

ALERT



virus

'Deepfaked Democracy': The Next-Generation Menace to Global Politics

İhsan Faruk Kılavuz

'Deepfaked Democracy': The Next-Generation Menace to Global Politics

İhsan Faruk Kılavuz

© TRT TRAINING AND RESEARCH DEPARTMENT

ALL RIGHTS RESERVED

WRITTEN BY

İhsan Faruk Kılavuz

PUBLISHER

TRT TRAINING AND RESEARCH DEPARTMENT

September 2024

TRT TRAINING AND RESEARCH DEPARTMENT

AHMET ADNAN SAYGUN STREET NO:83 34347

ULUS, BEŞİKTAŞ

İSTANBUL / TÜRKİYE

TRT WORLD LONDON

PORTLAND HOUSE

4 GREAT PORTLAND STREET NO:4

LONDON / UNITED KINGDOM

TRT WORLD WASHINGTON D.C.

1819 L STREET NW SUITE 700 20036

WASHINGTON DC

researchcentre.trtworld.com

The opinions expressed in this discussion paper represent the views of the author(s) and do not necessarily reflect the views of the TRT World Research Centre.

"The assumption that seeing is believing makes us susceptible to visual deception."

Kathleen Hall Jamieson

Introduction

Although the rapid advancement of artificial intelligence in digital age has led to significant innovations and improvements across various sectors, it has also given rise to many serious concerns. Among the most troubling concerns is the development and growing use of deepfake technology. What was initially viewed as a novel advancement, celebrated for its ability to manipulate voice, images, and videos convincingly, has quickly evolved into a tool with far-reaching, malicious implications. Deepfakes, once seen as an impressive demonstration of AI's potential, are now being exploited for harmful purposes on a large scale.

The misuse of deepfake technology has dramatically expanded, ranging from the creation of fake and deceptive content aimed at harming individuals' private lives and reputations to the broader and more insidious goal of undermining the integrity of democratic processes. The potential for misuse has grown at an alarming rate. On an individual level, deepfakes can be weaponised to fabricate false narratives that inflict irreversible damage to a person's reputation, privacy, or personal relationships. These fake representations can lead to devastating personal consequences, including public humiliation, social isolation, and emotional distress, all while the victims struggle to prove the authenticity of their identity or the truth behind the manipulated content.

However, the societal impact of deepfakes is even more alarming. Their potential to cause widespread damage extends beyond individual harm, posing a severe threat to democratic institutions and governance. The manipulation of audio-visual content, combined with the rapid dissemination of such material via social media

and other digital platforms, distorts reality and fosters disinformation, leading to a deep erosion of trust in public institutions. This erosion of trust is particularly dangerous for democratic societies, where faith in institutions and the accuracy of information is vital to the proper functioning of elections, policymaking, and public debate. In this context, deepfakes represent a fundamental and critical threat to the foundations of democracy itself.

By distorting the line between truth and falsehood and by enabling the creation and spread of content designed to deceive or mislead, deepfakes threaten to undermine the very pillars upon which democratic societies rest. If left unaddressed, this technology could severely damage public confidence in electoral systems, democratic governance, and the media, essential for maintaining a healthy, functioning democracy and exacerbate social polarisation.

Therefore, deepfake technology must be viewed not only as a threat to individuals but as a broader societal risk that carries significant implications for democracy. Legal measures and regulatory steps to curb the individual-level harms caused by deepfakes must be expanded to encompass the broader threat they pose to democratic stability. This expansion should include the development of comprehensive legal frameworks that specifically address the unique dangers deepfakes present. These frameworks are essential to ensure that the democratic pillars of the current liberal order are safeguarded and preserved for the future. In this way, society can work toward mitigating the destructive potential of this technology and protecting both individual rights and democratic institutions.



(İsmail Aslandağ - Anadolu Agency)

Rewriting Reality: Deepfakes and Questionable 'Reality'

As with every aspect of human life, AI technologies have increasingly influenced communication, bringing both positive, life-enhancing benefits and particular challenges in the relationship between humans and machines. Among these challenges, 'deepfakes'—hailed as one of AI's most valuable technological advancements—stand out as one of the most pressing.

