



**ASSOCIATION OF
COMMUNICATION SCHOLARS
& PROFESSIONALS OF NIGERIA
BOOK SERIES**

Deepfake and Ethical Concerns in Journalism Practice in Nigeria

*Aliyu-ohiare Fati; &
Onakpa Mohammed Shaibu*

To cite this article:

Aliyu-ohiare Fati & Onakpa Mohammed Shaibu (2025), Deepfake and Ethical Concerns in Journalism Practice in Nigeria in Digital Communication and Governance in Africa: Reimagining AI in Communication and Governance ACSPN Book Series 10, pp.89— 114
Retrieved from <http://www.acspn.com.ng>

(Institutional or Subscribed access may be required)

Please note that this is an electronic, pre-print version of this article published by the Evan Brothers (Nigeria Publishers) Limited for Association of Communication Scholars & Professionals of Nigeria (ACSPN).

The ACSPN Book Series is exclusive to the Association of Communication Scholars & Professionals of Nigeria.

ISBN: 978-978-690-135-0

Copyright © 2025
Editor and Contributors

Layout/Design: Ashroph De Grafico Ventures.

All rights reserved No part of this book may be used or reproduced in any form or any electronic or mechanical means, including information storage and retrieval systems, without permission in writing from the publisher, except by reviewer who may quote brief passages in a review.



Deepfake and Ethical Concerns in Journalism Practice in Nigeria



*Aliyu-Ohiare FATI; &
Onakpa Mohammed SHAIBU*



A deepfake is a digital fabrication of both image and sound that uses a neural network to mimic a person's speech patterns, movements, and expressions. This can result in a fake video where someone appears to be saying or doing things that never actually occurred. The rise of deepfake videos and images has had a significant impact on social and political dynamics. This deceptive content has raised concerns about the ethical implications of new technologies on journalism. Journalistic ethics guide photographers on what actions are

acceptable when processing photographs. The foundation of journalistic ethics includes accuracy, truth, visual transparency, and respect for the digital photographic process. However, the increasing demand for instantaneous news updates has led news organisations to compromise the fundamental element of accuracy. Some news organisations have even argued that speed of reporting holds more importance than established photography ethics in the rapidly evolving digital era. This approach not only raises ethical questions but also contradicts the long-standing principles upon which journalism is built.

In the face of the increasing amount of user-generated content and amateur image manipulation, it is vital to recognise that these practices consistently violate journalistic ethics. The use of image manipulation techniques like editing and deepfake technology presents a clear ethical dilemma. These manipulations severely distort the truth, sway public opinion, and undermine the trust that journalism is based on. It is crucial for these unethical practices to be met with strict consequences. The rise of deepfake technology has created unprecedented challenges for photojournalism and news agencies. As news organisations work to stay competitive in the digital age, the harmful effects of deepfake technology cannot be overstated. McCammon (2022) maintains that this poses a significant threat to the integrity of visual storytelling and worsens the already shaky position of traditional newsrooms. In a time where the demand for instant news is high and digital news outlets have a clear advantage, the business model for traditional newsrooms is facing major disruption.

It is important to acknowledge the significant change in this situation and the intense pressure it puts on news organisations. As they strive to provide immediate news and meet the relentless demand for up-to-the-minute updates, newsrooms are faced with a difficult dilemma: how can they maintain accurate

reporting and visual credibility while also adapting to the requirements of the digital era? This is a question that requires thoughtful consideration, innovation, and a renewed dedication to journalistic principles. The current state of journalism is experiencing a conflict between the need for speed and the necessity of accuracy. The emergence of amateur photo editing methods and the unsettling existence of deepfake technology pose substantial ethical challenges for news organisations. As the digital era continues to reshape the media landscape, it is essential for newsrooms to find a delicate balance between providing real-time experiences and upholding unwavering commitment to journalistic ethics. Only by doing so can news organisations preserve their credibility, rebuild public trust, and navigate through these uncertain times with integrity and resilience (Singh & Dhiman, 2023).

Background and Definition of Deepfake Technology

Deepfake technology represents a significant threat to the news industry. This sophisticated technology, based on artificial intelligence advancements, can create highly realistic visuals and audio that have the potential to deceive global audiences. Its ease of use allows for rapid creation of manipulated content, exacerbating concerns about its widespread manipulation. The implications of deepfakes include troubling applications, such as the creation of explicit adult videos using non-adult material, raising ethical concerns and potentially damaging individuals' lives. Additionally, the malicious use of deepfakes can propagate fake news, compromising the integrity of trusted sources and undermining public trust (Whyte, 2020).

