Daniel Andrés Birchner 1/18/2025

POST-TRUTH AND POLITICS: THE GROWING IMPACT OF DEEPFAKES

Although the term "deepfake" does not currently have a widely accepted definition, in order to arrive at a transversal understanding, we can approach it through Nina Schick's (2020) description, which defines it as: “a type of digital media (audio, video, and images) entirely or partially generated by artificial intelligence with a malicious or disinformative purpose” [1]. Thus, a deepfake consists of the creation or modification (with malicious intent) of a person’s physical appearance or even their voice, and placing it digitally in places where they were never present or making them say things they never said, simply by making their image move as instructed, with AI mediating this entire process.

As mentioned earlier, it is not currently possible to reach a consensus on the definition of deepfake in the relevant scientific literature. However, it is possible to find certain common ideas across almost all definitions, which are detailed below:

- The term "deepfake" is a combination of the terms deep learning and fake news [2] [3].
- A deepfake consists of the alteration of videos, audio, or images through AI (via machine learning processes, and more specifically, deep learning) [4] [5].
- A deepfake is an intentional montage, meaning that the altered digital medium has a clear human intention, assisted by AI [6] [7].
- A deepfake can have a positive purpose (technology demonstration, practical jokes) or a negative one (identity theft, public defamation) [8] [9]. The latter negative perception of intent is predominant.

---

[1] Schick, N. (2020) – *Deepfakes: the coming infoapocalypse*. Editorial Tamang Ventures, Nueva York, EUA. p. 9.
[2] *Ídem.*
[3] Botha, J., Pieterse, H. (2020) – *Fake news and deepfakes: a dangerous threat for 21st Century Information Security.* Council for Scientific and Industrial Research (CSIR), Southafrican Government.
[4] Le Cunn Y. *et al.,* (2015) - *Deep Learning* en Revista Nature, Vol. 521.
[5] Twang, T. (2020) – *Deepfakes, a grounded Threat Assessment*. Centro para la Seguridad y la Tecnología Emergente (CSET).
[6] Schick, N. *op. cit.*, p. 10.
[7] Twang, T. *op. cit.*
[8] Mahmud, B. U. y Sharmin, A. (2020) – *Deep insights of Deepfake Technology* en Revista DUJASE, Vol. 5.

It is important, however, to contain the dimensions of the concept. As previously stated, a deepfake is an intentional montage, assisted by AI, with possible disinformative or malicious purposes on the part of the creator (excluding "positive" purposes, such as mere entertainment or technology demonstration) [10]. This conceptualization excludes montages that do not require AI or those made in films to recreate the faces of actors who have already died (such as Princess Leia in the movie "Star Wars: Rogue One"). It also excludes montages made using video editing programs that do not involve AI. Additionally, thanks to the mentioned total/partial assistance of AI, deepfakes do not require a large volume of human intervention [11], as AI handles the most complex tasks automatically, allowing the human creator to focus on the intent behind the deepfake.

There is also a sub-category classification regarding the level of sophistication of the deepfake. The true deepfake is a deep manipulation of video and audio that uses advanced AI tools to replace faces or speeches with high realism. This technique relies on deep learning algorithms, making it difficult to detect. The cheapfake, on the other hand, is achieved with basic edits such as altering playback speed or subtitles, without the need for AI intervention. Due to its simplicity, it is usually easier to identify than the deepfake. In terms of purpose, according to statistics from the Dutch technology company Deeptrace [12], 96% of current deepfakes (video) are used for pornography, with much smaller proportions used for comedic (prank) purposes or political manipulation. Additionally, although still marginal, the U.S. National Artificial Intelligence Program, in its 2021 report, already warned of the potentially dangerous (intentional) use of deepfakes for political purposes[13].