While deepfakes are recognised for their innovative potential, the significant risks they pose have heightened immediate concerns surrounding their use.

These concerns have been further exacerbated by the dramatic increase in the creation and accessibility of fake content brought about by deepfakes. In this context, AI-powered deepfakes are generating a misleading world where the 'artificial' is perceived as 'real,' thereby increasingly questioning the concept of 'reality.' (Karakoc&Zeybek, 2022) In other words, deepfakes abstract and synthesise the fake and the real from their literal meanings, creating a new definition of 'reality'.

This results in a reliability dilemma regarding what is seen, heard, and thus trusted as information or truth. Consequently, the adage 'seeing is believing' has lost validity due to the rise of deepfakes.

Deepfake techniques come in various forms and can

be applied to any digital media. The most prominent and widely used techniques include:

- 1. Face Swapping:** Users can place the face of a target onto another body, applying this technique to still images and videos. Simplified versions of this technique are accessible online via specialised apps.
- 2. Re-enactment:** The face of a target is mapped onto a user, enabling the manipulator to control the target's facial expressions and movements.
- 3. Lip Syncing:** Users overlay mouth movements onto a target video, and when paired with audio generation, this technique can make it appear as though the target is saying something they never actually said.
- 4. Motion Transfer:** A person's body movements from a source video can be transferred to a target in an authentic recording.
- 5. Image Generation:** Users can create new images, such as faces, objects, landscapes, or rooms.
- 6. Audio Generation:** A synthetic voice can be generated from a small sample of an authentic voice, combined with lip-sync tools to dub audio into existing clips.
- 7. Text Generation:** Users can produce artificial text, ranging from short comments on social

media or forums to long-form news or opinion articles. Artificial comments are primarily effective due to the broad range of acceptable variations in online content. (Smith&Mansted, 2020)

Although deepfakes have become widely integrated into human life through these techniques and have elevated the risk of disinformation to unprecedented levels, the digital age has already introduced significant risks associated with false information before deepfakes. What distinguishes deepfake technology from ordinary falsehoods and amplifies the risk is its ability to manipulate reality quickly and cheaply. In other words, this technology enables the mass production and public dissemination of 'synthetic content' nearly indistinguishable from genuine video and audio recordings. (Karakoc&Zeybek, 2022)

Therefore, deepfakes can be readily exploited for personal, professional, or political gain when left unchecked. (Inter-Parliamentary Union, 2024)

Given that we now live in a 'post-truth era' where high-persuasion disinformation campaigns are conducted using digital deception and objective truths have become less influential in shaping public opinion compared to emotions and personal beliefs, deepfakes have undermined individuals' tendency to trust what they see with their own eyes and hear with their own ears.

As a result, the notion that 'seeing is believing' has become invalid due to deepfakes, giving way to numerous forms of misuse.

From Innovation to Exploitation: The Pitfalls of Deepfake Misuse

While deep fakes may not be inherently harmful and have legitimate and valuable applications in various social and economic fields—such as video game dubbing, entertainment, customer support services like call routing and reception, advanced medical research, healthcare, and infrastructure management—they also pose significant risks that cannot be ignored. (Smith&Mansted, 2020)

The primary risk associated with deepfakes arises from their ability to spread misinformation that appears to come from reliable sources, making them a powerful tool for malicious purposes. Indeed, the misuse of deepfake technology is increasingly shifting towards harmful applications.

In this context, public figures with numerous online videos and images, such as celebrities and politicians, are primary targets of deepfakes and are often the ones most adversely affected by their misuse. Notably, some of the earliest public instances of deepfakes occurred in 2017, when Reddit users employed AI-based face-swapping tools to superimpose celebrities' faces onto pornographic videos, with notable examples including fake explicit content featuring celebrities like Gal Gadot, Emma Watson, and Taylor Swift. (Smith&Mansted, 2020)

However, the misuse of deepfakes is not limited to fake pornography. This technology can also infringe on privacy by manipulating a person's image or voice without their consent, leading to significant emotional and psychological harm. Additionally, a person's facial features or voice can be replicated for fraudulent activities, identity theft, and other crimes, resulting in severe personal and financial damage.

