Academic Journal of Social Sciences (AJSS)
Vol. 8, Issue 2 (July – December 2024) PP. 20 – 34
Article Accepted: *15.09.2024* ***Article Published:*** *31.12.2024.*
DOI: *https://doi.org/10.54692/ajss.2024.822205*

Deepfake Deception and the Need for Media Literacy in Pakistan

Huda Imran [1] * and Dr. Madiha Maqsood

Abstract

Social media sites have given rise to a new shift when it comes to making and sharing content. It has assisted in bringing many voices into the mainstream. However, this shift has also brought various challenges, especially with the invention of deepfakes. These are AI-generated videos that make it almost impossible for users to distinguish between fact and fiction. The method used in this research is content analysis of existing literature between 2018 to 2024, which provides information on media manipulation and public perception. The paper explores deepfakes through the lens of the agenda-setting theory that, in the past, studied the role of the media in influencing public opinion, and evaluates its applicability in today's world. Based on the findings of the study, deep fakes are a threat that influences public opinion and participates in the spread of false narratives that shape public agendas. It was only a matter of time before deep fakes became a means of spreading fake news in Pakistan as well. Given the fact that most of the population in the country is passive in its media consumption, this research addresses the need to improve the level of media literacy. This initiative will enable people to assess the legitimacy of what information they are consuming from the internet. The study recommends for improvement of literacy in media in a way that people in Pakistan should be able to critically evaluate the realities of media content.

Keywords: Deepfakes, Fake News, Misinformation, Agenda- Setting, Media Literacy

1 Huda Imran, (Mphil), Mphil Scholar, Department of Media and Development Communication, University of the Punjab, Lahore, Punjab, Pakistan. Corresponding Author: hudaimran529@gmail.com

2 Madiha Maqsood, (Ph.D.), Assistant Professor, Department of Media and Development Communication, University of the Punjab, Lahore, Punjab, Pakistan. Corresponding Author: hudaimran529@gmail.com

1. Introduction

The power of media today is not in the hands of a few large media outlets but in the hands of several **small** channels. The way news spreads nowadays is through the virality of the content. New media and digital media have decentralized the control that was in the hands of traditional media. People can now connect and discuss various issues online, making information sharing more democratic (Wu, Y., et al., 2021). With social media, anyone can contribute to the content, raising the issue of its credibility (Deshpande, R., & Ogale, S. 2024). Such being the case, it can be seen that social media is more effective than mass media; therefore, agenda-setting theory has found a new home. This theory describes how anyone in the digital age can decide which matters will be emphasized and in what manner (Kalpokas, I., & Kalpokiene, J. 2022). The Pakistani digital narrative today is perhaps the best illustration of McLuhan's famous mantra that "the medium is the message" primarily due to the "WhatsApp culture", in which people trust the disseminated information blindly, without actually verifying it from a legit source (Talpey, C., 2022; Ahmed M., et al, 2021)

Being one of the disciplines of digital innovation, artificial intelligence, or AI, has opened up new opportunities for creativity generation and information sharing. This has also resulted in the creation of deepfakes. The term deepfakes is a contraction of the two words 'deep', an abbreviated form of the term deep learning, and the word 'fakes'. The term deepfakes was created by a user named 'deepfakes' on Reddit at the end of 2017 (Cole, Samantha, 2018). Deepfakes involve replacing one person with another within an image, video, or any other existing visual, hence eroding truth and reliability. As a result, users are unaware of whether the content that is being posted on social media platforms is genuine or not, which poses a threat to society, if not handled properly.

Deepfake on social media platforms negates the situation because the local audience is usually passive and believes everything they see or hear without cross-checking it with credible sources. It can act as a tool for cyber-terrorists to disseminate fake information, sway people's stances, and harm the character of others (George, A. S. et al., 2023). Though artificial intelligence affects different areas, the problem of deepfakes has emerged as a challenge, especially, concerning its negative impact on the youth of Pakistan (Falak, M. A., 2022). As more than one-third of the Pakistani population with over 180 million Internet users are now seeking information or news through social networks, the nature of information provision in the country has drastically changed (World Bank, 2023).