However, the ramifications of deepfakes go beyond individual harm and reach into the realms of historical manipulation and fundamental security threats. With the ability to fabricate convincing audio-visual representations, malicious actors can exploit deepfakes to rewrite historical events, distorting

collective memory, and potentially redefining our understanding of the past. This distortion of history could have catastrophic consequences for societies, as it challenges the core foundations upon which our knowledge and understanding are built

In addition to the historical and social implications, there are grave concerns regarding the use of deepfakes in espionage, corporate sabotage, and the erosion of personal privacy and security. By seamlessly merging artificial imagery and fabricated audio, deepfake technology has the capacity to deceive even the most discerning individuals, making it an attractive tool for espionage operations. This poses a significant threat to national security and the integrity of confidential information. Moreover, corporations may suffer severe financial losses due to deepfake-aided sabotage, which allows for the creation of fabricated evidence and the manipulation of narratives (George & George2023).

As we face the challenges posed by deepfakes, it becomes evident that decisive measures must be taken to address their misuse. Establishing effective regulations and frameworks to oversee the creation and spread of deepfakes is crucial in preventing their harmful impact. Furthermore, it is crucial to increase public awareness of the existence and potential risks of deepfakes. Educating people about the complexities and dangers of this technology can empower them to analyze the content they encounter and make informed decisions. Additionally, it is important to prioritise the development of strong measures that can accurately detect and mitigate the negative effects of deepfakes. Progress in artificial intelligence and machine learning can contribute significantly to this effort, as they enable the development of advanced algorithms capable of reliably identifying manipulated content. By investing in research and development, we can equip ourselves with the necessary tools and techniques to effectively combat the harmful effects of

deepfakes (Samuel-Okon et al.2024).

The rise of deepfake technology presents a significant danger to the news industry and society in general. Its capability to create highly convincing images and sound could damage truth, honesty, and confidence. Immediate steps, such as proper regulations, greater awareness, and the creation of strong defences, need to be taken to safeguard individuals, maintain historical truth, protect national security, and ensure the privacy and safety of personal information. It is only through united efforts and ongoing advancements that we can lessen the risks of deepfakes and maintain the genuineness and truthfulness of our information environment (Whyte, 2020).

The rapid spread of deepfake videos is a major concern due to their exponential dissemination. A single manipulated video can garner more than 40,000 views within 24 hours, posing a significant challenge to journalistic integrity. Videos created by shallow learning exhibit distorted characteristics, such as over-smoothness and crude mismatched faces, undermining their factual communication. The term "deepfake" comes from 'deep learning' and 'fake', referring to videos or images created by deep learning AI models. While shallow learning videos raise ethical concerns, deep learning makes replication faster and more convincing. Digital platforms are investing significant computational resources to screen potentially deepfake videos before posting, and the widespread use of social media ensures that the negative impact of deepfake technology will permeate various aspects of society. Consequently, deepfake raises ethical concerns in areas such as media ethics, revenge porn, political affairs, conspiracy, and public trust (Frank, 2022).

Importance of Journalistic Ethics

It is crucial to recognise the significance of ethical considerations in journalism due to the widespread interest in news and the

rising popularity of journalistic practices. As such, it is important for individual journalists to adhere to the principles of "Journalistic Ethics," which encompass their moral obligations and professional responsibilities. These principles are typically outlined in a comprehensive code of ethics established by professional associations to provide guidance for the journalistic profession.

By following these guidelines and demonstrating professional conduct, journalists can earn and cultivate the essential trust of the public in news outlets. When journalism is seen as solely profit-oriented, the significance and influence of these ethical duties, such as the critical responsibilities of confirming credible information, relentlessly pursuing the truth, reducing harm to the subjects of stories, and maintaining complete transparency, may unfortunately decrease (Oso et al, 2024).

Readers must be vigilant in recognising sensational or biased journalism and take a stand by boycotting such sources. This erosion of trust in the media could have dire consequences, leading to a decline in the value of journalism and undermining the reliability of information. This not only affects journalists and news organisations but also impacts the public who depend on trustworthy sources for making important decisions. The dissemination of false information threatens the core principles of democracy, such as transparency and an informed public. It is crucial for readers to understand the severity of this situation and take steps to support truth and accuracy. By being aware and developing media literacy skills, readers can play a key role in upholding the integrity of journalism and ensuring a well-informed society (Fawzi et al.2021).

The term "yellow journalism" was coined in the mid-1890s and gained widespread use in the 1920s. As newspapers focus more on making money than on true journalism, metatheses play a

significant role in leading away from authentic reporting. These metatheses become mere traits, compromising the integrity of news and presenting a false democratic front, as if packaging news as mere products. In today's world, various digital tools allow people to have more control over their actions and navigate through a flood of daily information. Therefore, it is no longer possible to ignore the wrongful use of journalism, as the potential for responsible engagement now exists (Huang, 2023).

