2023

Original: 03/18/2022

3rd Revision: 04/14/2023

By

Jesse Walter

Deepfakes, while problematic, also offer unique learning opportunities

**Deepfakes, while problematic, also offer unique learning opportunities.**

When we hear or read of a ‘deep fake’, also written as ‘deepfake’, it is either in the form of entertainment or some negative connotation such as the case with U.S. Speaker of the House, Nancy Pelosi, and her ‘slurred speech’ video. This video, dubbed a deep fake, which was released in May of 2019 onto Facebook, Twitter, and YouTube, had over a million views on the first day. The original video has since been removed, but many references can be found in media postings, like the [Washington Post’s](https://www.youtube.com/watch?v=sDOo5nDJwgA&t=2s) coverage shared via it’s YouTube channel. Over the last few years, several deep fake videos have entertained us, brought actors back from the dead, and even exposed us to the potential misuse of the technology. One popular warning example, which predates the Pelosi video, is that of the BuzzFeed (BuzzFeed, 2018) [video](https://youtu.be/cQ54GDm1eL0), with Jordan Peele, “You Won’t Believe What Obama Says In This Video!”. This video has a dual meaning while it’s supposed to be humorous, it also demonstrates both the ease and sophistication of technological deception. The purpose of this paper is to not deter the use of deepfake technology, but rather to understand its use in improving learning and even understanding.

Radio can easily be taken for granted today. It is an old technology and there are many other ways to listen to music or communicate. While the origins of radio date back to the 1800s, there is still some debate over who invented the first radio, Nikolai Tesla or Guglielmo Marconi(Radio Fidelity, 2019). The impact of the technology’s use, by Orson Welles created chaos in 1938, when a dance music program was interrupted with an adaptation of H.G. Wells’ *The War of the Worlds* The interruption informed the public of scientists who observed hydrogen explosions emanating from the surface of Mars. A variety of scientists were mentioned in the authentication process of the announced observation. The program was further interrupted to inform the public that aliens had invaded earth. Eventually the hysteria calmed down and fortunately radio technology continues to inspire innovations today.

Like those who fell under a hysterical spell during the radio broadcast of (Rosenberg, 2019)War of the Worlds, so have many people today been fooled by deepfakes. The dangerous uses afforded by deepfake technology will be addressed as the users and the tech matures. Emerging AI deepfake technology is changing the world, in ways which extend beyond propaganda or political misinformation. Current Deepfake technologies, as well as the understanding around their many uses, are rapidly evolving and like the radio will mature with society.

Deepfakes have been around for a long time, but the term *deepfakes* is more recently “The term “deepfake” was first coined in late 2017 by a Reddit user of the same name. Somers, M. (2020, para. 10)” In 1997, Christoph Bregler, Michele Covell, and Malcolm Slaney created the first system which fully automated the manipulation of existing video footage for facial reanimation. “The program was originally intended for applications in movie dubbing, enabling the movie sequence to be modified to sync the actors' lip motions to a new soundtrack.” In other words, the technology is intended to translate a movie from one language to another, which automatically alters the images to align the actors’ lip movements with the corresponding alternate audio track of in another language. A more simplified understanding for the term deepfake states “an image or recording that has been convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said (Merriam-Webster, n.d., para. 1).” In other words, if a person manipulates a photo by adding a family member or friend who was not actually present on that occasion, then they have created a deepfake. Merriam’s definition, while simplified without all the artificial intelligence lingo, makes it clear that deepfakes are also referencing the use of more sophisticated software such as image manipulation with tools such as Adobe’s Photoshop or the open-sourced software, Gimp. Recently, the Department of Homeland Security released and published the PDF document, *“Increasing Threat of Deepfake Identities Addendum – Examples”*, which currently categorizes deepfakes as either the following, Puppet Deepfake, Mouth Swap Deepfake, Synthetic Media, Audio Deepfake, and Unknown. The term *synthetic* is an accepted shorten version of *synthetic media* which “includes all types of media (text, image, voice, video) created partially or fully by AI algorithms” (Riparbelli, 2023, “A short definition of synthetic media,” para. 1). There are a few digital tools to aid the most novice of content creators such as Reface, developed by a Ukrainian company, this “is an easy-to-use deepfake app that allows you to make deepfake videos and gifs in one tap” (Salia, 2021, para.12). In other words, a person my use this application to take an existing video, say from their favorite movie and replace the actor’s face with their own. The frequently asked questions section of the article answers the question, confirming that deepfakes are a form of artificial intelligence, stating “Yes. Deepfake technology uses deep learning and artificial intelligence to create convincing but entirely fictional photos from the input. Generally, deepfakes refer to any videos in which the face or voice is swapped with someone else’s” (Salia, 2021, FAQ Section, para. 2). In other words, deepfakes are not necessarily created using artificial intelligence, and rather than a human painstakingly editing each of the 7,200 frames (at 60 frames per second) within a 2 minute video, with the use of a video editor such as Adobe After Effects, the artificial intelligence automates this process to complete it quickly and often just as accurately as a human.