Dealing with media literacy in the modern digital world is crucial to possessing or developing media literacy that would enable people to use and examine media critically. This also fosters a more enlightened and knowledgeable audience base that has no trouble distinguishing between an original piece and a fake replication (Ali, A.,

& Qazi, I. A. 2023). This research aims to analyze the agendas that deepfake set as well as review the aspects of public awareness, media literacy, and the potential impact of deepfakes via social networking sites in Pakistan. It aims to determine the current weaknesses of the media in the country to deepfake manipulation and to improve the ability to counter it using media literacy.

In conclusion, this paper aims to synthesize literature on the effects of deepfake and the effectiveness of international media literacy interventions in altering audiences' perceived ability to distinguish genuine content from fake. Considering the recent discovery of deepfake technology and its relationship with public perception, the study uses a broad theoretical lens to stage the review topic and introduce the idea of media literacy programs in Pakistan to transform the target audience into critical analyzers.

1.2. Research Objectives

1. To investigate how deepfakes affect public perception and opinion, and contribute to the dissemination of fake news and misinformation.
2. To analyze the importance of media literacy as a tool to help individuals in Pakistan critically assess and identify deepfake content.

1.3. Research Questions

RQ1: How do deepfakes influence public opinion and contribute to the spread of false narratives in Pakistan's social media landscape?

RQ2: Why is media literacy crucial in combating the influence of deepfakes in Pakistan?

2. Literature Review

The paper employs the thematic analysis of scholarly articles, blogs, and books published from 2018 to 2024. The themes consist of deepfakes and media manipulation, public perception and vulnerability, media literacy and its issues, implications for future research, and agenda-setting theory. This approach explains how deepfakes disrupt the traditional role of media and set agendas. The review focuses on deepfakes in Pakistan and the audience that usually does not verify the information on their own. This underlines the need to enhance the media literacy initiatives and the thinking capacity of the population of Pakistan.

2.1. Misinformation, disinformation, and fake news

Today, we are experiencing three major problems in the social media environment, that are misinformation, disinformation, and fake news. Misinformation is unintentionally spreading news to the people while disinformation is done intentionally to trick people. Fake news is the type that consists of the production of false information with the aim of misleading readers. These are getting out of hand due to no control over information disseminated on social media networks or its verification. For instance, during the COVID-19 pandemic, fake news in the form of conspiracy theories, circulated rapidly causing societal fear (Rocha, Y. M., et al., 2021).

According to Fraga-Lamas and Fernández-Caramés (2020), people post misinformation in the form of images, videos, texts, and even audio materials.

Fake news or information propagates faster than factual news or information. Thus, the public should not engage with fake news for an extended period as this is not healthy for them (Chesney & Citron, 2020). The process of repeated spread of such false stories, especially through social media platforms, is known as the "continued influence effect". This was especially apparent during the recent COVID-19 outbreak. However, the "illusory truth effect" explains that the more exposure the public gets to fake news, the more they believe in it (Sharma, Zhang & Liu, 2022).

2.2. Impact and Challenges of Deepfake Technology

Dan and his team in 2021 described a radical change in social and social media interactions due to the advancement of new media. Artificial intelligence is now very much involved in the dissemination of fake news through realistic videos and pictures (Westerlund, 2019). Normally, the public thinks anything displayed visually is true, so they accept the fake information they get from the Internet (Collins et al., 2020). This is why people should not share unconfirmed or false information with others. Fake news involves the falsehood and the deceit of the new generation of artificial intelligence in certain aspects that appear real. This can mislead those who are unaware of the issue at hand (Chan et al., 2019). The spread of fake news tends to undermine the public's credibility, therefore it is not advisable to believe anything on social media platforms. Also, it can degrade the credibility of people and companies. These concepts can also shift people's minds and can be dangerous to democracy (Chesney, B., & Citron, D. 2020).

Deepfakes are of several types, mainly involving video, audio, and images forged using artificial intelligence. Face swap and voice synthesis are two of the most common types of deepfake that are widely used today. Face swaps are a bit more sophisticated than morphs where the features of one person's face are superimposed on another's body resulting in a very realistic albeit fake picture. This one works slowly and deliberately to enhance some qualities such as faces, expressions, and lighting to give an appearance of credibility (Gupta, 2019).