History and Evolution of Deepfake Technology in Nigeria

The deepfake technology, also called synthetic media, is a rapidly advancing type of artificial intelligence that allows for the creation of manipulated and misleading audio-visual content. It is closely linked to the concept of 'predictive technology', which involves the seamless blending of artificial intelligence techniques like face-swap and the powerful Generative Adversarial Network (GAN). The use of GAN enables the machine learning system to generate incredibly authentic data, including images, videos, audio, and printed material that is difficult to detect as fake. This groundbreaking technology, as described by Motherboard, revolutionizes the manipulation of visuals by producing incredibly realistic content that perfectly mimics real media, making it increasingly difficult to distinguish fabricated elements from reality (Whyte, 2020).

The basic structure and mechanism for creating a deepfake video involves the use of deep learning, an algorithm that imitates the working of the human brain with multiple artificial layers, and manipulation of the Generative Adversarial Network (GAN) neural networks to create a fully realistic video. Deepfake of Nigerian political office holders as of 2019 can be traced back to 2018, when the founder and president of United Peoples' Congress (UPC) in Nigeria, Nnamdi Nwokocha Aghai, released a video of himself as a presidential candidate having a pleasant bulk debate with himself. This video went viral on social media and it

was later discovered to be a fake video created with a deepfake app called Fotogenic. In September 2019, The New York Times used deepfake technology to build "The Facebook Falacy" from The New York Times Magazine into a moving, talking, and realistic digital replica of "Liz," a victim of digital deception on Facebook. The use of deepfake technology in this instance showcased the potential dangers of manipulated videos and the need for increased awareness and skepticism when consuming media online. It served as a powerful example of how deepfake videos can be used to spread misinformation and deceive the public, highlighting the importance of developing robust tools and strategies to detect and counteract the spread of deepfakes. As technology continues to advance and deepfake techniques become more sophisticated, it is crucial to remain vigilant and critical of the content we encounter, to ensure that we are not misled by these deceptive videos (Gregory, 2022).

Origins of Deepfake Technology

Originally, the terms "fake" and "deepfake" were inspired by 4chan and referred to incredibly realistic non-consensual sexual images and videos made using AI-based image synthesis. Since then, deepfake has expanded to encompass a wide range of generated artificial video and audio content, including all types of spoofing. The most common deepfakes on the Internet use neural networks to impersonate real-life figures and place them in situations they were never part of, such as making a public figure say something they never actually said. This has raised concerns about potential threats to democracy and public trust. As technology advances, deepfakes are expected to become even more sophisticated and widespread, posing significant challenges for society and individuals. The rapid progress in AI has made it easier for those with malicious intent to create convincing deepfakes that are hard to distinguish from reality. These manipulated media not only undermine the credibility of public figures and institutions but also have the potential to incite

violence, spread misinformation, and manipulate public opinion on a large scale (Meikle, 2022).

The rise of deepfake technology raises important ethical and legal questions regarding privacy, consent, and the boundaries of free speech. It is crucial for governments, technology companies, and society as a whole to come together to develop robust strategies and tools to detect, combat, and mitigate the harmful effects of deepfakes. The development of cutting-edge algorithms, machine learning models, and forensic techniques is essential to ensure the integrity of media in the digital age. Additionally, media literacy and critical thinking skills need to be strengthened at all levels of education to empower individuals to distinguish between genuine content and manipulated deepfakes. As deepfake technology continues to advance, it is paramount that proactive measures are taken to safeguard against its potential misuse and protect the integrity of information in a rapidly evolving digital landscape (Wilkerson, 2021). In the meantime, however, the potential ramifications of poorly modeled speech synthesis have also raised an increasing number of concerns. Many experts, along with corps of technology and academic companies, keep releasing various datasets, benchmarks, tech demos, and publications. Infosec experts are also focused on ensuring that researchers' advances are in line with social and ethical considerations surrounding synthesised media.

Advancements and Applications

The rise of publicly available and incredibly advanced machine learning models has led to a surge in popularity for a wide range of applications. The GPT-3 model is a prime example, transforming the way we interact with technology. With its unmatched capabilities, GPT-3 can expertly handle user queries, act as a conversational partner, aid in copywriting, serve as a design tool, and even develop new skills autonomously. Its impact is felt in the education sector as well, as it can generate high-quality college

essays and create sophisticated source code snippets, pushing the boundaries of software development (Pereira et al 2023). Furthermore, GPT-3 is a highly effective tool for enhancing creativity in various fields. It has the ability to effortlessly create beautiful poems, displaying a level of artistic skill that can rival even the most talented poets. Additionally, GPT-3 helps writers improve their literary works by providing valuable insights into developing compelling characters, defining unique traits, and creating captivating scenarios similar to those found in popular video game styles. Beyond education and creativity, GPT-3 also plays a crucial role in the healthcare field. Here, it uses its exceptional capabilities to generate advanced NER (Named Entity Recognition) models, which are instrumental in accurately identifying and extracting important entities from large amounts of biomedical literature, revolutionising the processing of medical information and enabling groundbreaking discoveries (Akbasli et al, 2024).

