At the time, there were claims that AI became high stakes "all of a sudden" (Metz, 2019), that deepfake creators are only in it for the money, and "What we ended up with, was half a dozen poorly made clips which feel more like parody than anything particularly serious" (Boyd, 2021). One specific *Vanity Fair* article from writer Charlotte Klein (2023) states, "Fake Biden speeches have been a favorite on social media of late, with simulations of the president talking about everything from hip hop to drugs and video games. But some are also using the technology to spread misinformation, such as a deepfake video of Biden criticizing transgender women." This shows a distinct difference between the positives and negatives AI can be used for. Later in this article, Klein (2023) mentions how deepfakes of the past were just beginning throughout the course of said 2020 election cycle, and many ways of coding AI have changed and improved due to those

earlier renditions, all moving toward the presidential elections in 2024. The University of Ohio (2023) furthers the difficulties in fact-checking these deepfakes. The main rules for fact checking, according to the University's own page on the topic, include hours of research along with keeping in mind where you get your news. The University writes, "Research shows that Facebook users engage with misinformation — which often takes the form of fake news — 70 million times per month on average. This is a decline from the 2016 peak of 200 million monthly fake news engagements, but still no small figure" (University of Ohio, 2023). Lots of people did not take AI deepfakes very seriously a few years ago, but now that time has passed and improvements have been made, the believability of deepfakes is much greater and more serious.

Over time, our government has already made some rule changes regarding the use of AI technology for deepfakes. In a speech to *The Cybersecurity, Information Technology, and Government Innovation Subcommittee*, the professor of computing at MIT, Aleksander Madry, says this:

> Tools like Midjourney (an AI tool for generating hyper-realistic images using a natural-language interface) are transforming the world of visual design and art," and "Seizing this opportunity and bringing about a future that is positive and empowering requires discussing the role of AI in our society and nation, what we want AI to do (and not do) for us, and how we ensure that it benefits us all. (Madry, 2023)

As mentioned in an article backed by the *Regulatory Transparency Project of the Federalist Society* deepening this speech, Matthew Feeney claims, "Although Deepfake technology is relatively new, it [may create] unique problems. Accordingly, lawmakers and officials should proceed with caution when considering Deepfake technology. Absent careful consideration, legislation intended to prevent the malicious use of Deepfake technology could stifle its valuable

uses." Even though Feeney (2021) doesn't say what valuable uses he is referring to, he isn't shy about the misuses of deepfake technology. He also references the previous problems our government has faced because of AI. "Even in freer countries politicians, candidates, and their allies have used manipulated media to their benefit" (Feeney, 2021). While he does not explicitly state AI deepfakes as the root of all adulterated media, Feeney (2021) does draw a connection between the increased uses of both false information and deepfake technology.

After all, AI can also do other things besides wishing you a happy birthday. Except for some being programmed better or worse than others, AI can take on physical bodies, with arms, legs, and heads. The addition of extremities to AI presents a new problem, they could move towards taking our jobs, too. Calum McClelland, Head of Operations at *IoT for All,* relates our current technological struggle with that of the industrial revolution. He writes in an article posted at the beginning of this year, "The Luddite movement occurred all the way back in 1811, so concerns about job losses or job displacements due to automation are far from new." Luddites, being the term for factory workers who were aggravated and concerned by the advancement in technology, technology that would take over their jobs permanently. McClelland (2023) draws connections between the industrial revolution and the modern era, with AI assistants becoming more prevalent. "Despite these fears and concerns," McClelland continues, "every technological shift has ended up creating more jobs than were destroyed." The same should not be said for artificial intelligence, because steam engines and light bulbs couldn't think for themselves like AI can.