Moreover, deepfakes can be weaponised more insidiously when combined with other technologies and social trends. (Smith&Mansted, 2020) For instance, fabricated audio evidence created with deepfakes could be used to influence or undermine legal proceedings, threatening the judicial system. (Inter-Parliamentary Union, 2024) They could also be employed to create false racist statements aimed at destroying someone's career or used by blackmailers to extort money or confidential information from individuals, as proving their falsehood is not always straightforward. People might prefer to accept financial losses rather than deal with the irreparable damage to their reputation caused by deepfake content.

While these risks are typically the first to come to mind when discussing deepfakes, it is vital to recognise the other, less discussed side of the issue: the misuse of deepfakes also presents a serious threat to democracies and warrants careful attention.

From Bot to Ballot: Deepfakes as a Risk to Democracy

Deepfakes present not only individual dangers but also collective threats that significantly impact society concerning the integrity of democratic processes.

Among the most urgent challenges to the established liberal order today is the erosion of trust in democracy and the growing scepticism toward democratic institutions—an issue exacerbated by the rise of deepfakes. In this context, deepfakes can potentially undermine critical events, such as elections, by manipulating political figures, spreading misinformation, and launching disinformation campaigns.

These actions erode public trust in democratic processes and intensify societal polarisation, thus threatening democracy's very foundations.

The rapid and cost-effective production and dissemination of false and misleading content through deepfakes further undermine societies' confidence in democratic systems, with this trust being a crucial pillar that sustains democratic governance. Moreover, deepfakes serve as a core tool in producing disinformation and online propaganda, granting unfair advantages or disadvantages to political actors, thereby complicating fair competition and manipulating voters in ways that prevent an accurate reflection of their will—a fundamental aspect of democracy.

While disinformation and deceptive online propaganda have long posed significant challenges to democratic processes by easily swaying public opinion and leading to unfair electoral practices, the ability of deepfakes to reduce the costs associated with creating and spreading fake content has intensified this threat, making such practices more pervasive and concerning. (Hanson&O'Connor&Walker&Courtois, 2019)

On the other hand, beyond generating new disinformation, deepfakes are also employed to enhance the effectiveness of traditional manipulation campaigns by addressing their weaknesses. Traditional online propaganda often relies on fake accounts or bots, easily detected due to their lack of credible digital identities or online interaction histories. However,

deepfake-generated content enhances the perceived authenticity of these efforts, thereby increasing their overall impact and effectiveness. (Smith&Mansted, 2020)

It is crucial to acknowledge that while the threat deepfakes pose to democracy has gained more attention recently, it is not an entirely new concern. Media manipulation has existed for as long as media itself, albeit with varying degrees of harm. (Burkell&Gosse, 2019) However, what is critical to understand is that deepfakes have significantly escalated this threat. Earlier forms of false content, such as 'cheapfakes,' were more rebuttable due to their low quality, limiting their adverse effects on democratic processes. (Pawelec, 2022)

In contrast, the advent of deepfakes has enabled the creation of sophisticated manipulation techniques that produce video, audio, and image content almost indistinguishable from reality, thus amplifying the harmful effects of disinformation on democratic processes.

In essence, the threats posed by deepfakes to democracies manifest in four primary ways: the increased vulnerability of elections to foreign interference, candidate manipulation, erosion of trust, and the deepening of societal polarisation.

1) Digital Trojan Horse: Deepfakes Catalyse Overseas Interference in National Electoral Integrity

The first significant threat posed by deepfake technology to national democracies is its potential to open the door for foreign actors to interfere in countries' democratic processes, particularly elections. (Yont, 2024) These foreign actors, including state-backed groups, intelligence agencies, and even organised crime syndicates, actively employ manipulation and propaganda to undermine the democratic integrity of rival nations.

While states have the resources and capabilities to conduct such activities, they often prefer to

use proxies to evade accountability. This is where deepfakes become particularly useful, as their effectiveness in creating fake political content allows these interventions to be carried out more covertly by proxies. With deepfakes making the production of false content cheaper and more accessible, non-state actors have gained significant influence in this arena. (Smith&Mansted, 2020)

The Russian-style propaganda model serves as a prime example of such proxy intervention, with deepfake technology playing a significant role. This model prioritises the rapid dissemination of misleading information over its consistency, and deepfakes have enhanced the volume and speed of this content. (Paul&Matthews, 2016) Moreover, deepfakes have improved the coherence and persuasiveness of false information, enabling actors employing the Russian-style propaganda model to interfere in the elections of other nations more effectively.