The GPT-3 model's impact extends beyond education and healthcare to encompass business. Its guidance reduces the need for complex websites, simplifying information dissemination. It also analyses the effectiveness of mental health interventions, providing insights for decision-making. In talent acquisition, GPT-3 efficiently matches job offers with skilled applicants, improving recruitment processes. The model's influence in various fields is significant, and its potential for the future is limitless as it continues to evolve (Akbasli et al 2024).

Ethical Considerations in Journalism in Nigeria

The media plays a crucial role in society by being responsible for delivering news stories fairly, accurately, and truthfully. Considered the voice of the people, the media has a moral duty to uphold principles such as truth, accuracy, objectivity, impartiality, fairness, and public accountability. However, the threat of misinformation and personal and political interests may lead

journalists to act against ethical standards. The use of Deepfake technology to create false news could have significant implications for journalistic integrity in Nigeria. The digital age has eroded the truth and the moral authority of the media, making the public more susceptible to manipulation. The media holds a prime position as an influence peddler, perception moulder, and conscience groomer, and must reflect impartially on the good and bad around them. The media should adhere to professional standards and ethical codes, and avoid influencing, sensationalising, or commercialising information about events of public interest. It is important for all media content to be attractively presented while maintaining stringent adherence to professional standards. (Kondurkar *et al.* 2024).

Accuracy and Truthfulness

In journalism, the technique of deepfake has particularly invasive consequences. We need to define what "authenticity" means when it comes to genuine journalists. The first basic rule is accuracy, in other words, truthfulness. The Journalist's duty is to provide a mirror to society. Some use the term "mirror" to refer to poor journalism when it sometimes exercises power. Journalism, the mirror, should reflect the truth. Like everything else, deepfakes are not a totally novel form of illegal tactics used to generate fake reality. However, they are the first that have the ambition to be authentic, falsely authentic. Some argue that the professional use of journalistic creativity is based precisely on the listener: not only the enormous quantity, but also the news content, not just the content (Tian *et al.* 2024).

As we shall see, there is much more to this. However, it is important to specify that, just like music, film, or advertising, showing absolute news has an authentic deepfake as collateral. Sometimes, there are tricks to generate presence and the capacity to recognise fake reality within it. Deepfakes are only one branch of a function that is needed, which includes humour,

cloning, and misinformation as a result of the automation of creativity and precisely the garments that come with the ability. This means a world awaiting fake reality to be detected. However, if journalism does not have an effect in this specific matter, the fate of the deepfake will, inevitably, be the same indifference that has so far afflicted those that have, although literally identical. This gravity is not tempting and cannot be accepted. To properly address this destiny reels in the production process, the very nature consents, however, to beat the whole weapon of deepfakes. They need to be inadequately documented to generate not only information but also misinformation, and to have a 'realistic' impact on and patrols are needed to assist the operators' accuracy of which they are of the "mirror effect" used and accredited. The listening needs to glimpse a difference, violence of which they are used. If forced to believe in order to generate informative depth, the effects of using deepfakes significantly distort the mirror. To effectively activate and avoid this, proper measures need to be taken (Tian *et al.*2024).

Privacy and Consent

The rise of deepfake technology raises serious concerns among experts in journalism ethics, particularly in relation to issues of privacy and consent. According to Stacy (2019), the concerns regarding deepfakes center around the issues of privacy and informed consent. Existing regulations aimed at protecting individuals' privacy rights may not be equipped to address the ethical implications of deepfake usage, especially in comparison to previous regulations focused on nonconsensual sexually explicit content. Additionally, individuals who consent to the creation of deepfakes may unknowingly waive their privacy rights regarding the use of material featuring them. Even if a disclaimer is included stating that the individual must consent to the use of their material elsewhere, it may not fully address the potential privacy harm caused by deepfakes. Therefore, obtaining consent does not seem to provide a solution to the

broader problem of privacy harm associated with deepfakes.

Informed consent is closely linked to privacy considerations in deepfake content. With the significant risks associated with deepfake technology, obtaining fully informed consent from the responsible party is a major challenge. It is important to prioritise individual autonomy in order to address the exceptional risks posed by this technology. Users may consider obtaining consent from individuals whose images or likenesses are used in their content and digitally recording that consent as a potential solution, although it may not be ideal. Implementing digital notes would require the protection of copyright or similar laws to ensure that bending the requirements of informed consent does not completely remove it from the individual (Bousselham & Mourhir 2024).