Voice synthesis on the other hand manipulates audio so that it appears that the subject said things that were never stated. This is achieved by analyzing the speech information of the target such as the dialect, pitch, and even accent to create new recordings with a similar voice (Bird, J. J., & Lotfi, A., 2023). The essence of these deepfakes are Generative Adversarial Networks (GANs), two neural networks, the generator and discriminator, used to generate and train the fake content to emulate real content thus making them even harder to distinguish from real media (AnalytixLabs, 2024).

2.3. Deepfakes threat to identities

Deepfakes can be employed to steal someone's identity and disseminate propaganda. Ukrainian YouTuber Olga Loiek found out that her face and voice had been copied by AI. The party who used her avatar created her into a Russian woman selling goods on the Xiaohongshu social application in China (Yan, 2024). These fake profiles with ten-thousand-plus followers had Loiek speaking Chinese and endorsing China-Russia relations which she did not understand. Yet another fake Loiek stood for Chinese idols and culture, spreading stories that would fuel the perceived communication between China and Russia. This manipulation highlights how deep fakes act as persuasive propaganda in setting the agenda and it underlines the ability of deep fakes creators to decide which issues shall be presented in what manner, to achieve a specific effect on the public. Deepfakes therefore, have the potential to create and mediate social reality that raises concerns for the masses, as they can not distinguish fake from real content (Vaccari & Chadwick, 2020).

2.4. Media Literacy and Global Initiatives

Media literacy is the process by which a person is empowered to decode media messages in any form possible. It enables a person to avoid content that is manipulated and fake by employing critical and analytical parameters. Media literacy provides users with such specifics as the reliability of the information source, hints for deception, and the tools used in media creation (Dame Adjin-Tettey, T., 2022). It provides options to assess the irregularity of motions in visuals along with the sound that is heard in the scene. Therefore, Media Literacy enables people to possess these skills, which puts them in a position to critically analyze and examine the accuracy of the information received thus minimizing the effects of fake information, hence promoting the safe usage of media (Wei, L., et al, 2023).

Many programs help people identify deepfakes, learn about fake information, and analyze sources and facts (UNESCO, 2023). The current measures towards countering deepfakes involve technology, sensitization, and collaboration. The Deepfake Detection Challenge (DFDC) supported by Facebook, Microsoft, Amazon, and the Partnership on AI represents one of the large-scale tasks that promote the development of advanced approaches to identifying fake videos. Such tools can detect disparities that are inherently characteristic of deepfakes, with the help of machine learning, neural networks, and forensic examination (Groh, 2022).

This means that there is a need to embrace technological solutions in the detection of deepfakes. CNNs (Convolutional neural networks) and GANs (Generative Adversarial Networks) are the most advanced, focusing on the physical and biological signals that are beyond the capability of AI imitation. There are numerous projects dedicated to deepfake detection on websites such as GitHub and many of those are open-sourced, encouraging cooperation in refining algorithms. It is therefore important to have content creators such as; Siraj Ravel who helps make

knowledge readily available by teaching videos on AI and deep fake detection (McCosker, 2022).

Other examples include Media Forensics (MediFor), and Semantic Forensics (SemaFor) developed under the U.S. Department of Defense's Joint Artificial Intelligence Center (JAIC). The goals and objectives of MediFor include designing methods and algorithms for verifying the fact-checking of the displayed photos and videos, while the goals and objectives of SemaFor include the identification of semantically suspicious videos. To mention a few, measures like SystemID, Content Authenticity Initiative, and the AETHER Media Provenance (AMP) system are being used by YouTube, Adobe, and Microsoft help to authenticate media (Alimova, A., 2021).

The other defense mechanism is public awareness, which entails the knowledgeable enlightenment of the masses, especially the media so that they can differentiate between real content and deepfakes. This project was initiated by Google-Poynter jointly with Stanford University and the Local Media Association, called MediaWise and its goal is to empower youth to distinguish between reality and fake. Facebook in cooperation with Reuters, created educational media literacy regarding the issue of manipulated media, where the participant is offered real cases with brief information about the development of the necessary technologies. The Washington Post, CNN, and the University of Cambridge introduced immersive experiences like Bad News Game. They are additional resources that help in raising awareness and educating the public about deepfakes (Jaiman, 2020).

