مجلة الحقوق

فصلية علمية محكمة - تصدر عن مجلس النشر العلمي - جامعة الكويت

تقنين تقنية التزييف العميق .

الدكتور/ عبدالعزيز بن صالح بن عبدالله العبود

جامعة الكويت
KUWAIT UNIVERSITY

P-ISSN: 1029 - 6069

E-ISSN: 2960 - 2742

العدد ٣ - السنة ٤٨

ربيع الأول ١٤٤٦ هـ - سبتمبر ٢٠٢٤م

JOURNAL OF LAW

A Refereed Academic Quarterly, Published by the Academic Publication Council - University of Kuwait

Taming the Faceless Beast: Legislating Deepfake Technology.

Dr. Abdulaziz Saleh Alobud

جامعة الكويت
KUWAIT UNIVERSITY

P-ISSN: 1029 - 6069

E-ISSN: 2960 - 2742

No. 3 - Vol. 48

Rabi I 1446 - September 2024

Taming the Faceless Beast: Legislating Deepfake Technology

Dr. Abdulaziz Saleh Alobud*

Abstract

Objective: To deal with the potential deepfake abilities demonstrated by its diverse applications, legal systems have the choice to block this technology completely, allow it and deal with its issues later, or legislate it. Chinese legal system, for example, chose to ban the technology and its applications. If the choice was legislating deepfake, then the question will be what is the most suitable legal ground to use? **Methodology:** To answer this question, legislators should consider the national laws and the culture of any given legal system by using analytic methodology to the current legislative. **Results:** Furthermore, it appears that one legal ground is not enough to legislate deepfake technology due to the vast applications of this technology. Also, the ideology of legal principles should be mature enough in order to be used as a legal ground when legislating deepfake. **Conclusion:** if the copyright law is not advanced in certain legal system, then that legal system cannot legislate deepfake based on copyright legal ground or the results will be insufficient. The deepfake legislating process should not imply prohibiting the technology; in fact, legislating deepfake can boost certain useful application.

Keywords: legislation, deepfake(s), deep learning, artificial intelligence.

* Assistant Professor at the Faculty of Law and Judicial Studies, Law Department.- Islamic University of Madinah, Saudi Arabia, Medina.

Email: Abdulaziz.Alobud@iu.edu.sa

- Submitted: 28/1/2021, Accepted: 11/4/2021.

To Cite P. 507

Introduction

A computer machine that can solve problems and learn from itself is said to have Artificial intelligence (AI). The science of AI is not about programming code to solve certain problems, instead, the goal is to give the machine the function of mimicking the human cognitive ability to address a problem and produce a solution within the general limitations given by the programmer. One of these human abilities is learning from experiences and mistakes, which is an ability called Machine learning. Machine learning is artificial intelligence, but with the ability to learn to function beyond the original coding or its operating system. Deep learning, on the other hand, is machine learning, but the learning process happens through a technique that mimics the behavior of the neuron network of the human brain.

The beginning of AI was merely a theory introduced in the academic work of Charles Babbage.(1) Another theory presented by Claude Shannon suggested that computers could play chess competitively.(2) A few years later, Shannon joined other researchers at an artificial intelligence convention at Dartmouth University in 1956, which is now considered a milestone in the history of AI.(3) AI science is continued to grow with the rapid development in computer engineering and computer science. Today, almost no company does not use some sort of AI in its business.

For the end-users, AI has been behind many technologies that they are unaware exists. One of the technologies used daily is Speech Recognition, which allows an end-user to speak his words to be written onto a computer or applied to a variety of other useful forms. Another useful technology is the virtual agents. Virtual agents exist in many forms to accommodate end-users. One is the chatbot that helps

(1) Maad M & Mijwil, History of Artificial Intelligence (2015). Available at https://www.academia.edu/40077427/History_of_Artificial_Intelligence (last visited May 2, 2020).

(2) *Id.*

(3) Luca Sgambi, Artificial Intelligence: Historical development and applications in civil engineering field (2008). Available at https://www.academia.edu/2878509/Artificial_Intelligence_Historical_development_and_applications_in_civil_engineering_field (last visited May 2, 2020).

customers on almost every service provider's website. Another form is the cell phone and platform assistance devices such as Alexa, Cortana, Siri, and Bixby. Other products of AI are for business entities, such as Biometrics and Robotic Process Automation (B & RPA).(4)

One of the most ambitious AI's technology is deepfake. Deepfake is often considered a deceptive technology in which photographs, voices, and facial expressions mimic authentic media. Although it began as a very basic function, it has developed and now can create a fictional speech that is delivered using the voice and the face of anyone. AI was developed to study and analyze the voice pattern, body language, and social behavior of an individual. Currently, AI just needs a photograph of the subject to create a speech or an act using the available information and filling out any missing details. AI itself is taking a different approach when creating deepfake media. The new approach is proceeding the information in a neuron network similar to human's brain neuron network.

Deepfake technology is available to anyone with an average knowledge of technology. Furthermore, since this technology develops more every day, it is anticipated that the tools for making deceptive media will become more user-friendly and more common. Currently, the application of deepfake technology is used in political media, art performances, entertainment, and is limited to the faces, and voices of celebrities.

Today, deepfake is advanced enough to do live face swapping.(5) These frequent advances in technology have aggravated many institutions, companies, and governments. The technology could change the reality we know and destroy the credibility of the media. Law enforcement is likely

(4) It is an automation business process where AI conducting business on behalf of its programmer. See James R. Slaby, Robotic Automation Emerges as a Threat to Traditional Low-Cost Outsourcing, HFS Research (2012). Available at www.hfsresearch.com/Robotic-Automation-as-Threat-to-Traditional-Low-Cost-Outsourcing (last visited Jul 31, 2020).

(5) Yuval Nirkin, Yosi Keller & Tal Hassner, FSGAN: Subject Agnostic Face Swapping and Reenactment (2019). Available at http://openaccess. thecvf.com/content_ICCV_2019/papers/Nirkin_FSGAN_Subject_Agnostic_Face_Swapping_and_Reenactment_ICCV_2019_paper.pdf (last visited May 2, 2020)

going to face challenging times recognizing whether videos and voice evidence is real. Furthermore, the copyright and intellectual property laws are likely going to be violated if proper consent has not been obtained before the video, voice, or the photograph of the individual is used. The technology could be used to defame the reputation and integrity of an individual or group of people.

For all these reasons and more, the deepfake applications have to be governed by law. The legislators in any legal system must be aware of the implications of this technology and act accordingly, or they will confront serious after-effects of their use. Neither complete prohibitions nor poor legislation is preferable in any legal and political system. Thus, the question is not whether the legislators should act? The question is how?

Statement of the Problem and Research Question

The main question of this paper is on what legal ground should the deepfake be legislated? The legislators must take into consideration the current and anticipated applications of this technology to predict the legal issues at hand. Even with the easy approach, which is the prohibition of the whole technology, what legal ground should be used to justify that ban. Copyright infringement, intellectual property rights infringement, tampering with evidence, defamation, national security risks, and public policy violation are all probable legal grounds for legislating and controlling the use of deepfake.

This paper will address what regulatory process should be adopted, short of an outright ban, especially since deepfake incorporate different applications and usage of applications. A new approach is better than banning this technology because it wastes the potential of these useful applications. Reviews of whatever law is adopted will also need to be conducted in every legal system after legislating this technology to ensure the purpose of the law is served and the goals are met.

Hypothetically, a user appropriates the photograph of an individual, "Yahya," to create a video that shows him speaking, moving his lips, and giving facial expressions that resemble Yahya's appearance. The

creator then adds a voice that resembles the voice of Yahya on that video by mimicking the real voice of Yahya. Consequently, Yahya appears in the video saying whatever the creator wishes him to say in Yahya's voice. Yahya's right of publicity and copyright should be protected by the law, and if the creator has no consent from Yahya to use his voice or his photograph, then Yahya's rights would be considered violated and infringed. Now, if Yahya was a candidate for a public office or he was the head of the state, for example, then the legal ground to legislate against the use of a deepfake in this scenario would be different than in the previous example.

Paper Purpose

This paper aims to achieve several objectives, as follows:

1 - Present legal theories on the most appropriate legal grounds to legislate deepfake technology and its applications.

2 - Open the door to legal studies and law reviews addressing the enforcement of laws, the legal detection methods of deceptive media, and the admissibility of evidence in the deepfake era.

3 - Show the potential damages due to a lack of deepfake legislation.

Significance of Paper

This paper gains its significance from:

1 - The scarcity of legal research addressing deepfake technology and its applications. This is understandable since the technology is relatively new.