Fairness and Objectivity

Prior to the release of official surveys and witness accounts, certain organisations may choose to generate a deepfake simulation of a fictitious news event, such as a fire in a public institution. Production companies are capable of manipulating videos in real-time using software like Adobe After Effects or Animo. Although these platforms may not be as advanced as the latest deepfake technology, they are still considered to be unjust and unethical. The significance of obtaining accurate and verified information through an impartial pursuit of the truth cannot be overstated. Journalists must ensure that they have taken into account the diversity of societies when reporting on a news event. Fair reporting necessitates thorough communication with sources, cross-verification of information, and thorough research for reliable data (Akbasli *et al.*, 2024).

Ultimately, the ethical element of fairness is closely tied to the tasks of checking, verification, and authentication. The term "challenge" accurately encompasses the difficult role of a

journalist in applying tools for creating hyper-real videos, which are beyond their control and carry significant power. This creates a barrier to verification and presents a complex issue as journalists must possess a wide range of skills, be well-informed of potential sources, and understand the necessary adaptations for experimentation in generating verifiable images using artificial intelligence algorithms. The inability to guarantee the certainty of a video's origin represents a setback in the journalistic mission and should prompt serious reflection (Raiaan et al.2024).

Challenges Posed by Deepfake Technology in Nigeria

In Nigeria, the increased use of the internet, including social media, has heightened the country's susceptibility to the potential impact of a deepfake scandal, posing a threat to national security. Nigeria's diverse population, encompassing various ethnicities and languages, has prompted ongoing debates about secession in different regions, often taking place on social media and sparking concerns among government critics. In 2016, the Nigerian Senate rejected a proposed Bill SB 408 aiming to regulate harmful content on social media, as it failed to address issues of libel. The emergence of deepfake technology could worsen this problem by producing realistic videos that could incite violence or civil unrest. This is particularly troubling given Nigeria's history of social unrest and political oppression. Similar to Senators who fear that seemingly innocuous social media posts could lead to protests, the Nigerian Government is prepared to potentially utilise deepfake technology to protect the country from ethnic division or to discredit politicians they believe are inciting it from external sources. (Nnamdi et al.2023)

Misinformation and Disinformation

The dissemination of false information can be more detrimental when it originates from individuals in positions of authority who have the ability to reach a large audience and influence many

people. Manipulation and manipulated media have become prevalent, with the availability of videos and deepfakes contributing to their proliferation. Many individuals now encounter a wealth of information and share it with the public without verifying its sources. This raises questions about the authenticity of news. The beneficiaries of false reports are organisations that sow disunity, advocate for divisive causes, prioritise their political agendas over the well-being of society, pose a threat to democratic values, and have connections to authoritarian regimes. The rise of digital production has underscored the importance of ethical considerations across various social sectors. Journalism, in particular, faces a slew of ethical dilemmas due to the widespread manipulation of content on social and digital media, as well as the increased dissemination of disinformation through these platforms (Igwebuike & Chimuanya, 2021).

While pointing out the spread of fake news and misleading information on the internet, the public is becoming increasingly critical and skeptical towards traditional media outlets and official sources. This scrutiny extends to fact-checking processes in public media, independent news websites, and investigative journalism platforms. However, the rapid rise of disinformation and manipulative content has led to journalism principles taking a backseat. This study aims to adapt online journalism practices to the current information landscape, which heavily features manipulated visual content. By shifting focus to critical ethical reflection and responsibility, the goal is to contribute to the literature, media, and public discourse. This adaptation seeks to identify new truths and raise awareness within the information community, thereby allowing journalism to fulfill its primary duties in the digital realm. The primary objective is to question the reliability of visually altered images such as deepfakes, and to explore the ethical responsibilities of journalists who produce, and the consumers who consume, similar content (Ojala, 2021).

Credibility and Trust in Journalism

The importance of the credibility and trustworthiness of news has been emphasised in multiple studies. Lee et al. (2012, p. 490) highlighted the need for intermediaries to maintain the trust of their audience in order to continue their role as market intermediaries. Without a foundation in facts, journalistic work cannot be considered legitimate, making the ability to discern facts from fiction crucial. The lack of trust in news has become so prevalent in the post-truth era that the term 'post-truth' was chosen as the Word of the Year by the Oxford Dictionaries in 2016. Bates (2018) also warned that the decline in trust and credibility in news reporting could present an existential crisis for journalism.

The debate on misinformation involves discussion on how to define and address it in order to prevent further damage to the media. Nagel (2018) emphasises the importance of trust in determining the value and reliability of information. Additionally, Ashoor, Peersman, and Harrington (2019) categorise 41 quantitative measures for assessing the credibility of news and social media content related to misinformation, including areas such as metaphors, frames, trustworthiness, and sentence complexity. These measures do not consider more advanced or straightforward methods for verification through software (Brahms, 2022).