The "Detect Fakes" project by Matt Groh at Northwestern University may be a more effective way of bringing awareness and critical viewing of AI-generated fake news as it provides concrete tips and ways to analyze them. The Detect Fakes portal allows users to watch clips from the DFDC dataset in high quality with deepfake and genuine videos and attempt to recognize specific indicators of deception like flawed faces and mismatched shadows. From this hands-on experience, the users should be able to differentiate between genuine and fake media. (Groh, 2022).

The study by Sarah Shawky El Mokadem shows that increasing media literacy among the public can reduce the dangers posed by deepfake videos and fake news. This study also showed that after exposure to media literacy lectures, there was an increase in the identification of fake news and deepfake videos. Deepfake recognition training reduced perceived credibility and willingness to share these videos among participants. This implies that targeted media literacy intervention programs designed to teach people to detect fake news created by artificial intelligence, need to be prioritized to contain the spread and effects of fake news (El Mokadem, 2023).

Through assorted activities, such as global preparedness and the involvement within the Partnership on AI, WITNESS appropriately confronts the deepfake threat. Instead, they stress the importance of cheap, publicly available, and simple detection

technologies for journalists and other civil society actors. The use of these tools is also important when it comes to integrating them into different forms of social media to boost the verification features. Moreover, WITNESS also participates in media literacy campaigns raising awareness about detecting and interpreting deepfakes in media products; at the same time, WITNESS uses technological solutions and policy initiatives in conjunction with raising public awareness and education against synthesized media and fake news (Gregory, 2020).

In conclusion, it is imperative to find solutions to social inequalities, information prejudices, and manipulation techniques that facilitate the formation of fake news and misinformation. For such programs to operate in situations such as in Pakistan, they have to be adopted to cover different situations, cultures, media environments, and political settings different from those in developed countries (Khan et al.,2020).

2.5. Misinformation and Media Literacy in Pakistan

Shifting the focus towards Pakistan, the media has transitioned from a few traditional to multiple and dispersed digital platforms. Over 150 million Pakistanis over the age of 15 have access to the internet today. Over the past few years, the shift has been remarkably towards social media especially Facebook, Twitter (now known as X), Instagram, and WhatsApp or similar applications in the geographical areas that have often suffered from inadequate access to traditional media. (World Bank 2023, Khan 2022).

As the media environment becomes more liberal and open, the population of Pakistan is free to participate in information-sharing networks as producers of content, sharers, or merely commentators on the media. It also opened up the possibility of establishing a ground for disinformation to be disseminated (Abbas, M., Khan, M. A, 2023). This problem is most prominently witnessed in underdeveloped rural areas where society struggles to filter fake news from real news because they have basic to zero information about the new media technology. Around 57% of the Pakistani public today is informed enough and possesses minimum proficiency to comprehend media content (UNESCO, 2023).

As a solution to this problem, the previously effective conventional channels have arisen to explain the harms of deepfakes in Pakistan. According to the study conducted by Sunvy et al., (2024), the threat of this technology has been given a massive highlight with over 50% of the news published in Pakistani newspapers covering deepfakes. Dawn, Pakistan Observer, The News International, The Express Tribune, and The Nation are some of the important newspapers that were scrutinized. Personally, Dawn contributed 11 articles, out of which 7 discussed deepfake threats. Pakistan Observer had two pieces and both were on deepfake threats. The News International contained 14 articles, 7 of which were analyzed for deepfake threats. The Express Tribune had thirteen articles in the news, of which six were concerning

deepfake threats. Out of 51 articles, 9 were from the Nation, and 5 had deepfake threat-related information (Sunvy et al., 2024).

This overall emphasis on deepfake threats in the Pakistani media reveals the country's understanding of deepfakes and their risks, particularly regarding politics. This is an indication that the coverage taken was very extensive with an understanding of how deepfakes can be used to influence public opinion and cause social problems. Yet, there is a lower focus on preventive actions for these risks, which indicates that there is increased demand for creating proper strategies for raising public awareness about the identification and battling deepfakes. All these problems can be solved with the help of media literacy initiatives, and the final eradication of this problem can only be achieved in the country by adopting a multi-faceted approach to enhancing the literacy level of the population.