An article highly skeptical of how well we should trust AI, that contradicts McClelland's arguments, lists the top five jobs most likely to be taken over by AI, whether it be a physical manifestation or a digital one. These are: cashiers, drivers, translators, customer service workers,

and warehouse workers (Rajnerowicz, 2022). Rajnerowicz uses studies based out of Oxford, most notably the "Oxford Study," which focuses on the future of employment in the U.S. He cites a graph ranking the most used words to describe human-like robots. Cute was by far the most used positive adjective used, and scary was the most used negative one. Another graph shows that 68% of people would be willing to let AI handle their personal finances, whereas only 54% would be willing to let robots handle the larger U.S. economy. As I mentioned, Artificial Intelligence and machine learning are lightyears beyond what the printing press used to be. Nothing from the industrial revolution could vacuum your floor and then help you with your math homework, but that's exactly where robots are heading. Rajnerowicz's report also adds a graph which mentions the least likely job to be taken over by AI, according to those surveyed; artists. The following will prove otherwise.

Artificial intelligence has recently learned how to mimic art. It has learned to understand what you tell it, and then create an artistic piece based off that description. Ben Meisner, a Council Member for *Forbes*, forces the question of "Why?" He writes in a 2022 article "But when it comes to photo and video manipulation, what is it we are ultimately working toward solving?" He is not convinced that AI art should be accepted as art, and he believes that in some cases it can step over the line of copyright infringement. Meisner (2022) argues that AI needs to become much more human-like for its art to be compared to that of humans. The writer takes the stance that "emotionless machine[s]" cannot create art with the same meaning and intention that humans can. The validity and truth of this claim is questioned, however, because plenty of artists have come out in an attempt to defend their jobs from the recently booming AI generative art.

According to one article, Shutterstock, a popular high-quality image sharing website has started selling AI generated art (Thomas, 2022). The art has become so realistic and imaginative,

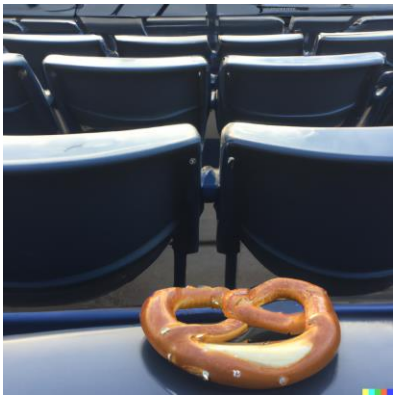
that companies like Shutterstock are profiting from its easy use. This bottom line proves that there is money to be made from AI art, regardless of ethics or other considerations. A podcast from *New York Times* debates whether Generative AI has the artistic talent to get out of control or not. In the podcast, the guest speaker Emad Mostaque, who is the creator of the AI program Stable Diffusion, responds to the question of "Where do you think this will be five years from now?" by saying, "[…] It'll be perfectly high resolution, no issues whatsoever." I decided to view certain AI programs firsthand to find out how unique they can be. While signing up to access *Dall-E,* an AI art fabrication program, I had to make sure I clicked a box that said, "Verify you are a human." Some prompts turned out to be exceptionally good images in terms of artistic value and composition, and I could not find any one of them through reverse Google search. While this isn't definitive, inhuman technology has the capability to convince the world of its computing power, and it shows how much room it has for growth. See some examples below.

There are certainly some things that AI cannot do; however, it is hard to argue against the unforeseen changes that I mentioned previously. Deepfake scams, AI chatbots, and changes to legislation regarding robotic companions were all inconceivable a few decades ago, but they have become a force to be reckoned with in the current day, and it will stay that way for the foreseeable future. I went into this writing thinking about how deadly AI can be to human civilization, and how if we are not careful, we could push mechanical systems past our reach. After hours of research, I see this as much less possible of an outcome; however, it is still a possibility for the future. If we do not take actions as a society to limit our use and development to AI, whether it be through regulations on computing power or accuracy of deepfakes, it can very easily get out of hand and take over our lives as we know them today. While I don't want a

droid to eat my food for me, I can see the benefits to handing some of our workload over to AI. AI needs to be kept in check, through government intervention and careful handling by the public, at the very least to keep AI at the level of hearing our voices and not destroying planets.