Regulatory Frameworks and Legal Implications

In the field of political communication studies, the term "deepfake" has emerged as a method for making fake news more easily. The term "fake news" gained popularity after the 2016 American presidential elections, when then-candidate Donald Trump frequently used it. The term was used to describe information that promotes one's agenda and discredits opposing viewpoints. Fake news often originates from websites that specialise in creating manipulated information to attract viewers

by playing on emotions such as faith, anger, fear, and more. This type of misinformation is created with the intent to deceive. As a result, international organisations are pressuring Facebook to use the term "false information" instead of "fake news," as Trump has often described it. Therefore, fake news refers to deliberately fabricated news with the intention of causing confusion and undermining the credibility of journalists. Fake news can take two forms: entirely fabricated news and partial news, which typically comes from biased sources and aims to misrepresent certain events (Cassam, 2021).

A deepfake refers to the creation of fake images, videos, and audios using AI technologies such as Generative Adversarial Networks for images and Voice cloning for audios. These manipulated media can result in fake news and depict individuals engaging in activities they never actually performed. The intention behind deepfakes is to deceive viewers and spread misinformation through social networks. This type of digital manipulation can compromise the integrity of video content and have serious repercussions for the reputation and privacy of the individuals depicted. The use of other people's identities in deepfakes can also be considered identity theft. The term "deepfake" has been used in the context of political communication to describe videos that use deep learning techniques to alter the appearance and movements of real people (Brahms, 2022).

Current Laws and Regulations

The authors of this groundbreaking research have conducted a thorough analysis on the impact of deepfake technology on journalistic ethics. They have outlined laws and regulations related to technology, academic advancements, and journalism ethics, all of which are violated by deepfakes. The press recently announced that social media giants like Facebook and YouTube have pledged to remove deepfake videos before the 2020 US

presidential election. Out of 195 countries, only Pakistan and Uzbekistan have laws explicitly prohibiting deepfakes. In many cases, existing laws focus on discussions about deepfakes rather than direct prohibitions, with some penalties under defamation or libel laws. Singapore, the United States, and China have provided guidelines on labeling and identifying manipulated content resembling deepfakes. In Europe and the United States, legislative bodies are creating laws to facilitate sharing deepfake technology information and enhance capabilities to detect and identify perpetrators. In Nigeria, there are currently no specific laws addressing the threat of deepfake technology (Cassam, 2021).

In 2017, the South Korean Supreme Court introduced stricter punishments, including demerit points in criminal cases, for exaggerated sentences using post-editing or artificial recording. The Canada National Research Coalition on Cyber Security released a report referencing international law-related case precedents and recommending a study on deepfake-specific regulation. In Japan, deepfake is not governed by criminal law, and although there is a policy opinion on the matter, the government has yet to establish an official legal position. In Pakistan, only name liability is enforced in relation to deepfake creation, while Uzbekistan imposes a fine of 50 million won or up to 3 years in prison for manipulating deepfake to deceive others for personal gain. Magazine Weekly Sinmun has raised concerns about the impact of deepfake on public trust and democratic values, advocating for legislative measures to address the risk it poses. The Nigerian government has also initiated efforts to regulate the use of deepfake technology through new legislation (Spector, 2020).

These laws and regulations are a positive development, but there is still a significant amount of work needed to fully address the ethical implications of deepfake technology in journalism. For

instance, there is a necessity for stronger penalties for individuals who intentionally create and distribute deepfake content. Additionally, there needs to be a clear definition of what qualifies as deepfake content and the corresponding consequences for its creation and spread. Another crucial consideration is the potential impact of deepfake technology on privacy laws and individual rights. It is necessary for the legislation to comply with the constitution and international human rights laws. Importantly, it is noted that the given requirements in these discussions were not effectively enforced by the law (Bufacchi, 2021).

Proposed Solutions and Future Directions

The comprehensive research highlighted the challenges in addressing complex threats and suggested that promoting cooperation between different areas of expertise could be highly effective. It stressed the importance of addressing and reducing the risks associated with media manipulation. The study also underscored the significance of collaboration among researchers, scholars, and professionals in creating and sharing scientific knowledge and theoretical insights. While acknowledging the difficulty of the task, it emphasised the importance of cultivating a new generation of communication leaders with the necessary skills to critically evaluate and navigate a large volume of content. These individuals must be able to assess and interpret content with a clear mind, knowledge, and judgment, while understanding the theoretical principles that inform their decision-making in various cultural contexts (Calleja *et al.* 2021).