2 - Only two legal systems around the world have taken legislating steps for deepfake technology, which means that there is a long journey ahead of legislating deepfake.

3 - The lack of legislation means there will be more harmful practices of this very powerful technology.

4 - Without the support of media and the public, the legislators cannot convince the political powers in every given political system to

adopt these laws. Thus, this paper and others are one way to educate the public on this significant issue.

Literature Review

Few papers on the legal ramifications of deepfake technology have been published. The only published paper has been "Deepfakes and Human Rights Litigation in New Zealand."(6) This paper focused on New Zealand's legal system and the use of human rights litigation for a party who argued that he/she had been abused by the use of this technology. No Saudi or Arabic studies have been similarly conducted on the subject or the field.

Introductory Chapter
The History and Applications of Deepfake

The History of Artificial Intelligence (AI) and Deepfake

Charles Babbage, a mathematician and mechanical engineer, introduced the idea of computer science in 1822.(7) Although Professor Babbage did not create artificial intelligence, he introduced the scientific concept in his theoretical work.(8) Despite the support of Britain's government, his work remained theoretical because he could not build the intelligent calculator he described.(9) Some would argue that France's Joseph Marie Jacquard was the first to invent a technology that could be considered computer science almost 20 years before Babbage.(10)

It is true that the concept of Jacquard's technology was used in computer science and might have inspired Babbage himself, but

(6) Phuoc Nguyen, Deepfakes and Human Rights Litigation in New Zealand Academia.edu, https://www.academia.edu/39752667/deepfakes_and_human_rights_litigation_in_new_zealand (last visited May 4, 2020).

(7) Mijwil, supra note 2.

(8) See Copeland, B. Jack, The Modern History of Computing, The Stanford Encyclopedia of Philosophy (Winter 2017 Edition), Edward N. Zalta (ed.). Available at https://plato.stanford.edu/archives/win2017/entries/computing-history/. (last visited July 13, 2020).

(9) *Id.* His machine was built later on several occasions. The first actual machine based on the Babbage design is displayed in The London Science Museum under the name (Difference Engine No.2).

(10) *Id.*

Jacquard's technology was created to supply handloom machines with premade patterns that were "programmed" on wooden cards in a form of small punctures.(11) On the other hand, the Babbage technology was motivated by the idea of creating a machine using artificial intelligence.

The Jacquard's method was used by Herman Hollerith in 1880 to calculate the United States census.(12) Hollerith was able to save the government $5 million.(13) Hollerith's calculator was a machine that used some aspect of artificial intelligence used for computing tasks and was built. One the other hand, Jacquard's machine was closer to a mechanical machine than a computing machine due to its purpose and functions.

If the previous history is to be considered as an early stage of computing, Alan Mathison Turing is considered the father of a new computing era.(14) From his early theoretical work, the Turing Machine, his contributions in World War II, to his creation of the Bombe Machine was a declaration of his prestigious position in the science of computers and artificial intelligence. Modern computers are based on his theoretical work of the Turing Machine, which was capable of calculating and computinganythingcomputable.(15)

Turing tried to answer the question, "Can machines think?" which was the first step in creating artificial intelligence. He proposed a test, later known as the Turing Test, that argued that a machine was intelligent if the machine convinced you it is human.(16) Although Turing, in 1950, did not present a physical machine that proved his theory of "the

(11) See Douwe Draaisma & Paul Vincent, Metaphors of memory: a history of ideas about the mind, 146-47 (2001). See also Robert T. Futrell, Donald F. Shafer & Linda I. Shafer, 1 Quality software project management, 631 (1 ed. 2002). (They considered Jacquard's technology software not hardware).

(12) Frank da Cruz, Columbia University Computing History A Chronology of Computing at Columbia University, Herman Hollerith (2011). Available at http://www.columbia.edu/cu/computinghistory/ (last visited July 14, 2020).

(13) *Id.* Available at http://www.columbia.edu/cu/computinghistory/hollerith.html (last visited July 14, 2020).

(14) See Draaisma, supra note 12, at 149.

(15) *Id.*

(16) Alan M. Turing, Computing Machinery and Intelligence, LIX Mind 433–60 (1950). Available at https://doi.org/10.1093/mind/LIX.236.433.

imitation game," he did predict that such a machine would be created by the beginning of the 21st century.(17)

The development of computer science was boosted by the invention of the transistor,(18) computer chip,(19) Dynamic Access Memory chip (DRAM),(20) the floppy disk,(21) and the development of the first computer language.(22) All these inventions and developments led to creating the prototype of modern computer interaction tools that, supported by the graphical user interface and the computer mouse, made the computers more appealing to the general public.(23) In the meantime, the science of artificial intelligence benefited from the accelerating developments in computer science.

In 1950, Claude Shannon promoted a theory that computers could play chess competitively.(24) Shannon joined others a few years later and contributed to the artificial intelligence convention at Dartmouth University in 1956, considered a milestone in the history of AI.(25) Shannon's theory and later work are important because he had introduced the fundamental principle behind the first form of artificial intelligence computing, Deep Thought, and Deep Blue.(26)

(17) *Id.* at page 441.
(18) Invented by William Shockley, John Bardeen and Walter Brattain in 1947. See Somanathan B. Nair, Digital Electronics and Logic Design, 360 (2004).
(19) The integrated circuit or (computer chip) was invented by Jack Kilby and Robert Noyce in 1958. See *Id.* at 361-62.
(20) In 1970 by Intel. See Robert A. Burgelman, Strategy is Destiny: How Strategy-Making Shapes a Company's Future, 32 (2002).
(21) In 1971 by Alan Shugart and his engineering team at IBM. See Eben Upton et al., Learning Computer Architecture with Raspberry Pi, 246 (2016).
(22) In 1953 by Grace Hopper. Another computer language was developed in 1954 by John Backus and his team at IBM. See Ashok Singh, Science & Technology, 179 (2007).
(23) Invented by Douglas Engelbart in 1964. See Mark Hall, Douglas Engelbart, Encyclopædia Britannica (2020). Available at https://www.britannica.com/biography/Douglas-Engelbart (last visited Jul 14, 2020).
(24) Mijwil, supra note 2.
(25) *Id.* supra note 4.
(26) Deep Thought and Deep Blue were both IBM computers. They were programmed for mastering a Chess game and then tested against Garry Kasparov, the best Chess player then.

Since Deep Blue and the famous rematch in 1997, artificial intelligence has developed immensely. Today, artificial intelligence is so highly advanced that it can learn from its own experiences and mistakes. Thus, the learning machine does not need the human input to perform what seems to be an artificial intelligence functions; instead, this machine can learn from scratch. As a result, its ability to handle certain tasks has evolved based on the power of processing the learning machine has.(27) That means that a machine with artificial intelligence can develop a method to communicate with another machine without human help, which has already happened.(28)

The Technical Processes Behind Deepfake

The technology of deepfake is heavily reliant on machine learning processes. One of these processes is Deep Learning which is the based technology behind deepfake hence the complexity of the learning process. Since the results and applications of deepfake is not real, the term "fake" replaced "learning," thus, the term Deepfake.(29) Deep learning and deepfake is dependent on itself to collect the data from a large pool of information provided by the internet. The internet is the main source of big data(30) and is the distribution platform of this technology as well.

Therefore, we have artificial intelligence, machine learning, deep learning, and deepfake. All these terms are related but distinguishable.

(27) For example, Google›s AI AlphaGo Beats best GO player in 2017. The board game Go has 10360 possible outcomes. See Christof Koch, How the Computer Beat the Go Master, Scientific American (2016). Available at https://www.scientificamerican.com/article/how-the-computer-beat-the-go-master. (last visited Jul 14, 2020).

(28) Andrew Griffin, Facebook robots shut down after they talk to each other in language only they understand, The Independent (2018). Available at https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html (last visited Jul 14, 2020).

(29) Hubert Davis, How Deepfake Technology Actually Works, ScreenRant (2020). Available at https://screenrant.com/deepfake-videos-explained-how (last visited Jul 14, 2020).

(30) Troy Segal, The Deal with Big Data, Investopedia (2019). Available at https://www.investopedia.com/terms/b/big-data.asp (last visited Jul 14, 2020).