Furthermore, rather than calling out the good, the media utilize their megaphones to inform the public that deepfake content is leveraged for insidious purposes. We see such reflections in headlines from Creative Blog’s author Joseph Foley (2023) which reads “20 deepfake examples that terrified and amused the internet” or even more recently FoxNews’ *Russia, Ukraine 'deepfake' video, audio are concerns for US intelligence* (Spunt, 2022). One ethically questionable example of a deepfake, which hit the social media headlines, included an Instagram video posting, by Bill Posters, mocking Facebook’s, now Meta, CEO Mark Zuckerberg. In this deepfake video the visage of Mr. Zuckerberg says the following, “Imagine this for a second, one man with total control of billions of people’s stolen data. All their secrets, their lives, their futures. I owe it all to Spectre. Spectre showed me that whoever controls the data, controls the future” (Posters, 2019). *Spectre*, mentioned in the video, is referencing an *“art installation”* or more simply put, an art project. This art project was designed, according to the (Department of Homeland Security), with the “aim to demonstrate how technology can be used to spread mis/disinformation (pg. 12, para. 1).” Instagram chose to leave this deepfake posting of Mark Zuckerberg up stating “We will treat this content the same way we treat all misinformation on Instagram. If third-party fact-checkers mark it as false… we will filter it from Instagram’s recommendation surfaces like Explorer and hashtag pages” (KPIX CBS SF Bay Area, 2019*)*. In other words, Instagram allowed the video to remain, but branded it as a deepfake by simply applying a hashtag label to it, #deepfake.
Mainstream media misses out on the opportunity to reinforce the shared positive life changing experiences afforded by this type of AI. Outside of the entertainment and political agendas, such as videos depicting a drunk (former) Speaker of The House, Nancy Pelosi, these perceivably deceptive technologies could be used to teach, improve, and preserve humanity through new educational experiences. Geraint Rees, Professor of Cognitive Neurology at the University of College London, defended the neutrality of artificial intelligence. In a World Economic Forum article, Rees (2019) stated, using healthcare examples, “AI is an enabling technology that is neither intrinsically good nor evil. Technology, like this depends on the context in which we create and use it.” (para. 8). In other words, AI is a tool and the distinction between good and evil lies solely in the intent of the user. Professor Rees’ comments reverberate similarly to the misunderstanding around guns. The gun is not evil. It’s not the gun which kills another life. It is the person using the gun who performs the action of killing. Bans on guns aboard aircraft did not deter the 9/11 hijackers who took the aircraft by force with box cutters. Take away one weapon, and a new weapon will be formed because the intent and the person driving that intent is the same. Once we mature our thinking around the uses of deepfake technology, then we witness the more wonderous side of the technology.

In the Superman lore, the Man of Steel uses an alien crystal technology to engage in a conversation with the visage of his long dead father. We do not have to wait for the arrival of alien crystal technology to provide us humans with interactive life-like visages of past members of society. Currently, through the aid of deepfake technology, people can hold conversations with the essences of the past. Deepfake AI has also enabled us to hear John F. Kennedy’s voice give a speech he wrote, but due to his untimely death was never able to deliver. People can learn about art from dead masters, like Salvador Dali. Albert Einstein will answer questions about his work and life. Through the aid of deep machine learning and other deepfake technologies, the likenesses, mannerisms, and in some cases the essence or synthetic variation of these individuals have been brought to life for the purpose of teaching through means of storytelling and personal engagement.