Without a strong sense of hope, recognition, and impartial commitment, the ability to manage and disseminate important messages that support vision and mission would be greatly compromised. It is this recognition that allows for more effective planning and assessment of pre-disaster risks and systemic

analysis. To achieve this, we have proposed a variety of measures tailored to address different aspects of this complex challenge. These measures include, but are not limited to: comprehensive education programmes focused on critical thinking, ethical decision-making, and media literacy; the creation of strong interdisciplinary research groups; the formation of international partnerships and collaborative networks; and the implementation of ethical guidelines and regulations governing media manipulation and image editing. By following these proposed measures and through interdisciplinary collaboration, we can advance our understanding and response to the challenges presented by media image manipulation. This joint effort will undoubtedly pave the way for a more informed, adaptable, and media-savvy society, equipped with the necessary tools to effectively navigate and interpret the intricate and constantly changing media landscape. With the unwavering commitment of all parties involved, there is great potential to strengthen societal resilience and protect the integrity and authenticity of our media environment, ensuring its continued ability to inform, inspire, and unite (Thomson et al.2022).

Policy and Regulation

Regulation and policy are crucial in preventing the misuse of deepfakes, and educating the public about their potential dangers is also important. Government interventions, such as requiring disclaimers when deepfake content is used in political campaigns, could help. Additionally, restricting deepfakes to specific uses and developing new media regulations to achieve these goals could be part of a broader effort to combat deepfakes. However, the current legislative framework makes it difficult for government actors to take proactive action, as technology companies are largely absolved from monitoring or altering the content they spread. It may be worth considering the private sector's role in voluntarily creating policies to limit the spread of harmful technological aberrations (Gregory, 2022).

Technological Solutions for Detecting Deepfakes

With the ongoing development of various deepfake applications, the increasing concern over the threat posed by deepfakes to the core principles of democratic society, particularly information integrity, has become more prominent. In 2019, there was a surge in the development of technological solutions designed specifically to detect deepfakes. These detection models are aimed at determining whether a given video is a deepfake by analysing the unique artifacts left behind during deepfake model training and the evidence of manipulation. The current technological solutions can be classified into three categories: model-based, artifact-based, and temporal-based. Model-based methods focus on identifying inconsistencies that are specific to the deepfake model used for manipulation, with a particular emphasis on deep learning-based manipulations (Jones, 2020).

Methods that use artifacts aim to detect any anomalies that may arise during the creation of deepfake videos. These anomalies can be the result of differences in the source materials and may manifest as inconsistencies in aspects such as alignment, occlusion, colour, noise, noise pattern, and frame quality. However, these methods often do not accurately reflect real-world properties. On the other hand, temporal-based techniques focus on identifying changes to artifacts that occur over the manipulation timeline. These techniques evaluate how frames change throughout the manipulated video, including factors such as illumination, reflection, and compression. In the case of deepfakes, these systems rely on a level of inconsistency for detection (Sandotra & Arora, 2024).

Conclusion and Recommendations

The emergence of deepfakes poses significant challenges to journalism and journalistic ethics. While the threat of deepfake hasn't completely materialised yet, the existing research has shown that they may proliferate in the near future. As they

become increasingly realistic and indistinguishable from real videos, it might undermine the editorial trust, enhance the spread of fake news and misinformation, and like the case of the Dangote deepfake???, may even succeed in impersonating high-profile figures and manipulate their reputations. To maintain public trust, the journalism industry can start to take some steps now to mitigate some of the harmful effects of deepfakes. As one industry observer wrote, "News organisations should prepare now for use of 'deepfake' technology for potential news manipulation" (Strong, 2018).

In addition to these measures, news organisations should also see it as their imperative to educate their audience about deepfakes. Not all deepfakes are sophisticated. Basic visual and audio cueing can help a careful observer tell a deepfake apart, but the audience may not be aware of these indicators. Simple deepfakes might slip through due to the audience's technological ignorance and news organisations' lack of prompt action due to technological indifference. This means a dual-pronged journalistic approach: technically-sophisticated organisations need to equip their journalists with the tools to immediately start tracking deepfakes being disseminated online while less-advanced newsrooms should focus on explaining and publicising what deepfakes are and how to spot them.

In order to have the greatest impact, creators of deepfake videos must also share or publicise their work. This presents a challenge to society as it encourages critical thinking and open discussion of verification systems and security protocols, as well as ethical concerns. To address this rapidly evolving phenomenon, society must acknowledge the widespread usage of deepfake technology and engage in meaningful conversations about its potential for artistic expression, information dissemination, and exploration of human rights. Through these discussions, we can better understand and navigate the challenges and opportunities

presented by this technological advancement (Sandotra & Arora, 2024).

While creating awareness about the existence and potential misuse of deepfake technology is an important first step, it is not enough to mitigate their harmful impact. It is crucial to critically analyse and openly discuss the logical and ethical inconsistencies present in suspected deepfake content in order to ensure transparency and accountability. Through in-depth analysis and thought-provoking conversations based on credible sources, we can enhance understanding of the risks associated with deepfakes. Additionally, reinforcing digital literacy training and sharing this knowledge is an opportunity to cultivate readers who are proactive, critical thinkers and who rigorously fact-check their sources. It is essential to foster a skeptical mindset and maintain constant awareness of this emerging phenomenon, as only through prompt and well-informed actions can we effectively monitor the accuracy and integrity of transmitted information. Let us remain watchful in our pursuit of truth and collectively combat the deceitful nature of deepfakes (Sandotra & Arora, 2024).