3. Theoretical Framework

Discussing the agenda-setting theory in the light of deepfakes in Pakistan's digital media is useful in identifying the changes that have taken place in the theory formulated in the context of decentralized information dissemination. Proposed first by McCombs and Shaw in 1972, the theory explains that media influences the public agenda on matters of concern. Earlier, only a couple of channels mediated and channeled public attention but now this power lies in the hands of a complex network of people through social media platforms (Kaplan, D. 2021; Wu et al., 2021). Thus, deepfakes, realistic manipulated videos, and images undermine the key assumptions of the traditional agenda-setting models due to the formation of false narratives that are difficult to differentiate from reality (Graca, N., & Gojakovic, A. L. 2021).

AI and digital creations have evolved to produce better work from AI algorithms, thus the need for an efficient media literacy system to detect and distinguish AI content is necessary. Deepfakes quickly went viral on social media and remain popular despite people not trusting them because of their high sensationalism (Vaccari & Chadwick, 2020). That is how they work with prejudices and phobias, working with beliefs and setting up an agenda using a storytelling approach ((Vasist, P. N., & Krishnan, S. 2022)). This skews the public perception of the actual events that happened, unlike the factual orientation of the 'agenda-setting" theory. Accordingly, an updated theory needs to be considered in this decentralized digital media environment and on the use of deepfakes for opinion shaping (George et al., 2023).

Algorithms and user-generated content in today's world are the primary issues that the modern agenda-setting theory should take into consideration since these aspects create personalized information environments and filter bubbles, strengthen people's prejudices, and draw their attention to misinformation. This may lead to the fragmentation of public agendas. To do this, the theory must also include the current social activity to media consumption and creation where the users and viewers are also involved. That is why the relationships between technology, media, and human

psychology must be understood properly. Thus, the modernized version of the agenda-setting theory should include the notion of critical media literacy and solutions for dealing with the manipulation of the content to provide a well-informed society (Pashentsev, E, 2021; Vaccari & Chadwick, 2020).

4. Methodology

This theoretical study utilizes a literature review method to understand the connection between deepfakes, media literacy, and public perception in the world to counter Pakistan's social media sphere. It uses a combination of academic materials such as articles, books, and other up-to-date sources of information using scholarly articles, focusing on the agenda-setting theory, the effects of deepfakes, and the part played by media literacy. Drawing from media studies, communication, and information technology as well as from the social sciences the study applies a multi-disciplinary perspective. Thematic analysis is used to extract and analyze the pattern and trends in the use of deepfakes in society, twisting the perception of the public, and the changing nature of media literacy.

The study also assesses McCombs and Shaw's (1972) agenda-setting theory with particular reference to the new world of digital media and deep fakes. The literature review also examines the missions and efforts of global and Pakistani media literacy that aims to determine how effective such campaigns are at enabling the audience to decipher between authentic and fake news. It offers a balanced view of the existing issues and possible responses to them.

5. Findings

The study primarily covers the area of deepfakes, media literacy, and agenda-setting theory with a special reference to Pakistan. It affirms that all the conventional theories of media including the agenda-setting theory are valid but they need to be updated for the new media environment. Deepfakes replace stories by providing fake ones, and thus change the role of media, empowering those who produce and distribute manipulations that affect attitudes and opinions to a great extent. Deepfakes especially have a high impact in Pakistan as many social media users believe everything they see on the internet. Such manipulated videos and images can alter the facts and cause societal issues, especially in the political and social realms which can lead to chaos in the society. The review also points out how media literacy becomes essential for people to differentiate between reliable and fabricated information.

The current global media literacy campaigns can be seen as offering a universal set of guidelines, however, their impact can be highly diverse, particularly when it comes to specifics of using media literacy in one or another country, such as Pakistan. For media literacy programs to be effective in the Pakistani context, cultural, language, and literacy factors need to be considered. Such programs should include simplified examples and local dialects in their utilization within the education system and the general community about deepfakes. Deepfakes can be easily identified with

technological help as technology is very much involved in the creation and surveillance of deepfakes.