The deepfake is part of deep learning. Deep learning is an advanced application of machine learning. Machine learning is artificial intelligence. Artificial intelligence mimics the intelligence of human behavior or thinking.(31) Machine learning mimics the ability to think; however, this ability does not need the human to program a set of rules to be followed, because the machine can study the presented data and come up with new information.(32) Consequently, deep learning is machine learning, but the way it processes the information mimics the human brain's neural process. As a result, deep learning is also known as deep neural learning or networking.(33)

The deep learning applications are a way to study any data—the bigger the data, the better results—and produce solutions or applications that help with many industrial products and processes, such as self-driving cars, fraud detection, abnormal marketing behavior, healthcare, image recognition, or falsifying voices or videos. Now, falsifying voices or videos, or "deepfake," is a process in which a machine uses a constellation of artificial intelligence(34) to study a targeted voice, photograph, or a video containing the target's pictures. By studying the target's voice, the machine can understand its sound features and apply them to a new voice. The outcome will be a voice matching the target's voice. The same procedure is true for faking videos.

In short, there are three main methods used to create deepfake videos. The oldest method is based on 3D technology where the target face is rendered by a 3D algorithm and passed onto a 3D model. The second method uses Generative Adversarial Network technology

(31) Jake Frankenfield, How Artificial Intelligence Works, Investopedia (2020). Available at https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp (last visited Jul 14, 2020).

(32) Jake Frankenfield, Machine Learning, Investopedia (2018). Available at https://www.investopedia.com/terms/m/machine-learning.asp (last visited Jul 14, 2020).

(33) See Marshall Hargrave, How Deep Learning Can Help Prevent Financial Fraud, Investopedia (2019). Available at https://www.investopedia.com/terms/d/deep-learning.asp (last visited Jul 14, 2020), see also Sally Adee, What Are Deepfakes and How Are They Created?, IEEE Spectrum (2020). Available at https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created (last visited Jul 14, 2020).

(34) The algorithm that is used by deep learning is called generative adversarial networks (GANs) and is more complex than the constellation of AI used in deepfake. See *Id.*

(GANs) to render deepfake faces. The issues with this technology are its cost, speed, and stability. A third method is a facial manipulation using GANs to match distribution points on the target's face. The third method creates an attention map on the face of the target, then reenacted the target's new face to match human facial expressions.(35) What is important is that this technology will keep developing in its core and by new application methods.

On the other hand, creating deepfake voices requires a different approach. First, the technology uses a relatively old technology—speech synthesis with a text-to-speech converter (TTS). The outcome of this process is an artificial human voice. Second, the AI is trained to learn the way the target speaks and the wave features of his voice. This process is easier and faster if there are a large number of recorded speeches of the target. Third, the wave features of the artificial human voice are matched to the target's voice.(36) Adobe Inc. has a project called "Adobe VoCo" which it anticipates will be the voice photoshop. Its project is to develop a software program that is capable of creating a human voice and manipulating it digitally.(37) Adobe's project is not, however, the only project in the market. Several projects are being developed and at least one of them, WaveNet is going to be open-sourced.(38)

Current Deepfake Applications

One famous application of deepfake technology was used in a 2015 movie in which one of the main characters died in the middle of production. The producers of the movie had to recreate at least three main scenes in which the deceased actor was playing his role. The time,

(35) See Nirkin, supra note 6, at 7185.

(36) Ye Jia et al., Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis (2019). Available at https://arxiv.org/pdf/1806.04558v4.pdf (last visited Jul 14, 2020).

(37) Benjamin Powers, Adobe is Developing Photoshop for Your Voice, Medium (2018). https://medium.com/s/story/adobe-is-developing-photoshop-for-your-voice-f39f532bc75f (last visited Jul 14, 2020).

(38) Aäron van den Oord & Sander Dieleman, WaveNet: A Generative Model for Raw Audio, DeepMind (2016). Available at https://deepmind.com/blog/article/wavenet-generative-model-raw-audio (last visited Jul 14, 2020).

the logistics, and the cost of creating such clips in 2015 were long, complex, and high. A few years later in 2019, a few videos began to surface on the internet. These videos were created by software using deepfake technology. The quality of the videos was low, and the contents were harmless, but experts in the field were alarmed.

First, they were concerned about the simplicity of creating those videos. The production of the movie videos of a deceased in the earlier film had a high cost and took a long production time.(39) In contrast, deepfake applications are simple and cost almost nothing to the end-user. In general, the end-user would only need a minute to create a deepfake video by using an App on his mobile phone.(40)

Secondly, this technology developed so rapidly, it took only a few years to create end-user Apps that were user-friendly. For example, in late 2019, a Chinese company developed an App with limited features that allow a user to deepfake their faces onto the face of several actors and actresses from selected movie scenes.(41) This application, ZAO, was followed by several Apps that also had some limitations. However, in 2020, a United States-based company developed an App that gave the end-user features to create high-quality deepfake videos.(42) This technology needed just three years to be developed for an end-user. It is safe to anticipate that this technology will be available to professional users with powerful processing machines in a very short time.(43)

(39) Eliana Dockterman, Fast & Furious 7 Filming Suspended, Time (2013). Available at https://entertainment.time.com/2013/12/03/fast-furious-7-filming-suspended-after-paul-walkers-death (last visited Jul 14, 2020), and Nolan Feeney, Fast & Furious 7: Paul Walker's Brothers Film Action Scenes, Time (2014). Available at https://time.com/64024/paul-walker-replacement-fast-furious-7 (last visited Jul 14, 2020).

(40) Mikael Thalen, You can now deepfake yourself into a celebrity with just a few clicks, The Daily Dot (2020). Available at https://www.dailydot.com/debug/impressions-deepfake-app (last visited Jul 14, 2020).

(41) Allan Xia, What You Need to Know About ZAO–the New, Viral Chinese deepfake App, The Daily Dot (2019). Available at https://www.dailydot.com/unclick/chinese-deepfake-app-zao/ (last visited Jul 14, 2020).

(42) Mikael, supra note 41.

(43) Disney scaled the resolution of deepfake to convert higher resolution to be 1080 1080 pixel. See Jacek Naruniec et al., High-Resolution Neural Face Swapping for Visual Effects, Disney Research Studios (2020). Available at http://studios.

Current applications of deepfake technology are limited to photographs and videos. However, voice and real-time deepfaking will be available soon to the end-user because of the speed of the development of deepfake technology.(44) In response, three large companies and one organization have united together to develop counter-technology that can detect deepfake and deceptive materials. Amazon, Facebook, Microsoft, and The Partnership on AI (PAI) created this project and named it Deepfake Detection Challenge (DFDC).(45)

The Potential Applications of Deepfake Technology

The current applications of deepfake technology are, at most, designed for end user's entertainment. However, this situation is expected to change. Soon, the only limitation of deepfake will be a user's imagination. With the proper tools and enough time, the applications of deepfake would exceed many expectations. Also, large entities such as governments and companies are probably become interested in this technology by now. Nonetheless, the interests in this technology may be motivated by good or bad faith.

The most obvious prospective application of deepfake is if deepfake used to alter the speech of political figures to pass or support a certain agenda. Many countries claim that they do not oppose freedom of speech, but none would permit public manipulation, falsification, and misrepresentation. The catastrophic scale of propaganda using deepfake could be tremendous if its use was adopted to carrying out the manipulation games that have become common in recent history.

= disneyresearch.com/2020/06/29/high-resolution-neural-face-swapping-for-visual-effects/ (last visited Jul 14, 2020), and James Vincent, Disney's deepfakes are getting closer to a big-screen debut, The Verge (2020). Available at https://www.theverge.com/2020/6/29/21306889/disney-deepfake-face-swapping-research-megapixel-resolution-film-tv (last visited Jul 14, 2020).

(44) See Andreas Rössler et al., FaceForensics++: Learning to Detect Manipulated Facial Images (2019). Available at http://www.niessnerlab.org/projects/roessler2019faceforensicspp.html (last visited Jul 14, 2020), see also Ye Jia, supra note 37.

(45) Facebook, Deepfake Detection Challenge Dataset (2020). Available at https://deepfakedetectionchallenge.ai/ (last visited Jul 14, 2020). See Eliza Strickland, Facebook AI Launches Its Deepfake Detection Challenge, IEEE Spectrum (2020). Available at: https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/facebook-ai-launches-its-deepfake-detection-challenge (last visited Jul 14, 2020).

Comparable application of deepfake, but on a smaller scale, is to use the technology as a method of defamation. Whether the target is unknown individual or public figure such as a celebrity, this technology could be used to fabricate speeches or other acts that could create confusion or discomfort, complicate the life, or jeopardize the safety of the target. An individual could be made to sound as though he or she had said something on a recorded video or audio that was harmful or controversial, but in reality, the video or the audio was fake and misleading.