Prominent examples include Russia's 'Internet Research Agency,' a troll farm based in St. Petersburg, which actively worked to influence American politics before the 2016 U.S. Presidential election, operating with a substantial monthly budget of \$1.25 million. (Smith&Mansted, 2020) Similarly, during the 2020 U.S. Presidential election campaign, Russian operatives, such as the group 'Storm-1516,' used disinformation campaigns, including shallow AI-generated content, to sow discord and undermine U.S. support for Ukraine. These efforts involved deepfake-like tactics to launder disinformation through fake whistleblowers and citizen journalists, making the narratives seem credible and widely shared among U.S. audiences.

Moreover, Russian operatives used a fake news platform, Peace Data, which employed AI to target left-leaning voters with manipulated articles and videos aimed at weakening support for Joe Biden in the 2020 U.S. presidential election. By using fake personas and digitally altered content, they amplified narratives designed to confuse voters and disrupt the electoral process. Additionally, deepfake videos targeted Ukrainian Support in the U.S., during election periods, Russian operatives produced AI-enhanced content that misrepresented political figures and their stance on issues like the conflict in Ukraine. These operations aimed to manipulate U.S. public opinion by fabricating conversations or statements that did not occur. Another example can be seen in Moldova, where President

Maia Sandu and her pro-Western government were repeatedly targeted by AI-generated deepfakes designed to mock them. Moldova's national security officials attributed these attacks to the Kremlin, viewing them as a new tactic in a longstanding interference campaign. (Necsutu, 2023)

2) The Role of Deepfakes in Political Smear Campaigns

The second major threat that deepfakes pose to democracies lies in their ability to facilitate smear campaigns against political candidates and figures, thereby undermining the core democratic principles of fair competition and transparency.

Deepfake technology allows for the creation of convincing fabricated content that portrays politicians engaging in acts they never committed—such as accepting bribes, making racist statements, or issuing diplomatically sensitive remarks. (Yont Gokcen, 2024) These digital fabrications are potent tools for distorting voter perceptions and unfairly influencing election outcomes.

As demonstrated in the 'Cambridge Analytica' scandal, the misuse of personal data to craft manipulative political narratives has already shown how misleading and inflammatory content can impact significant events like the U.S. Presidential elections and the Brexit referendum. With the advent of deepfakes, this kind of disinformation has become even more pervasive, posing a severe threat to the very foundation of democracy.

Indeed, numerous recent cases highlight this growing concern. One notable example occurred in 2019 when Italy's former Prime Minister Matteo Renzi appeared on a satirical news program, seemingly insulting his political opponents, threatening to resign, and using inappropriate language. The video sparked widespread outrage but was later revealed as a sophisticated deepfake parody. (Inter-Parliamentary Union, 2024) Similarly, U.S. President Joe Biden fell victim to a deepfake campaign in January, when Lingo Telecom, a telecom company, broadcasted a fake voice message purportedly from Biden, urging voters in New Hampshire to abstain from participating in the Democratic primary—a blatant attempt to manipulate his voter base with statements he never made. (Shepardson, 2024)

Furthermore, politicians have sometimes exploited deepfake technology for their own political gain—a tactic known as the 'Liars' Dividend,' which introduces another threat to democratic processes. For instance, during an election campaign in Poland, the political party of then-Prime Minister Donald Tusk shared a deepfake audio clip of his opponent on social media to gain a political advantage. (Scott, 2024) Similarly, in the United States, Florida Governor and former Republican presidential candidate Ron DeSantis used an AI-generated image of former President Donald Trump to discredit his rival. (Contorno & O'Sullivan 2023)

It is also important to note that the malicious use of deep fakes involving political figures is not limited to political motives. For example, deepfake videos featuring Sir Keir Starmer and Prince William have been used to promote financial scams on social media platforms like Facebook and Instagram. These AI-generated versions of the Prime Minister and the Prince of Wales were featured in over 250 advertisements promoting a fraudulent cryptocurrency trading platform aiming to deceive and manipulate investors. (Cluley, 2024)

3) Trust at Risk

Closely related to candidate manipulation, another significant damage deepfake content inflicts on democracies is the erosion of trust, a fundamental pillar of democracy.