To explain one example, CereProc, a company which creates text-to-speech solutions, was charged with the recreation of an authentic John F. Kennedy speech voice. The purpose of this challenge was to leverage the power of the Deep Neural Network AI to generate JFK’s voice to be used in delivering the last written speech he never gave. Fifty-five years later in *The Times*’ *JFK Unsilenced,* the essence of JFK’s voice delivered the speech which reached one billion people.

For a wider audience, the audio techniques used to resurrect JFK’s voice have since been adopted by many companies to help sufferers of amyotrophic lateral sclerosis, a neurodegenerative disease which affects nerve cells in both the brain and spinal cord, re-find their voices*.* “With help from the Motor Neuron Disease Association and companies including Intel, Microsoft, and Rolls-Royce developed a new tool called Quips. It uses artificial intelligence in combination with a technique called voice banking to learn the unique way its user talks before they've lost the ability to speak (Bonifacic, 2019, para. 2).” In other words, large corporations invested money to develop software which leverages AI to create a personalized synthetic voice based on hundreds and thousands of their own voice samples, which are stored in a digital bank. People with diseases like ALS and Parkinson’s benefit from having a digital voice when their own fails.

 In another example of positive uses for deepfakes, May 11, 2019, was not only Salvador Dali’s 115th birthday, but also his publicly released digital resurrection in the “Dali Lives (via artificial intelligence)” exhibit in St. Petersburg, Florida. Visitors “have the opportunity to learn more about Dalí’s life and work from the person who knew him best: the artist himself” (The Dali, 2019, para: 2). Of the synthetic resurrections, this artistic reflection of life is both an uncanny and exquisitely designed digital persona. Nathan Shipley, the technical director, explained the process of creating this digital-Dali. Over 6,000 frames from old film footage were analyzed by the AI system and it took over one thousand hours of machine learning to process the data. The *Dali Lives* exhibit is located at *The Dali Museum*, in St. Petersbug, Florida. It is here where several of Salvador Dali’s life size visages reside within large screens housed an interactive panels. While seemingly confined within the paneled display, it only takes one touch of a button for Dali’s visage to come alive in a fully reanimated form. The AI behind this deepfake analyzed over 6,000 frames of original Dali footage to generate his visage and provide over 190,500 possible combinations of interactions with visitors. This means virtually every experience with the Dali’s visage is different. Through these dynamic experiences, visitors learn about Dali’s art and his life from the artist himself.

 Another company specializing in AI-powered digital humans, also known as synthetic media or *synthetics* (shorten term), is Uneeq. “In partnership with the Hebrew University of Jerusalem and Greenlight Rights, we brought the great scientist back to showcase his life and work to a whole new generation of people” (Uneeq, n.d., para: 1). [Digital Einstein](https://www.digitalhumans.com/case-studies/digital-einstein)'s creation also showcased the possibilities of engaging learning experiences generated through AI. The synthetic’s mannerisms are based off the Nobel Prize recipient, Albert Einstein. This synthetic was available to engage with people over the internet, but is now available as a demonstration, which may be [booked through Uneeq’s website](https://www.digitalhumans.com/book-a-demo). I was fortunate enough to experiment with Digital Einstein demo and was impressed at that time by the voice recognition AI capabilities. I was able to ask a variety of simple questions, by talking into my computers’ microphone and which the Digital Einstein on my screen would animate and respond appropriately to the question asked. When engaging with this AI synthetic, you may ask Digital Einstein questions about science or about the real Einstein’s life. As interesting as Digital Einstein is, Uneeq has ambitions which extend beyond the visage of Einstein.