REFERENCES

- Akbasli, I. T., Birbilen, A. Z., & Teksam, O. (2024). Human-Like Named Entity Recognition with Large Language Models in Unstructured Text-based Electronic Healthcare Records: An Evaluation Study. *researchsquare.com*
- Brahms, Y. (2022). Philosophy of post-truth. *ethz.ch*
- Brooks, C. F. (). Popular discourse around deepfakes and the interdisciplinary challenge of fake video distribution. *Cyberpsychology*.
- Calleja, N., AbdAllah, A., Abad, N., Ahmed, N., Albarracin, D., Altieri, E., ... & Purnat, T. D. (2021). A public health research agenda for managing infodemics: methods and results of the first WHO infodemiology conference. *JMIR infodemiology*, 1(1), e30979. *jmir.org*
- Cassam, Q. (2021). Bullshit, post-truth, and propaganda. *Political epistemology*.
- Frank, L. (2022). Real-time Video Content for Virtual Production & Live Entertainment: A Learning Roadmap for an Evolving Practice.
- Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., ... & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media: a literature review and framework. *Annals of the International Communication Association*, 45(2), 154-174. *google.com*
- George, A. S., & George, A. H. (2023). Deepfakes: the evolution of hyper realistic media manipulation. *Partners Universal Innovative Research Publication*, 1(2), 58-74. *puirp.com*
- Gregory, S. (2022). Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism. *Journalism*.
- Huang, C. (2023). Popular Journalism in Contemporary China.
- Igwebuike, E. E. & Chimuanya, L. (2021). Legitimizing falsehood

- in social media: A discourse analysis of political fake news. *Discourse & Communication*. sagepub.com
- Jones, V. A. (2020). Artificial intelligence enabled deepfake technology: The emergence of a new threat
- Kondurkar, I., Raj, A., & Lakshmi, D. (2024). Modern applications with a focus on training chatgpt and gpt models: Exploring generative ai and nlp. In *Advanced applications of generative AI and natural language processing models* (pp. 186-227). IGI Global
- Meikle, G. (2022). Deepfakes.
- McCammon, M. N. (2021). The Emerging Threat of Synthetic Media: A Consideration of Journalists' Responsibilities. *The Handbook of Communication Rights, Law, and Ethics: Seeking Universality, Equality, Freedom and Dignity*, 292-302.
- Nnamdi, N., Oniyinde, O. A., & Abegunde, B. (2023). An Appraisal of the Implications of Deep Fakes: The Need for Urgent International Legislations. *American Journal of Leadership and Governance*, 8(1), 43-70. ajpojournals.org
- Ojala, M. (2021). Is the age of impartial journalism over? The neutrality principle and audience (dis) trust in mainstream news. *Journalism studies*. tandfonline.com
- Oso, L., Adeniran, R., & Arowolo, O. (2024). Journalism ethics: the dilemma, social and contextual constraints. *Cogent Social Sciences*. tandfonline.com
- Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., ... & Azam, S. (2024). A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*. iee.org
- Sandotra, N. & Arora, B. (2024). A comprehensive evaluation of feature-based AI techniques for deepfake detection. *Neural Computing and Applications*. researchgate.net
- Samuel-Okon, A. D., Akinola, O. I., Olaniyi, O. O., Olateju, O. O., & Ajayi, S. A. (2024). Assessing the Effectiveness of Network

- Security Tools in Mitigating the Impact of Deepfakes AI on Public Trust in Media. *Archives of Current Research International*, 24(6), 355-375. manu2sent.com
- Spector, B. (2020). Post-truth claims and the wishing away of brute facts. *Leadership*. sagepub.com
- Tian, S., Jin, Q., Yeganova, L., Lai, P. T., Zhu, Q., Chen, X., ... & Lu, Z. (2024). Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Briefings in Bioinformatics*, 25(1), bbad493. oup.com
- Thomson, T. J., Angus, D., Dootson, P., Hurcombe, E., & Smith, A. (2022). Visual mis/disinformation in journalism and public communications: Current verification practices, challenges, and future opportunities. *Journalism Practice*, 16(5), 938-962.
- Whyte, C. (2020). Deepfake news: AI-enabled disinformation as a multi-level public policy challenge. *Journal of cyber policy*.
- Wilkerson, L. (2021). Still waters run deep (fakes): the rising concerns of deepfake technology and its influence on democracy and the first amendment. *Mo. L. Rev.* Missouri.edu