Similarly, someone could superimpose a fake face or body on a pornographic video or photograph. There was an App called DeepNude that allows end-users to use a fully-clothed photograph of an individual and produce a naked picture.(46) The naked body is not the body of the victim, but the damage would be done since we are living in a time where the distribution of disinformation is a science.(47) As this technology becomes more powerful, its potential for damage grows. In a controversial study by Deeptrace,(48) the company claimed that technology had already used in more than 14,000 videos online, with 96% being nonconsensual pornography.(49)

On the other hand, many uses of deepfake can be both respectable and valuable. The technology could be used to construct the voice of an author reading his work and his writing with consideration of intellectual property law. It could be used, with proper permits, to produce videos of public figures recreating certain scenes. It also could be used to create new types of art, as Samsung's AI lab did by creating so-called "living portraits" from famous paintings.(50)

(46) See Sigal Samuel, A Guy Made a deepfake App to Turn Photos of Women into Nudes. It Didn't Go Well, Vox (2019). Available at https://www.vox.com/2019/6/27/18761639/ai-deepfake-deepnude-app-nude-women-porn (last visited Jul 14, 2020).

(47) Edmund L. Andrews, How Fake News Spreads Like a Real Virus, Stanford School of Engineering (2019). Available at https://engineering.stanford.edu/magazine/article/how-fake-news-spreads-real-virus (last visited Jul 14, 2020).

(48) A cybersecurity company, read more at https://deeptracelabs.com.

(49) Deep Trace Labs, Mapping the Deepfake Landscape, Regmedia, 7 (2019). Available at https://regmedia.co.uk/2019/10/08/deepfake_report.pdf (last visited July 6, 2020).

(50) Gregory Barber, Deepfakes Are Getting Better, But They're Still Easy to Spot, Wired (2019). Available at https://www.wired.com/story/deepfakes-getting-better-theyre-easy-spot (last visited Jul 14, 2020).

Another potential use of deepfake is in forensics. Deepfakes could be created to simulate a crime or crime scene to help investigators reconstruct a crime. Moreover, deepfake technology can be used with friendly and familiar faces or voices in the healthcare sector, on advertising platforms, in the entertainment business, or official guidance instructions and posts. For example, patients with dementia should react better with AI communicate with them in friendly or familiar faces and voices.

In addition, this technology can be used in the education fields to advance the psychological learning factors within the learning process of a second language. Some studies suggest that familiar faces in education could help with the learning process in general.(51) This application would not be conceivable without a proper legal structure around deepfake technology.

Lawmakers may not know the future, but they can predict potential problems and possibilities based on the facts on hand. Thus, being prepared for the potential of deepfake technology and its applications is important. All applications of deepfake technology need to be the subject of legislation precisely because this technology can be used in such a variety of situations. Some of these situations are beneficial and should remain legal, and other are harmful and should be illegal.

Chapter One
Current Legislations of Deepfake Technology

Shortly, there would be many legal systems and laws that aim to legislate and govern deepfake technology unless this technology is suddenly ceased to exist, which is unlikely. Currently, few of the world's legal systems had acted to legislate how deepfake can be used. This chapter explores those systems of laws that have been created as well as the principles and grounds. The goal of exploring these laws and legal systems is to understand the legal principles and the legal grounds behind these laws and legislations. There also will be consideration

(51) J. Jacob Kirksey & Michael A. Gottfried, Familiar Faces: Can Having Similar Classmates from Last Year Link to Better School Attendance This Year?, 119 The Elementary School Journal 223–243 (2018).

of the political and cultural motivations when studying the current legislation on deepfake.

The Chinese Law Criminalization of Publishing Deepfake

The Chinese government took an action against the use of deepfake by adding a new set of regulations that limits or even prevent the use of deepfake. The new rules were issued by government authorities and labeled as Administrative Provisions as an indication that more deliberate rules may be issued by the legislative body subsequently. The rules in force as of January 2020.(52)

The Legal Grounds for the Passage of the Chinses Law on Deepfake

These rules were issued and enforced by the power of the administrative branch of Chinses government. Cyberspace Administration of China (CAC) issued the "Regulations on the Management of Network Audio and Video Information Services" on November 18, 2019. This regulation was designed to legislate the use of audio and video services on the internet, Apps, and any other form of network platforms.(53) The language of this regulation gives the assumption that it targeted a large variety of audio and video services in any form of network or platform.(54) However, the main subject of this regulation addressed was deep learning and virtual reality.(55)

The law placed limitations on the applications powered by deep learning and virtual reality. These limitations affected the provider, the user, and the content. The law was justified based on jeopardizing national security, undermining stability and public order, and violating the legal rights and interests of others.(56) Any content that challenged

(52) LexisNexis, Administrative Provisions on Online Audio-visual Information Services Unveiled, Lexis China (2019). Available at https://china.lexiscn.com/latest_message.php?id=344149 (last visited Jul 14, 2020).

(53) Wǎngluò yīn shìpín xìnxī fúwù guǎnlǐ guīdìng (网络音视频信息服务管理规定) [Regulations on the Management of Network Audio and Video Information Services] (RMNAVIS) (2019), article 2.

(54) *Id.*

(55) *Id.* at article 11.

(56) *Id.* at article 9.

state power, endangered political security, or undermined public stability was criminalized under this law. Besides, spreading or initiating online rumors, distributing or creating pornography, defaming, infringing on the portrait, privacy, or intellectual property rights, or any other legal rights was also criminalized under this law.(57)

Noticeably, the prohibited actions mentioned by the law were merely examples, which meant that other actions could be criminalized as well. Thus, any use of deep learning and virtual reality or even altering and adding relevant functions that violated public policy, jeopardized national security, or infringed on the legal rights of others was prohibited and fell under this law's scope.

The Enforcement Methods of the Law

In terms of implementation and enforcement of this regulation, the Chinese government identified a three-factor equation to guarantee the enforcement. The most important and first factor was the provider. The provider in the provisions is the service provider of the network of the audio and video information. The second factor was the government represented by its agencies. The third factor was the network user.

The provider refers to any organization or individual who provides network service of audio and video to the general public. Thus, the service has to be a provider to the community in the form of audio and video information network. Otherwise, the provisions of this regulation are inapplicable.(58) For a provider to fulfill its duty as an enforcer of the regulations, it must comply with a set of orders and satisfy several requirements.

The provider must conduct his business in good faith and compliance with the law and socialist core values.(59) The regulation encouraged providers to do their best to improve operator's professionalism and the quality of service.(60) The regulation dictates that the provider should handle the users' information, their legal rights, and any security breach

(57) *Id.*
(58) *Id.* at article 2.
(59) *Id.* at article 4.
(60) *Id.* at article 5.

professionally.(61) Thus, no service provider can operate without proper qualifications as stated by the law.(62)

The regulation encourages the provider to establish a user's credit system.(63) The credit system is a common practice in China, and the Chinese government uses the Social Credit System.(64) In practice, the provider has to have an authentication system to verify the true identities of the users. The provider can choose any suitable authentication method such as an identification number, a mobile phone number, or a number issued by the provider or any other organization.(65)

For technologies based on deep learning and visual reality, which is the focus of this regulation, both the provider and the user have to report to the government when using these technologies to create, publish, and distribute deceptive audio or video. The provider is required to identify himself to the authorities and report any sort of use through his service, and the user required to do the same when he engaged in such practice.(66) To achieve this, the provider is required to have decent management over its network.(67) This management is detailed in this regulation.(68)

As part of the required management, the provider has to have a system that detects the illegal use of its networks such as producing or distributing fake audio or video.(69) In addition, a provider has to have the power to stop and remove the broadcasting and distribution of illegal audio and videos.(70) This regulation gives the provider the power to temporarily stop and remove materials when the user failed the identification requirement. The provider can resume the service for the user when he is identified.(71)

(61) *Id.* at articles 7 and 10.
(62) *Id.* at article 6.
(63) *Id.* at article 5.
(64) The outline of the Construction of the Social Credit System. Available at http://www.gov.cn/zhengce/content/2014-06/27/content_8913.htm
(65) RMNAVIS, supra note 54, at article 8.
(66) *Id.* at article 11.
(67) *Id.* at article 17.
(68) *Id.* at articles 12-16.
(69) *Id.* at article 12.
(70) *Id.* at articles 12 and 13
(71) *Id.* at article 12.

In addition to the required management, the provider has to build a system that accepts complaints and reports from users and non-users about the service.(72) Also, the provider must inform users of the service about the terms of use.(73) The regulation gave providers permission to warn, rectify, restrict functions, suspend updates, and ban a user if he violates the terms of use.(74)

Curiously, the government relies heavily on the provider regarding the implementation of this regulation. The law expects self-discipline and full cooperation when needed from the provider.(75) Thus, the provider should have full reports on the date and operation when needed.(76) On the other hand, the roles of government and user within the provisions of this regulation are limited in comparison to the provider.

