ARTIFICIAL INTELLIGENCE AND THE PROBLEMS OF DIGITAL MANIPULATION

Xudoyorov Shaxriyor

Termez University of Economics and Service, Faculty of Economics and Information Technologies, Information Systems and Technologies Department, 3rd year student

Scientific supervisor: Ernazarov Mirzohid

E-mail: xudoyorovshaxriyor57@gmail.com

Annotation: This article provides a comprehensive analysis of digital manipulation problems arising from the rapid development of Artificial Intelligence (AI) technologies in today's global information space. The research highlights the mechanisms by which AI systems influence the human mind, social opinion, and political processes — in particular, "Deepfake" technologies, "filter bubbles" formed through algorithms, and methods of cognitive manipulation. The main purpose of the article is to reveal the technical and socio-psychological aspects of digital manipulation, as well as to propose legal and technological solutions for combating these threats. The article discusses the transparency of AI algorithms, data security, and the issue of increasing digital literacy on the basis of scientific evidence. As a result of the analysis, the necessity of creating automated systems for detecting manipulation and developing international normative standards is justified.

Keywords: Artificial Intelligence, digital manipulation, Deepfake, algorithms, cognitive security, social engineering, data analytics, cybersecurity, neural networks, information warfare, filter bubbles, ethical standards, Big Data, machine learning.

Introduction

The current stage of human development is taking place in the context of the Fourth Industrial Revolution (Industry 4.0), and the main driver of this era is artificial intelligence (AI). AI has emerged as a technology with the ability not only to increase economic efficiency but also to fundamentally ease human life. However, like any powerful technological tool, AI also has its negative aspects, particularly the potential for digital manipulation. Today, the capabilities of processing data, creating personalized content, and predicting human behavior using AI have reached unprecedented levels. This in turn has given rise to a new type of threat in the digital space — threats that cast doubt on the truthfulness of information, divide society, and violate personal privacy.

Digital manipulation is the process of covertly influencing users' decision-making, controlling their feelings, and shaping their worldview in the interest of specific goals. Whereas previously manipulation was carried out through traditional media, today AI algorithms develop manipulative strategies tailored to each individual on the basis of their psychological profile. The relevance of this problem is that distinguishing fake content (text, audio, video) created by AI from the real thing is becoming nearly impossible.

At a time when digital transformation processes are rapidly underway in Uzbekistan as well, increasing the digital literacy of the population and developing a strategy to protect against possible manipulative influences caused by AI is considered a priority task. In this article, we analyze the problems of AI and digital manipulation not only from a technical point of view but also from social, political, and moral perspectives. The scope of the problem is so broad that it covers areas ranging from state security to the fundamental rights of the individual.

Discussion

To understand the link between artificial intelligence and digital manipulation, it is necessary first of all to focus on the working principle of algorithms. Modern social networks and search engines operate on the principle of the "attention economy." The main goal of these systems is to keep the user on the platform for as long as possible. To do this, AI algorithms analyze every step of the user — what interests them, which post they stop at most, what they "like," and even what topics they avoid. As a result, "echo chambers" are created for the user. In this process, a person begins to see only information that matches their views, which weakens their ability to think critically and makes them susceptible to manipulation.

In scientific circles, several main directions of manipulation through AI are being discussed.

First, there is Deepfake content created using generative neural networks. Deepfake is a technology that uses AI to transfer the appearance or voice of one person to another with extreme accuracy. This technology can serve to discredit political figures, spread false statements, and create social panic. For example, in 2022 the spread of a fake video purporting to show Ukrainian President Volodymyr Zelensky surrendering demonstrated how real the danger in this area is.

Second, the problem of micro-targeting. The 2016 US presidential election and the Cambridge Analytica scandal demonstrated how AI algorithms can be used to influence the political will of millions of people on the basis of their personal data. By dividing

people into psychological types, algorithms show each group special advertisements that match their fears or aspirations. This process is carried out so subtly that the person believes they made the decision independently, when in reality they are acting in the direction determined by the algorithm.