One of the tools is Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), which help perceive the physical and biological inaccuracy in the content. The issue of deepfakes is quite active on repositories such as GitHub, where multiple repositories are created to improve Deepfake detection algorithms. Others and influencers on YouTube, like Siraj Ravel, play their part by educating people regarding the concept of AI and deepfake detection (McCosker, 2022). Media Forensics known as MediFor, and Semantic Forensics or SemaFor are a few examples of programs made under the guidance of the U. S. Department of Defense's Joint Artificial Intelligence Center whose main objective is to determine the authenticity of media content alongside identifying the logical contradictions (Alimova, 2021). Some of the top organizations involved in this process include YouTube, Adobe, and Microsoft through systems such as SystemID, Content Authenticity Initiative, and AETHER Media Provenance (AMP) system (Alimova, 2021). Education and media literacy of the people in the country continue to be an essential strategy. Programs such as MediaWise, a project developed by Google, Poynter, Stanford University, and the Local Media Association work on teaching youth how to differentiate the truth. A media literacy course should be created in the local language on manipulated mass media, developed by Facebook and Reuters, and on the Washington Post's, CNN's, and the University of Cambridge's Bad News Game and interactive approaches (Jaiman, 2020).

Specific examples of Deepfake pre-detection training include the "Detect Fakes" Project by Matt Groh from Northwestern University can be used in the local context as it provides practical experience in detecting deepfakes alongside training in sharpening one's logical thinking processes (Groh, 2022). Based on the studies of Sarah Shawky El Mokadem there is such great importance of a media literacy approach to such threats because targeted programs drastically reduce the credibility and the likelihood of sharing news that can be fake or deepfake videos (El Mokadem, 2023). The organization known as WITNESS works across the world to provide detection instruments for individuals so that they can be used in Pakistan for Social media accounts as part of verification tools. The tool incorporates technological measures, lobbyism, and awareness campaigns to deal with synthetic media and fake news (Gregory, 2020).

In conclusion, the findings highlight the demands for enhancing the levels of public critical thinking and media literacy against the background of deepfakes as a threat, as well as stressing the importance of the development of contextually relevant programs to address this threat in Pakistan.

6. Conclusion

This paper establishes that there is an urgent need to combat the threats that deepfakes present within the Pakistani digital media sector. Thus, the classical concepts, agenda setting, for instance, are still valid but require application under the conditions of the decentralized flow of information. Deepfakes change the role of media in disseminating true information by presenting fake messages as authentic and by doing that, they can heavily influence the opinion of the society.

Pakistan's society is prone to deepfake deception because people are mostly a passive audience engaged with different media platforms limiting their critical thinking. The twisting of facts via deepfakes leads to societal problems mostly during political and social debates. It is, therefore, necessary to improve the media literacy of the public in Pakistan. Incorporated media literacy programs, within schools and any other community-related programs, can enable individuals to distinguish between fake and real information leading to a more enlightened society.

Frameworks from organizations such as MediaWise, the Deepfake Detection Challenge, or those established by big players like YouTube, Adobe, or Microsoft can prove to be helpful. However, they have to localize their products which means that they have to take into account the cultural, linguistic, and educational settings of Pakistan. To extend its abilities to combat the use of fake news, Pakistan should integrate more sophisticated detection methods with its already comprehensive media literacy campaigns. This overall approach is needed to develop critical thinking skills in people that will enable them to move through the information society with open eyes and enable them to distinguish between real news and fake news.

7. Gaps and Future Recommendations

This study, however, has some limitations, which could be useful for improvement in future studies. It relies mainly on what is currently available in previous literature, which may be outdated on the latest innovation in the generation of deepfake and media literacy. Even though this study employs the theoretical framework of the agenda-setting theory, there isn't sufficient literature from Pakistan. This makes it hard to isolate Deepfakes' influence separately for the Pakistani audience and analyze its impacts. Moreover, the importance is given exclusively to the problem of media education; however, no assessment of the current local campaigns describing the new strategies adequate for Pakistan is given. This idea is also relevant to evaluating the psychological consequences of deepfakes for the change in opinion and general behavior more systematically. Last but not least, the research does not point out how governments can and are responding to the issue of deepfakes through policy and regulations.