For democracy to function effectively, it relies on a baseline level of trust within society, primarily upheld by ensuring 'information security.' The issue of trust erosion emerged concurrently with modern democracy. As democracy has been increasingly embraced, demanded, and recognised as a cornerstone of the liberal order, it has become a prime target for political exploitation. Undermining democracy's integrity and weakening its institutions is most effectively achieved by eroding public trust in it, its institutions, and its practices. Before the rise of deepfakes, trust erosion was primarily based on 'lies told by people.' However, with the advent and growing prominence of deepfakes in political life, 'lies not told by people' are becoming the primary drivers of trust erosion.

Deepfakes undermine trust in two main ways. First, by injecting falsehoods into the marketplace of ideas, they mislead and misinform the public, disrupting the healthy functioning of democratic processes. Second, they generate widespread 'collective scepticism,'

causing people to view uncomfortable but accurate information with excessive doubt. Cognitive biases already predispose individuals to resist such truths, and the scepticism fuelled by deepfakes intensifies this tendency, providing an easy excuse to dismiss unwelcome evidence, even when it is entirely truthful. (Chesney&Citron, 2018) In other words, as fake content becomes more prevalent, people may increasingly doubt the validity of the information they encounter, even if it is accurate, making it harder to trust what they see or hear—ultimately undermining the healthy functioning of democracy.

On the other hand, it is also crucial to highlight that the widespread erosion of trust in society can be a powerful tool for autocratic regimes, presenting a significant threat to democracy. (Gregory, 2021)

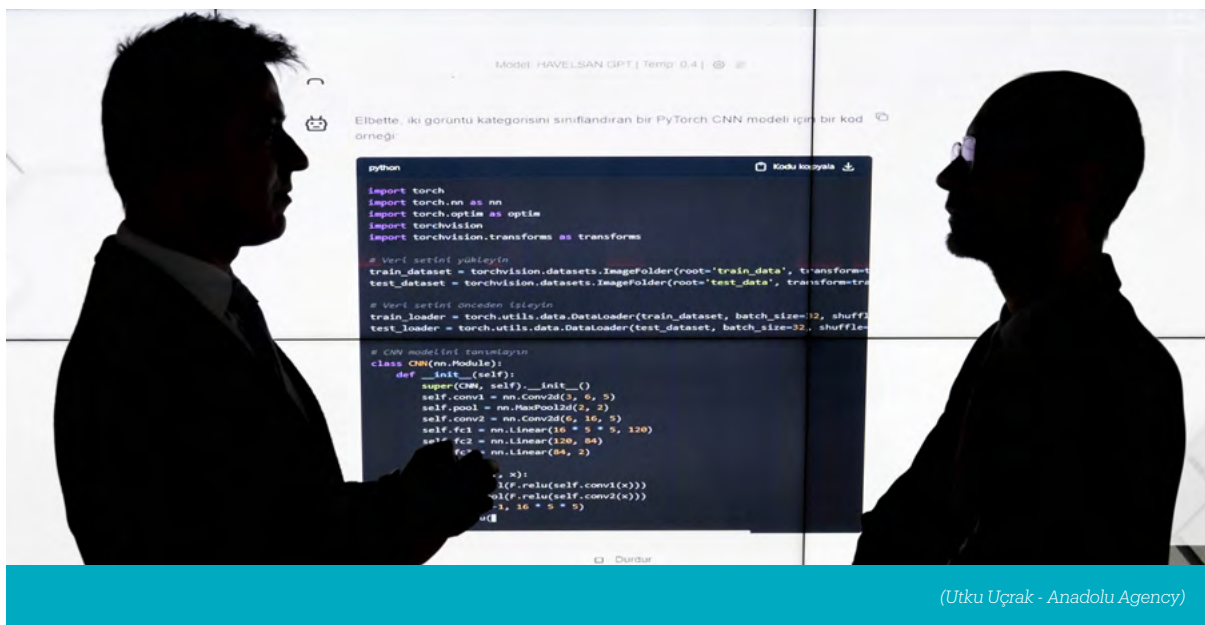
When trust in democracy, its practices, state institutions, the media, and even in one another is undermined, it creates an ideal opportunity for autocratic leaders to step in and fill the void.