To explain, Uneeq rose out a program which provided access to information for persons with disabilities. This was achieved through the creation of online human-to-digital-human speech. The team quickly recognized the mental health benefits in using synthetics in the medical field. People dealing with mental health struggles are more comfortable engaging with syntheticsover humans. Synthetics will not judge the human sharing their mental or medical information. Synthetics can give a voice to someone who no longer has a voice. Preserving and teaching history through artificial intelligence, in the holographic visage of those who experienced it long after they are dead, may seem like something out of a comic book or science fiction film, but the USC Shoah Foundation’s has done with the *Dimensions in Testimony* project as it “enables people to ask questions that prompt real-time responses from pre-recorded video interviews with Holocaust survivors and other witnesses to genocide.” Most often visitors will experience the synthetic by interacting with a life size screen embedded into a panel, similar to a kiosk. These flat screens visages certainly look beautiful and function quite well, however the animation *blending* could use some work. Blending is a term used in gaming and film industries, which literally refers to blending the animation sequences where one ends and the next begins. As a character animator, the blending bothered me, but the general audience would most likely not think twice about it. If the exhibit is showing the hologram versions of the survivor, then this is a must to experience. The quality of the holographic representation of each Holocaust survivor rivals modern Hollywood digital effects and the exhibit does not disappoint in modern technical savviness. As museum visitors gather in an auditorium, they are presented with a hologram of an elderly person sitting center stage. Like magical effects out of George Lucas or Steven Spielberg film, the hologram, seemingly alive, communicates and interacts with its audience. The hologram is that of an older woman and she is so seemingly real in appearance, actions, and sound that at first one believes they are interacting with a live person. It is as if Hollywood truly captured the “essence” of this woman from the cracks in her speech to secondary gestures. The audience members may ask questions about the essence’s life and the audience receives a genuine response which projects a real recorded emotion of the individual.

To achieve the Hollywood level of effects in this project, it required one of the brightest minds out of Hollywood. With the help of Steven Spielberg's masterful story telling and the next generation of natural language processing, Dimensions in Testimony was able to achieve this interactive biography of Holocaust survivors. This technique is redefining inquiry-based education. ​​​​“Inquiry-based teaching focuses on moving students beyond general curiosity into the realms of critical thinking and understanding. You must encourage students to ask questions and support them through the investigation process, understanding when to begin and how to structure an inquiry activity” (Prodigy, 2017, para. 4). In other words inquiry-based learning incorporates engagement with the students, which in this case happens to be the audience, and the process of asking questions helps support critical thinking processes. As I think about this, it seems clear to me that there are layers of Connectivism Learning Theory (CLT) at hand here. With the CLT approach learning occurs through the collaboration and sharing of idea, opinions, and views. Ashish Jaiman, Director of Product at Microsoft Bing, also champions for positively motivated cases where deepfakes, such as AI-Generated synthetic media *“*have clear benefits in certain areas, such as accessibility, education, film production, criminal forensics, and artistic expression” *(*Jaiman, 2020, para. 3*).* Microsoft and Google are both leveraging AI for synthetic voice for a multitude of narrated devices, ranging from GPS navigation to fully interactive digital synthetic assistants. These synthetics provide opportunities for automated educational experiences, which will not only preserve historical accounts from people, but also their essence. So rather than reading about the historical events and the people who were affected by those events, the learner may engage in conversation with the human’s essence or synthetic media which has been created after the human replicated. This leveraging of technology ironically brings the learning process back into the old ways of tribal storytelling, when people huddled around the campfire while the elders tell their accounts and interpretations of events. Perhaps, one day in the future, history lessons will be taught by asking the AI questions and the learner will receive the information in a dazzling spectacle of a lifelike holographic long dead great-grandparent. Like the essences in the Dimensions in Testimony, your holographic great-grandparent answers queries of nearly forgotten ancestral recipes or the details about the era they lived in.

In another example, Harvard graduate and co-founder of Edyoucated & TechLabs, David Middelbeck, championed the need for change in the fundamentals of learning as societies adapt to these emerging and disruptive technologies. During his TedTalk Re-inventing Education for the Digital Age, Middlebeck describes the pattern of technological growth as a race between technology and education. “Whenever technology raced ahead, it first leads to social pain and inequality until at some point in time the whole educational system gets turned upside down in order to keep up” (TEDx Talks & Middelbeck, 2019).In other words, when new technology arrives, sometimes it is misused until understood and that often there is a divide between those who have access to the technology and those who do not. He draws parallels to the emergence of the printing press and the effects it had on society to the disruptive technological trends of modern times. As Chairman of TechLabs, a European nonprofit organization, Middelbeck’s vision to bring technology to education is digitally reaching and helping hundreds of students learn coding every year for free. TechLabs’ mission “aims to equip young individuals with tech domain expertise as well as methodical and soft skills that are highly relevant now and in our (more and more automated) future workforce” (TechLabs, 2022, para 4). In other words, this means TechLabs understands that as AI automation is more prevalent in the work domain, humans will be displaced if they do not have skills which support the needs. TechLabs provides an educational path to learn about these emerging technologies and gain the required skills to leverage them.