1. “Pretzel on a seat in a baseball stadium no one else is there” 2. “Cars in a line painted like a rainbow portrayed as a watercolor painting” 3. “Painting of a magical forest with a purple sky and blue colored grass and the trees are 15 feet tall in the style of Claude Monet”

-- (Sourced from DALL-E 2)



References

Agomuoh, F. (2023, March 30th). ChatGPT: how to use the AI chatbot everyone's talking about. *Digital Trends*. https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot

Boyd, C. (2021, April 16th). Deepfakes were going to change everything. And then they didn't. MalwareBytes Labs. https://www.malwarebytes.com/blog/news/2021/04/deepfakes-were-going-to-change-everything-and-then-they-didnt

Crash Course Computer Science. (2017, October 18th). Hackers & Cyber Attacks: Crash Course Computer Science #32 [Video]. YouTube. https://www.youtube.com/watch?v=_GzE99AmAQU

Etienne, V. (2021, July 19th). What to Know About the Controversy Surrounding Anthony Bourdain's A.I. Voice Used in 'Roadrunner'. *People Lifestyle*. https://people.com/food/what-to-know-about-anthony-bourdains-a-i-voice-controversy-in-new-doc-roadrunner/

Feeney, M. (2021, March 1st). Deepfake Laws Risk Creating More Problems Than They Solve. *Regulatory Transparency Project of the Federalist Society*. https://rtp.fedsoc.org/paper/deepfake-laws-risk-creating-more-problems-than-they-solve/

Klein, C. (2023, March 6th). "THIS WILL BE DANGEROUS IN ELECTIONS": POLITICAL MEDIA'S NEXT BIG CHALLENGE IS NAVIGATING AI DEEPFAKES. *Vanity Fair*. https://www.vanityfair.com/news/2023/03/ai-2024-deepfake

Madry, A. (2023, March 8th). Hearing: "Advances in AI: Are We Ready For a Tech Revolution?" *Cybersecurity, Information Technology, and Government Innovation Subcommittee.* https://oversight.house.gov/wpcontent/uploads/2023/03/madry_written_statement100.pdf

McClelland, C. (2023, January 30th). The Impact of Artificial Intelligence - Widespread Job Losses. *IoT For All*. https://www.iotforall.com/impact-of-artificial-intelligence-job-losses

Meisner, B. (2022, December 9th). Is AI-Generated Art 'True Art'? Implications And Considerations For Businesses. *Forbes*. https://www.forbes.com/sites/forbesbusinesscouncil/2022/12/09/is-ai-generated-art-true-art-implications-and-considerations-for-businesses/?sh=1598366a3013

Metz, R. (2019, June 12th). The fight to stay ahead of deepfake videos before the 2020 US election. *CNN Business*. https://www.cnn.com/2019/06/12/tech/deepfake-2020-detection/index.html

Newman, L.H. (2023, March 15th). AI-Generated Voice Deepfakes Aren't Scary Good—Yet. *Wired.* https://www.wired.com/story/ai-voice-deep-fakes/

Ohio University. (2023). *Fake News, Misinformation, & Fact-Checking*. https://onlinemasters.ohio.edu/masters-public-administration/guide-to-misinformation-and-fact-checking/

OpenAI. (2023). *Dall-E 2*. https://openai.com/product/dall-e-2

Quach, K. (2023, March 11th). Thanks to generative AI, catching fraud science is going to be this much harder. *The Register*. https://www.theregister.com/2023/03/11/ai_scientfic_fraud/

Roose, K. (Host), Newton, C. (Host), Mostaque, E. (Guest). (2022, October). Generative AI is Here, Who Should Control it? [Audio podcast]. https://open.spotify.com/episode/1C6LhuMgaO1hHiblOAKLBw

Rajnerowics, K. (2022, December 19th). Will AI Take Your Job? Fear of AI and AI Trends for 2023. *Tidio*. https://www.tidio.com/blog/ai-trends/

The Majority Report W/ Sam Seder. (2023, February 27th). *Can AI Deep Fakes Be Stopped*? [Video]. YouTube. https://www.youtube.com/watch?v=JVMd9SQr4Lw

Thomas, R. (2022, November 22nd). Wait, AI art is good now? *The Face*. https://theface.com/culture/wait-ai-art-is-good-now-artifical-intelligence-technology-dall-e-life

Verma, P. (2023, March 5th). They thought loved ones were calling for help. It was an AI scam. *The Washington Post*. https://www.washingtonpost.com/technology/2023/03/05/ai-voice-scam/