As individuals and voters become uncertain about what to trust, they become more vulnerable to manipulation and exploitation by autocratic actors. These regimes can exploit this erosion of trust to present their controlled flow of information as the only truth, coercing society into accepting these manipulated narratives. In an environment where confidence in independent information sources has diminished, the propaganda of autocratic leaders becomes increasingly compelling. Furthermore, the resulting social distrust and chaos can be leveraged by autocratic regimes to legitimise their promises of 'restoring order.'

As public faith in traditional democratic institutions wanes, autocratic regimes can consolidate their power by promoting narratives of 'strong leadership' and 'stability.' Therefore, trust erosion should be viewed not only as a weakening of democratic values but also as a threat that could facilitate the rise of autocratic rule.

4) From Lies to Polarisation: Deepfakes Deepen Social Fault Lines

Finally, a crucial aspect of the threat posed by deepfakes to democratic principles is their role in generating disinformation and false propaganda, which exacerbates existing ethnic, religious, and political tensions in regions with active social fault lines.



This technology enables the creation of compelling but entirely false content, which can heighten social tensions and fuel conflicts. Indeed, attempts to exacerbate societal polarisation through contextually misleading images and false claims spread via social media have been observed in countries with sharp and active social fault lines, such as India, Myanmar, Sri Lanka, and Bangladesh. (Smith&Mansted, 2020)

Furthermore, similar manipulative tactics have also been noted in countries where societal polarisation is generally perceived as more manageable and

controlled. For example, the fabrication of a video falsely depicting former U.S. President Barack Obama calling Republican Donald Trump a 'total and complete dipshit' exemplifies how deepfakes can intensify polarisation within American political discourse. (Vincent, 2018)

In conclusion, regardless of the potential for social conflict, the creation and dissemination of deepfake fake content can undermine democratic processes, rendering them a profound threat to the stability of democracies.

Legal Responses must Be Created to Deepfake's Subversion of Democracy

The measures taken by authorities against the threats posed by deepfakes primarily focus on the risks they present to individuals. In other words, while states combat deepfake content, they prioritise safeguarding their citizens' interests since the photos and videos used to create deepfakes often contain individuals' data. In this context, states typically approach the issue from a human rights law perspective when implementing protective measures for their citizens, as producing a deepfake of an individual without legitimate interest or explicit consent constitutes a violation of privacy law, which is one of the core areas of human rights law. (Yont, 2024)

However, it is crucial to emphasise that since faked content can be used not only to harm an individual personally but also has the potential to undermine collective democratic efforts, separate regulations and measures aimed at preventing deepfakes from targeting democracies should be enacted alongside existing individual-focused deepfake protections.

In other words, while the digital society is sufficiently vigilant against the risks deepfakes pose to individuals, and the legal regulations and measures currently in place are primarily centred on this issue, it is essential also to consider the risks deepfakes pose to democracies. To protect democratic principles and

institutions, additional measures and legal frameworks should be implemented, as the current ones are insufficient in addressing these broader risks.

To effectively mitigate the harmful effects of deepfakes on democracies, the first crucial step should involve detecting and flagging such content. In this regard, governments could impose legal obligations on mainstream and social media companies, particularly their policy departments, to implement technologies that identify deepfakes and alert users—a practice already adopted by some platforms. Mandating these technologies, rather than leaving them to the discretion of media companies, would contribute significantly to preventing political manipulation through misleading content.

Additionally, the deliberate use of deepfakes, especially in ways that threaten election security or mislead the public, should be met with strict criminal penalties. Legal provisions could be introduced to impose severe fines and prison sentences on individuals involved in producing or disseminating such content. At the same time, regulations should address those who negligently spread false or suspicious information via mainstream or social media. Although their penalties may be less severe, holding them accountable is essential to curbing the spread of disinformation. No one contributing to disinformation should escape consequences by claiming a lack of intent. As Judge Oliver Wendell Holmes pointed out over a century ago, falsely shouting fire in a crowded theatre can cause unnecessary panic and disrupt order. (Chesney&Citron, 2018) Similarly, today's social media platforms have become the modern 'crowded theatre,' with deepfake content as the 'false fire alarm.' Therefore, tackling these issues is essential to prevent societal unrest.