In 2020, the Center for Security and Emerging Technologies, a think tank at Georgetown University in the USA, released its report "Deepfakes: A Grounded Threat Assessment." In this report, the entity highlights the growing concern within the U.S. national security community regarding the potentially dangerous uses of deepfakes, particularly those related to manipulating videos of political figures and their consequences, both national and international. It emphasizes that the proliferation of

---

[9] Graber–Mirchell, N. (2021) - *Artificial Illusions: Deepfakes as Speech* en Revista Intersect, Vol. 14, N° 3.
[10] Botha, J., Pieterse, H. (2020) – *Fake news and deepfakes: a dangerous threat for 21st Century Information Security.* Council for Scientific and Industrial Research (CSIR), Southafrican Government.
[11] Westerlund, M. (2019) – *The Emergence of Deepfake Technology: A Review* en Technology Innovation Managment Review, Vol. 9.
[12] Cafranc, P. (2019) - *Deepfake, cuando lo que vemos ya no es fiar*, pág 6.
[13] National Program for Artificial Inteligence (2021) - *Deepfake Guide.*

deepfakes corresponds to an era where “we can no longer believe what we see” (we can no longer believe what we see), a time when much of the information is disseminated via digital social platforms that are easily manipulable on one hand and designed to polarize/compartmentalize public opinion on the other.

It is necessary to recognize that deepfakes spread and are considered in a defined context and universe, that of fake news. According to the International Federation of Journalists[14], the term fake news is used to conceptualize the dissemination of false news that triggers a dangerous cycle of intentional disinformation[15]. They can take various formats, such as written press, audio, or videos, online or in the form of rumors.

To ensure conceptual accuracy, there is a tendency to use some terms interchangeably with fake news: false news and propaganda. Fake news are not necessarily false news, as the latter are not initially conceived with the purpose of disinformation but can be false due to omissions or misinterpretations. False news can be reviewed and corrected accordingly, while fake news, from its conception, is intended explicitly to misinform, and they are generally unverifiable, with no possibility of confirming their truthfulness. False news can appear in reputable media outlets, while fake news are typically spread through questionable or biased sources.

When deepfakes have a political intent, they can be considered within the realm of fake news, i.e., intentionally false news. It is important to highlight that false or contentious news are not new, but the innovation of fake news lies in their primary medium of dissemination (social media) and in the social-informational context known as post-truth.

The concept of post-truth refers[16], according to the Oxford Dictionary (1992), to the symbolic framework where objective facts are less important in shaping public opinion, prioritizing appeals to emotion or popular beliefs[17]. In other words, a perspective shaped not by rational and objective criteria for analyzing reality, but by interpretations based purely on emotional and volatile viewpoints, which are easily manipulated by news or information that appeal to sensationalism rather than critical thinking. This

---

[14] International Federation of Journalists - *¿Qué son las fakenews?*
[15] Ídem.
[16] *Posverdad* is the term in spanish.
[17] Available in: https://www.oxfordlearnersdictionaries.com/spellcheck/english/?q=postruth

definition is also confirmed in the Royal Spanish Academy's version, which defines post-truth as "*deliberate distortion of reality, manipulating beliefs and emotions to influence public opinion and social attitudes*." [18] This definition will be important in understanding the effect of deepfakes on public opinion, alongside the cognitive processing errors studied in Cognitive Psychology. It is a key concept in understanding why deepfakes pose a threat to National and International Security, paralleling the technological context that supports it.

POLITICAL EFFECTS (CASES)

The political use of deepfakes is less frequent compared to their applications in pornography or parody, but when used for political manipulation, they can have profound consequences on the continuity of a government, electoral processes, or even the course of a military conflict. In these cases, the key lies in manipulating public opinion through falsified images or videos to cast doubt on the legitimacy of a government, influence support for state policies, induce votes for or against a candidate, defame political figures, or even direct economic perceptions for profit.

The following examples showcase cases where deepfakes either produced or were on the verge of producing relevant political changes. Governments are becoming increasingly aware of the power of these techniques, employing experts who aim to detect and debunk them before they gain significant traction. Both deepfakes and cheapfakes can effectively manipulate public opinion, sometimes without requiring advanced technical resources. However, when targeting a more informed or smaller group, greater realism is required to achieve the desired effect. The cases are presented in order of political impact, from least to greatest.