Future research should also include a collection of quantitative obtained through research done in Pakistan to estimate and analyze the perception, reception, and impact of deep fakes on the local audience. This can also be done through questionnaires, interviews, and case studies to have an outside view of the findings

made. Thus, there is a need to develop and assess culturally appropriate and linguistically competent media literacy programs for Pakistan. They should be integrated into the systems of education and extension services within the community. In addition, the paper can broaden the understanding of how deepfake works at the psychological level to shape attitudes and behaviors among the public. It is also wise to consult experts in technology on how best to design the new generation of tools that can be used in the detection and elimination of deepfakes. Thus, a breakdown of an effective strategy to deal with the emerging threat will also depend on interacting with policymakers to develop policies that will work against the creation and dissemination of deepfakes. Given the overall issues that deepfakes bring to the media environment, a new agenda-setting theory must be taken into account that highlights the decentralized structure of social media, AI's contribution to content production, and the need for people to begin to learn how to verify and discern authentic information.

8. References

Abbas, Z., Khan, R., Khan, M. Z., & Imran, M. (2023). Cyber laws and media censorship in Pakistan: An investigation of governmental tactics to curtail freedom of expression and right to privacy. Journal of Cyber Policy. https://doi.org/10.1177/09732586231206913

Bird, J. J., & Lotfi, A. (2023). Real-time detection of AI-generated speech for deepfake voice conversion. ResearchGate. https://www.researchgate.net/publication/373364324_Real-time_Detection_of_AI-Generated_Speech_for_DeepFake_Voice_Conversion

Kaplan, D. (2021). Public intimacy in social media: The mass audience as a third party. Media, Culture & Society, 43(4), 595-612.

Chesney, R., & Citron, D. K. (2020). Deepfakes: A looming challenge for privacy, democracy, and national security. University of Pennsylvania Law Review, 168(4), 1165-1325.

Collard, A. M. (2024). 4 ways to future-proof against deepfakes in 2024 and beyond. World Economic Forum. Retrieved from World Economic Forum.

Deshpande, R., & Ogale, S. (2024). The role of social media in the spread of misinformation and fake news. ResearchGate. Retrieved from https://www.researchgate.net/publication/379872133_'The_Role_of_Social_Media_in_the_Spread_of_Misinformation_and_Fake_News'

El Mokadem, S. S. (2023). The Effect of Media Literacy on Misinformation and Deep Fake Video Detection. Arab Media & Society. Retrieved from https://www.arabmediasociety.com

Fraga-Lamas, P., & Fernández-Caramés, T. M. (2020). Fake news, disinformation, and deepfakes: Leveraging distributed ledger technologies and blockchain to combat digital deception and counterfeit reality. IT Professional, 22(2), 53–59. https://doi.org/10.1109/MITP.2020.2977589

George, A. S., & George, A. H. (2023). Deepfakes: The evolution of hyperrealistic media manipulation. Partners Universal Innovative Research Publication, 1(2), 58-74.

Gregory, S. (2020). Insights from WITNESS' global preparedness work and the Partnership on AI's SteerCo on Media Integrity work on the Deepfake Detection Challenge. WITNESS Blog. Retrieved from WITNESS Blog.

Groh, M. (2024). Detect DeepFakes: How to counteract misinformation created by AI. Retrieved from https://detectfakes.kellogg.northwestern.edu

Rocha, Y. M., de Moura, G. A., Desidério, G. A., de Oliveira, C. H., Lourenço, F. D., & de Figueiredo Nicolete, L. D. (2021). The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. Journal of Public Health, 1-10.

Graca, N., & Gojakovic, A. L. (2021). Artificial intelligence in the context of intellectual capital and intellectual capital in the context of artificial intelligence. Independent Scientific Researchers, Belgrade, Serbia. https://doi.org/10.34190/EAIR.21.022

Kalpokas, I., & Kalpokiene, J. (2022). Deepfakes: a realistic assessment of potentials, risks, and policy regulation. Springer Nature.