The Duration of the Law, Sunset Provisions, and Callousness.

This regulation was issued on November 18, 2019, and entered into enforcement on January 1, 2020. There was no time limit on the law and no sunset provisions. However, the regulation was issued by agencies of the Chinese executive branch that could reform the existing law or write a new one anytime.

The regulatory rules are remarkably strict. First, identification of all network users that uses deep learning and virtual reality. Second, a provider must cooperate with the authorities and provides them with full reports and operations logs without question. Third, the regulation criminalizes any violation of it along with five other laws listed.(77) Fourth, the regulation criminalizes any use of technology based on deep learning or visual reality without pre approval and notification to the authorities.

State of California Law

The State of California was the first legal system to address deepfake technology in the world. On October 3, 2019, the governor of California

(72) *Id.* at article 15.
(73) *Id.* at article 14.
(74) *Id.*
(75) *Id.* at articles 2 and 17.
(76) *Id.* at article 17.
(77) *Id.* at article 18.

approved two bills aimed at controlling the use of deceptive audio or visual media.(78) The two bills were to be added to the California Civil Code.(79)Both bills addressed different facets of deepfake technology. Assembly Bill 602(80) legislated the use of deepfake in pornography,(81) and Assembly Bill 730(82) legislated the use of deepfake in politics.(83) The effective date of both bills is January 1, 2020, since neither bill "...call[ed] for elections, [..] providing for tax levies or appropriations for the usual current expenses of the state."(84) Otherwise, the effective date would be different.

The Legal Grounds for the Passage of the CA § 20010 and § 1708.86 of the Civil Law Code

The common legal ground used to pass sections 20010 and 1708.86 was to protect the public from deceptive audio or visual media. Nevertheless, the subject matter of section 20010 was deceptive audio or visual media that was created, distributed, published, or broadcast in election campaigns. Section 1708.86, on the other hand, legislated deceptive pornographic audio or visual media.

Section 20010 governed four actions regarding deceptive audio or visual media including producing, distributing, publishing, or broadcasting(85) any campaign material that contains superimposed photographs or deceptive audio or visual media.(86) Further, these actions had to be motivated by "actual malice."(87) This section is only

(78) Assembly Bill. No. 602 and 730.
(79) Cal. Civ. Code § 20010 and § 1708.86.
(80) Cal. Civ. Code SEC. 1 § 1708.86.
(81) Cal. Civ. Code SEC. 1 § 1708.86(b)(1), and (2).
(82) Cal. Civ. Code SEC. 3 § 20010.
(83) Cal. Civ. Code SEC. 3 § 20010(a), and SEC. 4 § 20010(b).
(84) Cal. Const., art. IV, § 10(a).
(85) Cal. Civ. Code SEC. 3 § 20010(a), and SEC. 4 § 20010(b).
(86) *Id.*
(87) *Id.* "Actual malice" is a United States Supreme Court term that was first used in New York Times Co. v. Sullivan (1964). The Supreme Court definition of "actual malice" is: "with knowledge that it was false or with reckless disregard of whether it was false or not." The New York Times Co. v. Sullivan, 376 U.S. 254 (1964). This bill defined this term as "the knowledge that the image of a person has been superimposed on a picture or photograph to create a false representation, or a reckless disregard of whether or not the image of a person has been superimposed on a picture or photograph to create a false representation." Cal. Civ. Code SEC. 3 § 20010(a)(2).

applicable when these actions have done by a person, firm, association, corporation, campaign committee, or organization.(88)

The prohibition imposed by section 20010 is only active within a certain time, within 60 days of any election.(89) Also, this section had a sunset clause that limited its application until January 1, 2023.(90) In addition to the time limit, the bill excludes individuals and institutions from the prohibitions. The section had two separate subsections: Subsection 3 was dedicated to deceptive photographs and subsection 4 concerning audio and visual media. However, for the sake of simplicity, the following exceptions were merged from both subsections.

The first exemption was for licensed personnel under the Federal Communications Act(91) as long it is being performed within the granted functions by the license.(92) Second, this section should not conflict with the obligations, the rights, or the immunity of interactive service providers operating under section 230 of the Communications Decency Act (CDA).(93) Third, the section did not apply to any publisher or newspaper, magazine, or other periodical's employee as long as these periodicals were published regularly and does not publish campaign advertising or communication as its primary purpose.(94)

This section does not apply to television or radio broadcasting stations, including cable or satellite television operators, programmers, or producers as long as the television or radio broadcast states that the media authenticity is questionable.(95) In case of the internet or electronic publications of any newspaper, magazine, or another periodical, they were exempted from the requirements of the section on the condition

(88) *Id.* SEC. 3 § 20010(a), and SEC. 4 § 20010(b).
(89) *Id.* SEC. 4 § 20010(a).
(90) *Id.* SEC. 1 § 35(b).
(91) The Federal Communications Act of 1934 (47 U.S.C. Sec. 151 et seq.).
(92) Cal. Civ. Code SEC. 3 § 20010(d)(1).
(93) *Id.* SEC. 4 § 20010(d)(1). Communications Decency Act (CDA) is federal law and section 230 is safe harbor that protects provider or user of an interactive computer service from the liability of another provider's content, 47 U.S.C. § 230, a Provision of the Communication Decency Act.
(94) Cal. Civ. Code SEC. 3 § 20010(d)(2).
(95) *Id.* SEC. 4 § 20010(d)(2).

that they plainly state that the deceptive material is not accurate and does not represent the truth.(96) Fourth, if the deceptive audio or visual materials are published in a satire or parody platform, it is exempted fromthissection.(97)

However, it is mentioned many times that this section is to protect the public prospective from deceiving information regarding the election when this information seems believable to a reasonable person. If the deceptive media, audio, video, and photographs, does not cause confusion or does not appear authentic to a reasonable person, then this section is not applicable. Because if the media shows clearly to a reasonable person that it is meant to be deceptive, then there is "no harmnofoul."(98)

Accordingly, the deceptive medium was not prohibited under this section if there were a proper warning and disclosure that the media was inauthentic.(99) If the deceptive medium is a picture containing campaign materials, then the picture must contain the following statement in the same font size as used in that campaign picture: "This picture is not an accurate representation of fact."(100) If the deceptive medium is a video, then the video must contain the following statement: "This video has been manipulated," in readable font size for the average viewer.(101) If the deceptive medium is an audio, then the following statement must be read every two minutes: "This audio has been manipulated."(102)

Section 1708.86 is written to protect an individual from having any part of his body used in pornography products without his consent. An individual is protected by this section even with the use of a disclaimer or in the absence of malice.(103) To have a legal action against an individual regarding this section requires two conditions. First, the individual who created sexually explicit material of the victim knew

(96) *Id.* SEC. 4 § 20010(d)(4).
(97) *Id.* SEC. 4 § 20010(d)(5).
(98) *Id.* SEC. 4 § 20010(e).
(99) *Id.* SEC. 3 § 20010(b), and SEC. 4 § 20010(b).
(100) *Id.* SEC. 3 § 20010(b).
(101) *Id.* SEC. 4 § 20010(b)(1) and (3)(A).
(102) *Id.* SEC. 4 § 20010(b)(1) and (3)(B).
(103) *Id.* SEC. 1 § 1708.86(d) and (e)(1)(B)(ii)(II).

or should have known that the victim did not consent to the creation of such material.(104) Second, if a person intentionally publishes, makes available, or distributes sexually explicit material to the public without the consent from the victim, it is actionable.(105)

This section also gives immunity from liability under certain circumstances.(106) These circumstances are related to the investigations or reporting of unlawful activities under this section, or any other laws. That including a political or newsworthy work. However, if the victim is a public figure, and the value of the news is because of his position in the public, then the news is not newsworthy; and thus, is not protected under this section.(107)

The Enforcement Methods of the Law

Enforcement of section 20010 is undertaken in two steps. The first step is the right to submit a request addressing the court to issue a preliminary injunction(108) to prevent the publication, distribution, or broadcast of deceptive media.(109) A candidate for public office who was the subject of deceptive media can bring a civil suit against the responsible party by this section.(110)

A similar approach to enforce section 1708.86 is used. The enforcement of this section is sought through civil litigation. Although there is no preliminary injunction outlined in its provisions, this law does not prevent the victim from seeking an injunction based on section

(104) *Id.* SEC. 1 § 1708.86(b)(1).
(105) *Id.* SEC. 1 § 1708.86(b)(2).
(106) *Id.* SEC. 1 § 1708.86(c).
(107) *Id.* SEC. 1 § 1708.86(c)(2). However, this provision could be challenged because it could override some constitutional rights such as the freedom of speech and freedom of the press.
(108) This right is guaranteed by California Section 527 of the Code of Civil Procedure, which states: "preliminary injunction may be granted at any time before judgment upon a verified complaint, or upon affidavits if the complaint in the one case, or the affidavits in the other, show satisfactorily that sufficient grounds exist therefor."
(109) This right is for any registered voter in cases of a superimposed a picture or photograph; however, only a candidate for public office whose image, voice, or video of himself is deceptive can use this right. Cal. Civ. Code SEC. 3 § 20010(c)(1) and SEC. 4 § 20010(c)(1).
(110) *Id.* SEC. 3 § 20010(c)(2) and SEC. 3 § 20010(c)(2)

527 of the Code of Civil Procedure. The victim [plaintiff] can request the profit gained from the deceptive sexual material, economic and non-economic damages, punitive damages, lawyers' fees and costs, and any other available relief such as injunctive relief.(111) The victim can also request statutory damages instead of economic and non-economic damages.(112)

The Duration of the Law, Sunset Provisions, and Callousness.