 Furthermore, AI tools can be used as swords or plowshares; the choice is left up to humanity to figure out. In 2019, the online learning platform, *Udacity*, which has over 100,000 courses, took a step towards turning the sword into a plowshare. Producing Massive Open Online Course (MOOC) content for a demographic variety of students is very time consuming, especially for videos, so the Udacityteam is *“*investigating a new machine learning framework that automatically generates lecture videos from audio narration alone*”* (Griffin, 2019, para. 2). Udacity’s utilization of AI automates and accelerates learning by converting these lectures into synthetic instructors. Udacity’s researchers, who work on synthetic AI, are aware of the potential misuses of the AI technologies and hope to catalyze advanced developments in the deep learning technologies to produce video content. Understanding and discovering the effects of long term deepfake technology and its usage will take time understand as the technology matures.

Consequently, figuring out the effects of Deepfake technology on education is what the College of Engineering, at “Carnegie Mellon University, in collaboration with RAND corporation and Challenger Center seeks to spend the next year on a National Science Foundation-funded project investigating the impact of deepfakes and the ability to fool student learners of different populations” (College of Engineering, Carnegie Mellon University, 2021). The demographics under study include higher education, K-12, and adult learners. It will be interesting to see what results from the research.

 Concluding thoughts, I’d like to share that I was not thoroughly on board with the idea of AI driven learning. Much of my concern was due to lack of exposure to the positive and progressive uses of AI. As we gain knowledge in AI, we are either strengthened in our views or we further our understanding which enables us to change perspectives. From chatbots, androids, or synthetic personalities like the librarian from the Time Machine movie or HAL (Heuristically programmed Algorithmic computer) from A Space Odyssey, what was once science fiction is now a reality. The essence of human experiences are more than static photos or words in a textbook. Education has adopted many technologies over the years; AI is no different. Over the course of writing this paper, I have turned my skepticism into positive hope with this technology. These technological tools are currently driving disruption in our society, but they are also lighting many new paths to more engaging and personalized learning experiences. With the right vision and proper stewardship, artificial intelligence technology will lead the way to bridging learning experiences and creating common understanding between the many peoples of the world. So rather than ending in a story where AI destroys humanity, perhaps it may just save humanity and bring us together instead.

References:

BuzzFeedVideo. (2018). You Won’t Believe What Obama Says In This Video! In *YouTube*. <https://www.youtube.com/watch?v=cQ54GDm1eL0>

Bonifacic, I. (2019, December 18). Quips is an AI to help ALS patients speak with their own voice. Engadget. Para: 2. <https://www.engadget.com/2019-12-18-rolls-royce-quips-als-mnd-speech-ai.html>

CereProc. (2018). JFK Unsilenced | CereProc Text-to-Speech. Www.cereproc.com. <https://www.cereproc.com/en/jfkunsilenced>

College of Engineering, Carnegie Mellon University. (2021). The Effect of Deepfake Technology on Education. In *YouTube*. <https://www.youtube.com/watch?v=ocqH1bOGdhI>

Department of Homeland Security. (n.d.). *DEEP Increasing Threat of FAKE Identities ADDENDUM -EXAMPLES.* In Department of Homeland Security (pp. 1–19). Retrieved March 12, 2022, from <https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_addendum_0.pdf>

DHS. (2021, October 26). *AEP Deliverables | Homeland Security*. Www.dhs.gov; Department of Homeland Security. <https://www.dhs.gov/aep-deliverables>

Foley, J. (2023, March 10). *20 deepfake examples that terrified and amused the internet*. Creative Bloq; Creative Bloq. <https://www.creativebloq.com/features/deepfake-examples>

Griffin, M. (2019, August 2). *Edtech company Udacity uses deepfake tech to create educational videos automatically*. By Matthew Griffin Futurist and Keynote Speaker. Para: 2, <https://www.fanaticalfuturist.com/2019/08/edtech-company-udacity-uses-deepfake-tech-to-create-educational-videos-automatically/>