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. Public Opinion Quarterly, 36(2), 176-187.

McCosker, A. (2022). Making sense of deepfakes: Socializing AI and building data literacy on GitHub and YouTube. New Media & Society, 26(5). https://doi.org/10.1177/14614448221093943

Ahmed M., Qamar Riaz M., Qamar M., & Asghar R. (2021). FAKE NEWS SHARED ON WHATSAPP DURING COVID-19: AN ANALYSIS OF GROUPS AND STATUSES IN PAKISTAN. Медиаобразование, (1), 4-17.

Talpey, C. (2022). Stop trusting WhatsApp. Geek Culture. https://medium.com/geekculture/stop-trusting-whatsapp-9c7aadfb536d

Sharma, K., Zhang, Y., & Liu, Y. (2022). COVID-19 vaccine misinformation campaigns and social media narratives. Proceedings of the Sixteenth International AAAI Conference on Web and Social Media, 920-931. Retrieved from https://doi.org/10.1609/icwsm.v16i1.19346

Sunvy, A. S., Reza, R. B., & Imran, A. A. (2024). Media coverage of DeepFake disinformation: An analysis of three South-Asian countries. INFORMASI, 53(2), 295-308. https://doi.org/10.21831/informasi.v53i2.66479

Ali, A., & Qazi, I. A. (2023). Countering misinformation on social media through educational interventions: Evidence from a randomized experiment in Pakistan. Journal of Development Economics, 163, 103108. https://doi.org/10.1016/j.jdeveco.2023.103108 Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0304387823000639

World Bank. (2023). Individuals using the Internet (% of the population). World Bank Data. Retrieved from https://data.worldbank.org/indicator/IT.NET.USER.ZS

Wu, Y., Wu, Y., Guerrero, J. M., & Vasquez, J. C. (2021). Digitalization and decentralization driving transactive energy internet: Key technologies and infrastructures. International Journal of Electrical Power & Energy Systems, 126, 106593.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society. https://doi.org/10.1177/2056305120903408

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology in Society. Retrieved from https://www.sciencedirect.com/science/article/pii/S0160791X19300774

Vasist, P. N., & Krishnan, S. (2022). Deepfakes: An integrative review of the literature and an agenda for future research. Communications of the Association for Information Systems, 51(Article 14).

Cole, S. (2018, January 24). "We are truly fucked: Everyone is making AI-generated fake porn now". Vice. Archived from the original on 7 September 2019. Retrieved 4 May 2019.

Gupta, B. (2019, December 26). Deepfake: A deep learning approach in artificial content generation. Medium. https://medium.com/@gupta.brij/deepfake-a-deep-learning-approach-in-artificial-content-generation-a626ceebe48f

Jaiman, A. (2020, September 7). Media Literacy, an Effective Countermeasure for Deepfakes. Medium. Retrieved from https://medium.com/the-innovation/media-literacy-an-effective-countermeasure-for-deepfakes-c6844c290857

Alimova, A. (2021, July 12). Defending Against Deep Fakes Through Technological Detection, Media Literacy, and Laws and Regulations. The International Affairs Review. Retrieved from https://www.internationalaffairsreview.com

Pashentsev, E. (2021, November). The malicious use of artificial intelligence through agenda setting: Challenges to political stability. In Proceedings of the 3rd European Conference on the Impact of Artificial Intelligence and Robotics ECIAIR (pp. 138-144).

Falak, M. A. (2022, May 16). Youth, abuse, and deepfake. Express Tribune.

AnalytixLabs. (2024, January 4). Detecting Deepfakes: Exploring Advances in Deep Learning-Based Media Authentication. Medium. Retrieved from https://medium.com/analytixlabs/detecting-deepfakes-exploring-advances-in-deep-learning-based-media-authentication-15af8a5b6aef

Yan, A. (2024, February 29). Ukrainian YouTuber's face cloned with AI for propaganda in China. South China Morning Post. https://www.scmp.com/news/world/article/3212145/ukrainian-youtuber-says-her-face-and-voice-have-been-cloned-ai-present-her-russian-sell- products