Section 20010 is enforceable from January 1, 2020, to January 1, 2023.(113) Thus, on January 1, 2023, this law would cease to exist unless it is reintroduced on a new date with new law or the sunset date is canceled. However, the sunset date does not apply to civil cases that arise from violating the provision of section 20010 before the sunset day. In comparison, section 1708.86 does not have a sunset date, so the law will be in the books until it is revoked or reformed.

The legislators who wrote section 20010 tried to protect the constitutional and legal rights of the people without jeopardizing any other rights. The language and principles considered in the section were proportionate with the laws and regulations in the United States and the State of California. However, section 20010 does have some inconsistencies, which is apparent when comparing subsection 3 to subsection 4.

Subsection 3, for example, was reserved for deceptive photographs of a candidate for public office. On the other hand, subsection 4 was for a deceptive image, video, or audio.(114) Subsection 4 could replace subsection 3 if the specific rules in subsection 3 merged with the rules of subsection 4. Besides, the right to request a preliminary injunction is guaranteed to any registered voter under subsection 3, but it is limited to the candidate for the public office in subsection 4, though there is little difference between the violations cited.

Legislators drafted section 1708.86 with more consistent provisions and interconnected thoughts. This could be because this section was

(111) *Id.* SEC. 1 § 1708.86(e).
(112) *Id.* SEC. 1 § 1708.86(e)(1)(B)(ii).
(113) *Id.* SEC. 1 § 35(b).
(114) *Id.* SEC. 4 § 20010(b)(2).

written without a sunset provision or because it was the first law of its kind; thus, the lawmaker did not have to deal with numerous regulations and provisions to ensure the enforcement. Either way, this section will be remembered as the first law regulates the use of deepfake in pornography.

Future Laws on Deepfake

Currently, there is no information regarding laws, law projects, or drafts being written in any other jurisdiction than in China and California. However, the European Union understands the danger of deepfake technology. Through its executive branch, the European Commission, the European Union addressed the online disinformation practice on April 26, 2018.(115) The European Commission did not issue any laws or regulations, but it did issue advice on how existing laws and regulations could be used to protect the public from misinformation.

Chapter Two
Legal Grounds for Legislating Deepfake Technology

Deepfake technology can be legislated in many ways and on a variety of legal grounds. However, such legislation would be complicated to develop because politics, social peace, and human needs all must be considered. Besides, excessive legislating could result in a bureaucratic nightmare.(116) Thus, when discussing the legal grounds to legislate deepfake technology, lawmakers need to take into consideration those most vulnerable, along with community morals, ethics, rights, and groups of people or individuals.

(115) COM/2018/236 final, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Tackling online disinformation: a European Approach.

(116) See Abdulaziz Alobud, Developing Saudi Private International Law to Accommodate E-Commerce 196 (2019). Available at https://www.researchgate.net/publication/336230121_DEVELOPING_SAUDI_PRIVATE_INTERNATIONAL_LAW_TO_ACCOMMODATE_E-COMMERCE_TRANSACTIONS_GROWTH (last visited July 10, 2020).

Human Rights Violations

The Universal Declaration of Human Rights[117] provided a range of basic human rights. Some of these rights will be violated by some uses of deepfake technology. The following rights are not necessarily explicitly mentioned in the UDHR, but they are very likely to be violated by the unlawful use of deepfake.

Privacy

The right to privacy is a basic human right that has been acknowledged and is protected by the UDHR[118] and many legal systems. The only exception to intervene with the privacy of an individual is when it is supported by law. Using deepfake technology to create deceptive media that harms an individual is an invasion of privacy. Privacy does not only mean protection from observation; it also means protection from being disturbed.[119] Moreover, the right to privacy is related to the right of publicity since publicity is the opposite of privacy. An individual has the right to revoke some aspect of his privacy and go public, but that must be by his choice and by his free will.

Protecting the right of privacy can be used alone or with a collection of other rights to undergird legislation on deepfake. Deepfake does not need to be banned completed, but certain uses of deepfake that jeopardize the right of privacy should be legislated. The existing protections of the right of privacy in a particular legal system needs to be examined to determine what additional protections are necessary for the rise of deepfake.

Human Dignity

Under human rights law, human dignity is one of the important human rights, though the definition of "human dignity" is subject to interpretation. In its simplest terms, two components shape the meaning

(117) The Universal Declaration of Human Rights (UDHR). Available at https://www.un.org/en/universal-declaration-human-rights/ (last visited July 10, 2020).

(118) *Id.* Article 12

(119) Bryan A. Garner & Henry Campbell Black, Black's law dictionary, 1349 (8 ed. 2004).

of human dignity.(120) The first component is the rights, and the second is that which is forbidden. An individual is entitled to groups of rights such as the right to respect and value. In addition, it is prohibited to unlawfully humiliate an individual, damage his reputation, or destroy his credibility.

Consequently, legislators can criminalize any use of deepfake technology that harms and damages the dignity of an individual. First, human dignity must be defined, if that is not already done. The law should consider the dignity of all affected individuals, even when they are not the subject of the deceptive media. The family of an individual who was the subject of a deceptive medium is affected by an attack on the dignity of a father, son, sister, etc. The law usually gives the judge the power of judicial discretion. Therefore, the definition of dignity and the amount of damage to dignity is determined on a case-by-case basis.

Protected Identity

The identity of a human is protected by criminal law and human rights law. Individual identity is part of the social existence of a human being. That identity should be protected from abuse and theft. If deepfake technology is used to deceive a reasonable person who, then, associates the deceptive media with an individual, which is a violation of human rights. However, if the deceptive medium is used to create false identification, then that is identity theft and fraud in addition to a human rights violation. Legislators, therefore, can use identity theft and misidentification as legal grounds to legislate against harmful deepfake applications targeting the identity of individuals.

Right of Publicity

As the right to privacy is a human right, the right of publicity is also a human right. An individual has the right to publicize any aspect of his life. However, he is the only one who has this right; thus, any publicity without his consent is unlawful even if the publicity is beneficial. Legal systems address the right of publicity in several different ways. Some

(120) See Doron Shultziner, Human Dignity - Functions and Meanings, 3 Global Jurist Topics 4 (2004).

limit the right of publicity to commercial use, and other legal systems do not recognize the right of publicity at all. Instead, they protect the right of publicity under the right of privacy.(121) Other legal systems do not recognize the right of publicity of a child when there is consent from his guardians.(122) The legislators should consider the legal system's position on the right of publicity concept before criminalizing deepfake applications that publicize an individual.

Defamation

Defamation is a false statement that harms the reputation of an individual.(123) This unlawful act is another legal ground to legislate deepfake technology. Defamation allegations are usually examined by a two-part test. First, is the media authentic and reflects the truth? Second, does the deceptive media harm the reputation of the individual? If the media is deceptive and was created by deepfake, then the media is not real and does not reflect the truth. However, the answer to the second part of the test predetermines whether to consider the creation of deceptive media as defamation or copyright infringement.

Copyright Infringements

Copyright is a power of ownership over tangible published and unpublished work. Copyrighted material covers a wide swath of products, from a personal photo or a voice. When creating deceptive media using deepfake, it is likely that copyrighted material is used and then modified. The owner of the original material subject to modification is entitled to seek remedies under copyright law. Thus, a person who modifies copyrighted material with no valid license from the owner of the material is accountable if the owner chooses to press charges. This means a government cannot enforce the copyright law without copyright infringement and a request from the owner of the copyrighted material.