Jaiman, A. (2020, August 14). *Positive Use Cases of Deepfakes*. Towards Data Science. <https://towardsdatascience.com/positive-use-cases-of-deepfakes-49f510056387>

Kelly, M. (2019, May 24). *Distorted Nancy Pelosi videos show platforms aren’t ready to fight dirty campaign tricks*. The Verge. <https://www.theverge.com/2019/5/24/18637771/nancy-pelosi-congress-deepfake-video-facebook-twitter-youtube>

KPIX CBS SF Bay Area. (2019, June 12). *Artists Create “Deepfake” Video Of Mark Zuckerberg To Gauge Facebook Response*. Www.youtube.com. <https://www.youtube.com/watch?v=yV7sR5jpQNs>

Lee, D. (2019, May 10). *Deepfake Salvador Dalí takes selfies with museum visitors*. The Verge; The Verge. <https://www.theverge.com/2019/5/10/18540953/salvador-dali-lives-deepfake-museum>

Merriam-Webster. (n.d.). *Definition of DEEPFAKE*. Www.merriam-Webster.com. Retrieved March 5, 2022, from <https://www.merriam-webster.com/dictionary/deepfake>

Posters, B. [@Bill\_Posters\_uk]. (2019, June 7). *I’m very pleased to announce that the BIG DADA / PUBLIC FACES series of synthetic media artworks are now available to* [Video]. [Instagram profile]. Instagram. Retrieved April,13,2023, from <https://www.instagram.com/p/ByaVigGFP2U/>

Prodigy (2017, January 19). *What Is Inquiry-Based Learning: 7 Benefits & Strategies You Need to Know.* Prodigy. Para: 4. <https://www.prodigygame.com/main-en/blog/inquiry-based-learning-definition-benefits-strategies/>

Radio Fidelity. (2019, May 2). *When Was Radio Invented? How Radio Changed The World*. Radio Fidelity. <https://radiofidelity.com/when-was-the-radio-invented/>

Rees, G. (2019, November 25). *Here’s how deepfake technology can actually be a good thing.* World Economic Forum. Para: 8. <https://www>.weforum.org/agenda/2019/11/advantages-of-artificial-intelligence/

Riparbelli, V. (2023, March 8). *The Future of [Synthetic] Media.* Synthesia. “A short definition of synthetic media” Para: 1. <https://www.synthesia.io/post/the-future-of-synthetic-media#what-is-synthetic-media>

Rosenberg, J. (2019, June 3). *How the “War of the Worlds” on the Radio Caused a Panic*. ThoughtCo. <https://www.thoughtco.com/war-of-the-worlds-radio-broadcast-1779286>

Salia, R. (2021, January 19). *Top 10 Deepfake App & Software for 2022 | Deepfake Elon Musk is Possible*. TopTen.AI. Para: 12. <https://topten.ai/deepfake-app-and-software-review/>

Somers, M. (2020, July 21). *Deepfakes, explained.* MIT Management Sloan School. Para: 10. <https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained>

Spunt, D. (2022, February 23). *Russia, Ukraine “deepfake” video, audio are concerns for US intelligence*. Fox News. <https://www.foxnews.com/tech/russia-ukraine-deepfake-video-audio-us-intelligence>

TechLabs. (2022). *About the Organization | TechLabs*. About the Organization | TechLabs. Para: 4 <https://techlabs.org/about>

TEDx Talks, & Middelbeck, D. (2019). Re-inventing Education for the Digital Age | David Middelbeck | TEDxMünster [YouTube Video]. In *YouTube*. <https://www.youtube.com/watch?v=ArI6albrkuY>

The Dali. (2019, January 23). *Dalí Lives: Museum Brings Artist Back to Life with AI*. Salvador Dalí Museum. <https://thedali.org/press-room/dali-lives-museum-brings-artists-back-to-life-with-ai/>

UCL. (2017, November 15). *Geraint Rees*. Institute of Cognitive Neuroscience. <https://www.ucl.ac.uk/icn/people/geraint-rees>

USC Shoah Foundation. (2019). *USC Shoah Foundation*. USC Shoah Foundation. <https://sfi.usc.edu/dit>

Wikipedia. (2022, March 5). *Deepfake*. Wikipedia. <https://en.wikipedia.org/wiki/Deepfake#:~:text=Deepfake%20technology%20has%20been%20developed>