In 2018, the United States witnessed a surprising turn in the perception of deepfake power when actor and filmmaker Jordan Peele released a fake video of former President Barack Obama. Titled "You won’t Believe what Obama says in this video!", the footage showed the former president making offensive statements about Donald Trump, with unsettling realism. Given Obama's prestige and Peele’s popularity, the video garnered massive media attention and served as an early warning about how high-precision

[18] Available in: https://dle.rae.es/posverdad?m=form

digital manipulation could be used to spread false information. Peele, in an act of transparency, quickly revealed his authorship, explaining that his goal was to raise awareness about how easily deepfake technology can deceive the public if used irresponsibly.

A year later, in 2019, the Speaker of the U.S. House of Representatives, Nancy Pelosi, became an involuntary subject of another manipulation using a cheapfake (without AI). The video, which spread quickly on social media, showed her speaking slowly and appearing drunk. While the technique was far simpler than the Obama case—only slowing down the footage—the impact was similar in terms of reach, causing various media outlets and political figures to question her state and capacity. Pelosi was forced to make official statements to debunk it, and the incident highlighted how even basic manipulations, without AI algorithms, can erode the reputation of key leaders and fuel narratives that damage their public image.

That same year, in Gabon, the prolonged absence of President Ali Bongo sparked rumors about his health and ability to govern, fueled by a video showing him speaking strangely and with unnatural gestures. Many citizens and opposition figures suspected it was a deepfake meant to hide the president's death or incapacity, creating such a level of distrust that it almost led to a coup attempt. Although it was later proven that the video was real, and Bongo's appearance was due to the effects of surgeries and medical treatments (which caused his absence), the case highlighted how vulnerable a government can be to the mere suspicion of audiovisual manipulation. Collective fear and the lack of consistent official information almost brought down a government, underscoring the potential of this technology to destabilize entire countries.

In 2021, Latvia experienced an incident that illustrated another risk of deepfakes: identity theft in high-level political or diplomatic meetings. Several European Parliament members, including Richard Kols, believed they were having a video call with Leonid Volkov, a Russian opposition leader and campaign chief for the late Alexei Navalny. However, they were actually speaking with an imposter using a real-time deepfake. The deception, attributed to two Russian comedians known as Vovan & Lexus, revealed the vulnerability of international institutions to increasingly

sophisticated digital manipulation strategies, which could alter sensitive geopolitical discussions if not detected in time.

Lastly, in 2022, the armed conflict in Ukraine provided the setting for the first major instance of deepfakes being used for psychological and military purposes. Videos emerged showing Ukrainian President Volodymyr Zelensky supposedly calling for his troops' surrender to Russian authorities, which were quickly debunked by Ukrainian authorities. Shortly after, a supposed mayor of Kiev convened European leaders in another fake recording. Both actions were attributed to Russia and reflected the growing use of deepfakes to demoralize the population and undermine the legitimacy of leaders in conflict scenarios. However, the rapid debunking demonstrated an advancement in awareness and the ability to identify and neutralize these manipulation attempts, marking a new chapter in the use of information (and disinformation) as a weapon.

The cases mentioned are among the most relevant, but they are not the only ones. Since 2019, deepfakes have significantly increased in the context of electoral campaigns around the world, highlighting cases in South Korea (2020) and the United States (2024), with the latter supposedly involving manipulations from China, Russia, and Iran. The U.S. case requires special attention, as by 2024, the use of deepfakes during elections has nearly become normalized in American society, alerting experts like Donie O'Sullivan. In a CNN interview[19], he stated that deepfakes, due to their ease of generation and impact, are becoming a powerful political disinformation tool accessible to anyone. He also pointed out that in creating deepfakes, whether audio or video, authors use the same medium that candidates use to promote themselves: their appearance in the media. Photographs, combined with massive speeches, provide all the necessary data to train the algorithms, resulting in a deepfake. Combined with the spread of political propaganda on social media, this forms a fertile ground for the propagation of these artificial productions. Speaking of deepfakes now means discussing a reality that, according to current trends, will only continue to grow. Mitigating this will depend not only on improving detection systems but also on educating the public about the potential effects and psychological vulnerabilities that deepfakes seek to exploit.

[19] Available in: https://edition.cnn.com/2024/01/24/politics/deepfake-politician-biden-what-matters/index.html