The legislators should consider the status of copyright law within the given legal system. The copyright law must contain the following

(121) *See* Restatement (Second) Of Torts §§ 652A - 652I.
(122) For example, see CAL. CIV. CODE § 3344, and tit. 42. Pa. Cons .Stat § 8316.
(123) Black's law dictionary, supra note 120, at 479.

elements. It must be thoroughly prepared and carefully written. Secondly, the copyright law must have the support of the executive branch to detect infringements and fairly execute the law. If the copyright law is impractical or inefficient, and it is used as a legal ground to legislate deepfake, the result would not be effective.

Violation of Public Policy

"Public policy" is a reference to a variety of concepts; however, in this paper, it refers to the legal doctrine of public policy. The legal doctrine of public policy is the collection of legal principles derived from the constitution, laws, regulations, directives, and enforcement policies in a national legal system.(124) Public policy varies greatly from legal system to legal system based on the ideologies and limitation placed on it in each country. Nevertheless, it is an effective ground for legislating deepfake technology.

If public policy is to be used as a ground for legislating deepfake, it must be interpreted in line with the given legal system. The Saudi legal system, for example, has deemed Islamic Law, Sharia, as the primary source of the principles that form the Saudi legal system.(125) Thus, public policy in the Saudi legal system is composed of Islamic law, national laws, statutes, regulations, and directives. The role of Islamic law in the Saudi legal system does not mean Islamic law is the same as Saudi law. Instead, Saudi law is a result of an application of Islamic legal principles.(126) This is an important example of the use of public policy in creating laws.

Violation of Islamic Law

For a legal system that uses Islamic Law as a primary source of legal principles, it provides a good legal ground to legislate some uses of deepfake technology while not fully following the strictures of Islamic law. For example, the use of deepfake to create deceptive pornography media is prohibited by Islamic law because pornography is prohibited,

(124) See Percy H. Winfield, Public Policy in the English Common Law, 92, 42 Harvard Law Review 76–102 (1928).
(125) The Basic Law of Governance, Royal Order No. A/90, 3/1/1992, article 1.
(126) See Alobud, supra note 117, at 193.

therefore, recreating it is also illegal. Islamic law can be used as a legal ground to prevent copyright infringement using deepfake. However, the most efficient practice is to use Islamic law as a ground to enact laws protecting copyright, then use these laws as grounds for legal action.

Freedom of Expression and Speech

Legislating deepfake technology and its applications does not mean preventing its use. Legislating deepfake or any subject matter is about protecting rights and enforcing obligations. Thus, the freedom of speech, freedom of the press, the freedom of expression or freedom of religion can be used as a legal ground to legislate restrictions on deepfake technology. For example, if the legislators used the freedom of the press as a legal ground, then a reporter who becomes the subject of a deceptive medium that seeks to deprive him of this right has the right to seek redress. Also, legislators must consider that freedom of the press means allowing some uses of deepfake technology that does not violate public policy such as creating deepfake videos of himself reporting the news.

All these legal grounds are suggestions that may be more or less suitable to legislating against deepfake technology and its applications. Additionally, some legal systems provide better legal grounds that work specifically with their political and social system. Moreover, not all legal grounds share same legal terminologies even when sharing same legal values. There are differences for one legal system to another in term of the legal terminologies and legislators should consider this fact when adopting laws legislating deepfake.

Conclusion

Findings

1 - Below is a list of the findings of this paper, either as they are presented on the previous pages observed during research.

2 - The applications of deepfake technology will exceed expectations in a short period.

3 - In many legal education systems, legal education regarding Information Technology Law, Computer Law, Internet Law, Cybersecurity Law, and Cyberlaw is lacking.

4 - Few legal systems address deepfake technology.

5 - There are different approaches to legislating deepfake. Chinese law has the most direct and confrontational approach by criminalizing the use of technology. Chinese law was so extensive that it criminalized any use of technologies based on deep learning and virtual reality without a state prior permission.

6 - The State of California is adapted to the constitutional and legal rights of the people of the State of California and the United States.

7 - All existing laws and regulations that address deepfake technology do not address the technology by its known name, deepfake. Instead, legislators took the opportunity to legislate all technology falling under the criteria of deceptive media.

8 - Chinese law targeted the technology behind deepfake instead of addressing a deepfake directly and add visual reality to it; on the other hand, California law targeted any altered and deceptive media.

9 - Using only one legal ground to legislate the uses of deepfake technology is not be the best option. The range of such applications is large and is unlikely to be covered by only one legal ground. The copyright law, for example, is inadequate in the legislation of deepfake because the material used in deceptive media is not protected under copyright law, or because an owner chooses to not seek redress for copyright infringement.

10 - There are legal grounds that can ease and or tighten the control over deepfake technology. For example, privacy as a legal ground can be used to prevent any form of invasion of individual privacy, with the result that the rules are strict. It also could be used to create laxer rules by only criminalizing a few forms of privacy invasion.

11 - Some considerations of legal grounds would help legislators encourage and control the uses of deepfake technology. Freedom

of expression, for example, can be a legal ground to motivate the use of deepfake technology to support a person's ability to express beliefs, opinions, philosophies, ideologies, and emotions. At the same time, the same legal ground can be used to prevent any use of deepfake technology to suppress the freedom of expression.

Recommendations

1 - In connection with findings number 1-,2-, and 3-, I recommend developing an interest in teaching modern law fields that incorporate technology. That would be beneficial for the legal system since computer-based technology will not wait for legislation.

2 - In connection with findings number 4-, 5-, and 7-, I recommend adopting a balanced approach when legislating deepfake. First, criminalizing all violations in the new regulation is not necessary. Instead, the violations of the regulation should be ranked as a major felony, misdemeanor, infraction, or any other violation ranks used by the legal system. Second, the new regulation must consider the rights established by the law to protect them and not reduce them. Third, legislators have the burden to find a balanced approach to address deepfake technology. I recommend using broad terminologies that cover two mains deepfake properties. First, deep learning technology is the technology that generates deepfake, and second, deceptive media, which deceives reasonable people to accept an item's authenticity.

3 - In connection with findings number 6-, legislators must be explicit with the definition of terms used in the law. Technology carries its new terminology that may be unfamiliar or familiar, though used differently. Some legislative systems address the definition of terms in the first few articles of law. Other systems may add the definition when the term is first used. Interestingly enough, the State of California legal system wavered between these two styles, as can be seen in section 20010 and section 1708.86. My recommendation is to place the definitions at the beginning and be consistent. Both approaches are acceptable as long as the used terms are specified and defined. However, the use of terms must adhere to the characteristics of public opinion.

4 - In connection with findings number 8-, 9-, and 10-, I recommend that any attempt to legislate deepfake should take into consideration a collection of legal grounds. Adopt the best combination of legal grounds in any given legal system and use the efficiency of its laws and regulations. Finally, I recommend revising all existing laws and regulations that are related to or affected by any uses of deepfake technology.

Bibliography

1 - Books, Articles, Textbooks and academic papers.

- Aäron van den Oord & Sander Dieleman, *WaveNet: A Generative Model for Raw Audio*, DeepMind (2016).
- Abdulaziz Alobud, Developing Saudi Private International Law to Accommodate E-Commerce 196 (2019).
- Alan M. Turing, *Computing Machinery and Intelligence*, LIX Mind 433–60 (1950).
- Allan Xia, *What You Need to Know About ZAO–the New, Viral Chinese deepfake App*, The Daily Dot (2019).
- Andreas Rössler et al., FaceForensics++: Learning to Detect Manipulated Facial Images (2019).
- Andrew Griffin, *Facebook robots shut down after they talk to each other in language only they understand*, The Independent (2018).
- Ashok Singh, Science & Technology (2007).
- Benjamin Powers, *Adobe is Developing Photoshop for Your Voice*, Medium (2018).
- Bryan A. Garner & Henry Campbell Black, Black's law dictionary (8 ed. 2004).
- Christof Koch, *How the Computer Beat the Go Master*, Scientific American (2016).