It is also important to consider potential concerns regarding freedom of expression. Measures to restrict disinformation on platforms that influence democracy, especially social media, do not inherently conflict with free speech. On the contrary, safeguarding freedom of expression depends on maintaining a well-functioning democracy free from the dangers of disinformation. However, banning all deepfakes could violate democratic principles and constitutional protections in many countries. (Yont, 2024) Therefore, balancing necessary restrictions and freedom of expression is vital.

Another appropriate legal measure would involve establishing a temporary regulatory framework to limit the use of deepfakes, particularly during election periods, to protect democratic processes.

Finally, alongside legal regulations, raising public awareness about deepfake technology and its potential misuse is critical. Media literacy programs and public awareness campaigns can help citizens better identify misleading content, making society more resilient to deepfake attacks. This is especially important in countering the 'sleeper effect,' where individuals, even after recognising content as fake, may still internalise its influence over time.

Collaboration with social media platforms is also essential, given their role in rapidly creating and spreading deepfake content. Social media companies often act faster than lawmakers and could offer valuable expertise in addressing this issue. For instance, Facebook's 2019 'Deepfake Detection Challenge' and the 'Deep Trust Alliance,' a global platform fighting disinformation, have worked to protect elections from deepfake interference.

Considering that legal measures tend to lag behind technological developments and social media platforms are well aware of the deepfake problem, it becomes clear that cooperation between policymakers and the technology industry is crucial. Moreover, social media platforms not only facilitate the creation and spread of false content, but their vast user base often consumes information without scrutiny. Recent events have shown how false claims can go viral with unprecedented success, amplified by the ubiquity of social media. (Chesney&Citron, 2018)

Applications like 'FaceSwap' and 'Zao' allow everyday users to create deepfake content, quickly gaining traction on platforms like Instagram, Facebook, X, and TikTok. (Smith&Mansted, 2020) For example, deepfake videos featuring Vladimir Putin and Kim Jong-un, allegedly warning the American public ahead of U.S. elections, spread rapidly on social media, confusing being debunked.

However, while deepfake content targeting democracies poses a common threat to all democratic societies, a unified legal response has been lacking due to differences in national legislation. For instance, while some U.S. states have taken initiatives to address the

issue, these efforts vary in scope, penalties, and focus, making them far from comprehensive. Therefore, rather than relying on individual state efforts, collective international action is needed to establish uniform regulations and ensure faster, coordinated responses.

Ultimately, addressing the threats posed by deepfake technology to democracies requires global cooperation

and legal measures. These measures should carefully consider issues such as the legitimacy of deepfake content, the algorithms used to create them, the transparency of data storage and protection processes, and the procedures for data deletion. By considering these factors, governments can more effectively safeguard democratic institutions from the dangers of deepfakes.



(Rasid Necati Aslım - Anadolu Agency)

Conclusion

All in all, beyond the threats deepfake content poses to individuals, it also shakes the foundations of democracy in various ways. Foreign interference in elections and smear campaigns targeting candidates can mislead voters, undermining the core of public discourse and societal decision-making. Moreover, the widespread use of deepfakes risks eroding public trust in the authenticity of the information, as people may grow increasingly doubtful of visual and auditory proof. (Yont, 2024) While deception in politics is not a new tactic, the degree of manipulation and the speed at which falsehoods can be spread have transformed

our relationship with the truth. (Inter-Parliamentary Union, 2024) Social media has played a significant role in this, providing deepfakes with the tools for rapid production and dissemination, thus turning them into a pressing security issue in political communication. Consequently, authorities must expand their efforts to safeguard democratic pillars, which are already under attack on multiple fronts, by addressing the dangers posed by deepfakes. This calls for preventive measures led by legal regulations and bolstered through collaboration with social media platforms to protect the integrity of democracy.