- Copeland, B. Jack, *The Modern History of Computing*, The Stanford Encyclopedia of Philosophy (Winter 2017 Edition), Edward N. Zalta (ed.).
- Deep Trace Labs, *Mapping the Deepfake Landscape*, Regmedia (2019).
- Doron Shultziner, *Human Dignity - Functions and Meanings*, Global Jurist Topics 4 (2004).
- Douwe Draaisma & Paul Vincent, Metaphors of memory: a history of ideas about the mind (2001).
- Eben Upton et al., Learning Computer Architecture with Raspberry Pi (2016).
- Edmund L. Andrews, *How Fake News Spreads Like a Real Virus*, Stanford School of Engineering (2019).
- Eliana Dockterman, *Fast & Furious 7 Filming Suspended*, Time (2013).
- Eliza Strickland, *Facebook AI Launches Its Deepfake Detection Challenge*, IEEE Spectrum (2020).
- Facebook, *Deepfake Detection Challenge Dataset* (2020).
- Frank da Cruz, Columbia University Computing History A Chronology of Computing at Columbia University, Herman Hollerith (2011).
- Gregory Barber, *Deepfakes Are Getting Better, But They're Still Easy to Spot*, Wired (2019).
- Hubert Davis, *How Deepfake Technology Actually Works*, ScreenRant (2020).
- J. Jacob Kirksey & Michael A. Gottfried, *Familiar Faces: Can Having Similar Classmates from Last Year Link to Better School Attendance This Year*?, 119 The Elementary School Journal 223–243 (2018).
- Jacek Naruniec et al., High-Resolution Neural Face Swapping for Visual Effects, Disney Research Studios (2020).

- Jake Frankenfield, *How Artificial Intelligence Works*, Investopedia (2020).
- Jake Frankenfield, *Machine Learning*, Investopedia (2018).
- James R. Slaby, Robotic Automation Emerges as a Threat to Traditional Low-Cost Outsourcing, HFS Research (2012).
- James Vincent, *Disney's deepfakes are getting closer to a big-screen debut*, The Verge (2020).
- Luca Sgambi, Artificial Intelligence: Historical development and applications in civil engineering field (2008).
- Maad M & Mijwil, History of Artificial Intelligence (2015).
- Mark Hall, Douglas Engelbart, Encyclopædia Britannica (2020).
- Marshall Hargrave, *How Deep Learning Can Help Prevent Financial Fraud*, Investopedia (2019).
- Mikael Thalen, *You can now deepfake yourself into a celebrity with just a few clicks*, The Daily Dot (2020).
- Nolan Feeney, *Fast & Furious 7: Paul Walker's Brothers Film Action Scenes*, Time (2014).
- Percy H. Winfield, *Public Policy in the English Common Law*, 42 Harvard Law Review 76–102 (1928).
- Phuoc Nguyen, Deepfakes and Human Rights Litigation in New Zealand.
- Restatement (Second) Of Torts § 46 (1965).
- Robert A. Burgelman, Strategy is Destiny: How Strategy-Making Shapes a Company's Future (2002).
- Robert T. Futrell, Donald F. Shafer & Linda I. Shafer, 1 Quality software project management (1 ed. 2002).
- Sally Adee, *What Are Deepfakes and How Are They Created?* IEEE Spectrum (2020).

- Sigal Samuel, *A Guy Made a deepfake App to Turn Photos of Women into Nudes. It Didn't Go Well*, Vox (2019).
- Somanathan B. Nair, Digital Electronics and Logic Design (2004).
- Troy Segal, *The Deal with Big Data*, Investopedia (2019).
- Ye Jia et al., Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis (2019).
- Yuval Nirkin, Yosi Keller & Tal Hassner, FSGAN: Subject Agnostic Face Swapping and Reenactment (2019).

2. Laws, Regulations and Law Cases.

- Cal. Civ. Code § 1708.86.
- Cal. Civ. Code § 20010.
- Cal. Const., art. IV, § 10(a).
- COM/2018/236 final, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Tackling online disinformation: a European Approach.
- Communications Decency Act (CDA), 47 U.S.C. § 230.
- Constitution of Pennsylvania State, Pa. Cons .Stat.
- LexisNexis, Administrative Provisions on Online Audio-visual Information Services Unveiled, Lexis China (2019).
- The Basic Law of Governance, Royal Order No. A/90, 3/1/1992.
- The Federal Communications Act of 1934 (47 U.S.C. Sec. 151 et seq.).

- The New York Times Co. v. Sullivan, 376 U.S. 254 (1964).
- The Outline of the Construction of the Social Credit System.
- The Universal Declaration of Human Rights (UDHR).
- Wǎngluò yīn shìpín xìnxī fúwù guǎnlǐ guīdìng (网络音视频信息服务管理规定)[Regulations on the Management of Network Audio and Video Information Services] (RMNAVIS) (2019).

تقنين تقنية التزييف العميق

د. عبدالعزيز بن صالح بن عبدالله العبود*

ملخص الدراسة

الأهداف: تقنية التزييف العميق بتطبيقاتها المتعددة والإمكانيات الحالية والمتوقع تحققها في المستقبل القريب يحتم على الأنظمة العدلية في العالم التصدي مبكراً بسن تنظيمات وتشريعات تعالج استخدامها وتحد من آثارها. وتلك الأنظمة العدلية أمام ثلاث خيارات في التعامل مع هذه التقنية. أولى هذه الخيارات هو عدم استباق الأحداث وعدم فعل شيء. وهذا الخيار يعني التعامل مع آثار التقنية ومحاولة معالجتها حين أو بعد وقوعها. الخيار الثاني هو المنع التام للتقنية وإغلاق الباب أمامها كما فعلت الصين في التصدي لها. هذا وإن كان الحل الأسهل، إلا أنه حل مؤقت يخلق العديد من الإشكالات المستقبلية وقد يحرمنا من تقنية فيها نفع إذا استخدمت بشكل حسن. الخيار الثالث هو دراسة هذه التقنية وتقييم أخطارها واستخداماتها ومن ثم وضع تنظيمات وتشريعات تساعد على الحد من الآثار الضارة وتسمح بالاستخدامات المفيدة والعادلة. ولأجل الخيار الثالث، يحتاج المنظم إلى أسس قانونية يبني عليها التنظيم والتشريع الخاص بتقنية التزييف العميق. **المنهجية:** هذا البحث يعالج السؤال المتعلق بالأسس القانونية المناسبة لبناء تنظيم أو تقنين يعالج تقنية التزييف العميق من خلال منهج تحليلي للتشريعات الحالية. يجيب البحث عن هذا التساؤل بأن يعرض نتائج الخيارات الثلاث السابق ذكرها ثم يجيب عن أفضل أسس قانوني يمكن أن يبنى عليه التنظيم أو التشريع الجديد. **النتائج:** والنتيجة تشير إلى أن أساس قانوني واحد لا يكفي لبناء تنظيم أو تشريع يعالج تقنية متعددة الاستخدامات كتقنية التزييف العميق. كما يخلص البحث إلى أن لكل نظام عدلي خصوصيته وبالتالي يجب على كل نظام عدلي أن يراعي تلك الخصوصية عند اختياره الأسس القانونية المناسبة لمعالجة وضع هذه التقنية. **الخاتمة:** حرية التعبير قد يصلح أن تكون أساساً قانونياً يراعى عن استحداث التنظيم أو التشريع الجديد في نظام عدلي معين، ولكن ذات الأساس القانوني قد لا يصلح في نظام عدلي آخر.

الكلمات المفتاحية: التزييف العميق، تشريع، الذكاء الصناعي، التعلم العميق.

* أستاذ القانون المشارك في قسم الأنظمة في كلية الأنظمة والاقتصاد في الجامعة الإسلامية في المدينة المنورة.

الإيميل: Abdulaziz.Alobud@iu.edu.sa

- تُسلّم البحث في: ٢٠٢١/١/٢٨، أُجيز للنشر في: ٢٠٢١/٤/١١.

للاستشهاد بهذا البحث انظر ص٥٠٧

دكتور/ عبدالعزيز بن صالح بن عبدالله العبود، أستاذ القانون المشارك بقسم الأنظمة في كلية الأنظمة والاقتصاد في الجامعة الإسلامية بالمدينة المنورة. حاصل على الدكتوراه في القانون من جامعة سافوك في الولايات المتحدة الأمريكية. عمل رئيساً لقسم الأنظمة وعميداً لكلية الأنظمة والدراسات القضائية بين الأعوام ١٤٤٢ إلى ١٤٤٥هـ.
البريد الإلكتروني: Abdulaziz.Alobud@iu.edu.sa

الاستشهادات:

العبود، عبد العزيز. (٢٠٢٤). تقنين تقنية التزييف العميق. *مجلة الحقوق*، جامعة الكويت، *٤٨*(٣)، ٤٦٩-٥٠٧.

To Cite:

Alobud, Abdulaziz. (2024). Taming the Faceless Beast: Legislating Deepfake Technology. *Journal of Law*, Kuwait University, *48*(3),469-507.