Bibliography

Burkell J., Gosse C., (2019). "Nothing new here. Emphasizing the social and cultural context of deepfakes", *First Monday*, 24(12). <https://firstmonday.org/ojs/index.php/fm/article/view/10287>

Chesney R., Citron D., (2018), Deepfakes: "A looming crisis for national security, democracy and privacy?", *Lawfare*. <https://www.lawfaremedia.org/article/deepfakes-loomng-crisis-national-security-democracy-and-privacy>

Cluley G., (2024), "UK Prime Minister Keir Starmer and Prince William deepfaked in investment scam campaign", *Bitdefender*. <https://www.bitdefender.com/blog/hotforsecurity/uk-prime-minister-keir-starmer-and-prince-william-deepfaked-in-investment-scam-campaign/?srsltid=AfmBOooV8GmMd9kP8Dtx8yzp7ocI2oqIz-969hP9BOBTeZpwREliziTwN%2F>

Contorno S., O'Sullivan D., (2023), "DeSantis campaign posts fake images of Trump hugging Fauci in social media video", *CNN*. <https://edition.cnn.com/2023/06/08/politics/desantis-campaign-video-fake-ai-image/index.html>

Gregory S., (2021), "Authoritarian regimes could exploit cries of 'deepfake'", *Opinion Wired*. <https://www.wired.com/story/opinion-authoritarian-regimes-could-exploit-cries-of-deepfake/>

Hanson F., O'Connor S., Walker M., and Courtois L., (2019), "Hacking democracies", *The Australian Strategic Policy Institute* <https://www.aspi.org.au/report/hacking-democracies>

Karakoc E., Zeybek B., (2022), "Is seeing enough to believe? Political deepfake contents as a distinctive form of visual disinformation", *Oneri Dergisi*, 17(57), 50-72. https://dergi-park.org.tr/tr/pub/maruoneri/issue/68283/908542#article_cite

MacDonald A., (2022), "The uses and abuses of deepfake technology", *Canadian Global Affairs Institute*. https://www.cgai.ca/the_uses_and_abuses_of_deepfake_technology

Necsutu M., (2023), "Moldova dismisses deepfake video targeting president Sandu", *Balkan Insight*. <https://balkaninsight.com/2023/12/29/moldova-dismisses-deepfake-video-targeting-president-sandu/>

Paul C., Matthews M., (2016), "The Russian 'Firehose of Falsehood' Propaganda Model", *RAND*. <https://www.rand.org/pubs/perspectives/PE198.html>

Pawelec M., (2022), "Deepfakes and democracy (Theory): How synthetic audio-visual media for disinformation and hate speech threaten core democratic functions", *Digital Society*, 1:19. <https://link.springer.com/article/10.1007/s44206-022-00010-6>

Scott M., (2024), "Deepfakes, distrust and disinformation: Welcome to the AI election", *Politico*. <https://www.politico.eu/article/deepfakes-distrust-disinformation-welcome-ai-election-2024/>

Shepardson D., (2024), "Lingo Telecom agrees to \$1 million fine over AI-generated Biden robocalls", *Reuters*. <https://www.reuters.com/technology/artificial-intelligence/lingo-telecom-agrees-1-million-fine-over-ai-generated-biden-robocalls-2024-08-21/>

Slaughter A., Castleberry A.C., (2019), "ISIS 2.0 and the information war", *Project Syndicate*. <https://www.project-syndicate.org/commentary/america-islamic-state-information-war-by-anne-marie-slaughter-and-asha-c-castleberry-2019-09>

Smith H., Mansted K., (2020), "Weaponised deepfakes: national security and democracy", *International Cyber Policy Centre*. <https://apo.org.au/node/303485>

Vincent J., (2018) "Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news", *The Verge*. <https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peelee-buzzfeed>

Yont S., (2024), "The deepfake menace: Legal challenges in the age of AI", *TRT World Research Centre*. https://researchcentre.trtworld.com/wp-content/uploads/2024/03/The-Deepfake-Menace_v2.pdf

"Danger of deepfakes for parliamentarians", *Inter-Parliamentary Union*. <https://www.ipu.org/news/news-in-brief/2024-02/dangers-deepfakes-parliamentarians>