Third, the lack of ethical and legal regulation of AI. At present, the algorithms of technology giants (Google, Meta, Amazon) work on the principle of a "black box." That is, even the creators of the algorithm cannot always explain why the algorithm decided to show that particular information to the user. This lack of transparency opens a wide path for manipulation. Another issue at the center of discussion is bias in AI algorithms. If the dataset on which the AI was trained contains racial, gender, or religious discrimination, the algorithm absorbs these errors and begins to make manipulative decisions.

Main Part

Manipulation technologies based on artificial intelligence.

The basis of digital manipulation is automated systems for collecting and processing data. Today, Machine Learning and Natural Language Processing (NLP) models create texts indistinguishable from human-written ones in a matter of seconds. For example, large language models such as GPT-4 can manage thousands of bots on social networks, participate in discussions, and propagate a particular idea to the public.

Deepfake and synthetic media: Deepfake technology is based on Generative Adversarial Networks (GANs). Here, two neural networks compete with each other: one creates a fake image, and the other checks whether it is fake. This process continues until the image reaches a level of perfection. As a result, fake videos and audio appear at a level that the human eye cannot distinguish from the real thing. This is not only an invasion of privacy but also a major challenge for the forensic system.

Algorithmic governance and cognitive manipulation.

Social network algorithms lead the user to addiction by acting on their dopamine system. "Infinite scroll" and personalized recommendations exploit the weak points of the human brain. Cognitive manipulation is mainly carried out through the following methods:

Confirmation Bias: The algorithm gives the user only information that confirms their opinion, which leads to intolerance toward opposing views.

Information Overload: The user faces so much information that they cannot distinguish the important from the unimportant and become inclined to accept ready-made, simple (often manipulative) conclusions.

Cybersecurity and AI: new generation threats.

AI manipulation is not limited to information; it has also become an integral part of cyberattacks. Through "smart" phishing attacks, AI can study the user's correspondence style and send convincing messages on behalf of their relatives or colleagues. This leads to financial fraud and theft of confidential information. Likewise, automated hacking attacks with the help of AI find and exploit vulnerabilities in systems faster than human operators.

Social and political consequences.

Digital manipulation creates a crisis of trust in society. If any video or audio could be fake, people stop believing in the truth. This condition is called the "post-truth era." In the political arena, changing the opinion of voters through AI strikes a blow at the foundations of democratic processes. In "hybrid wars" between states, AI is becoming the main weapon and is widely used to sow discord among the population of the enemy state and to create distrust toward the government.

Solutions to the problem: a technological and legal approach.

A comprehensive approach is necessary to combat manipulation:

Establishing AI audits: independent international organizations that check how algorithms work must be created.

Digital watermarking: any content created by AI must automatically have a hidden mark indicating that it is an "AI product."

Media literacy: lessons on fact-checking and identifying manipulation methods should be taught to students in schools and universities.

Legal liability: strengthening criminal liability for harm caused through deepfakes and increasing the responsibility of platforms for the content they distribute.

Conclusion

Although artificial intelligence has opened the door to enormous opportunities for humanity, its transformation into a tool of digital manipulation is a serious cause for concern. This research shows that the problem of manipulation is not merely a technical malfunction but a systemic problem with deep socio-psychological and political roots.

The influence of AI algorithms on the human mind is so powerful that it can pose a direct threat to personal freedom and social stability.

We may not be able to eliminate digital manipulation entirely, but we have the ability to minimize its impact. To do this, it is necessary, first, to establish public and state control over the algorithms of technology giants; and second, to strictly adhere to ethical principles (AI Ethics) in the development of AI technologies. For developing countries like Uzbekistan, developing a national strategy in this regard and modernizing local cybersecurity systems on the basis of AI achievements is extremely important.

The final conclusion is that artificial intelligence must remain a tool that assists humans, not one that replaces them. Building a society protected from manipulation in the digital space is the common task not only of programmers but also of lawyers, psychologists, sociologists, and every conscious user. In the future, building the relationship between AI and humans on a foundation of trust and transparency is a guarantee of the preservation of human civilization.

References

1. Bostrom, N. (2014). Superintelligence: Paths, Dangers, Strategies. Oxford University Press.
2. Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.
3. Floridi, L. (2019). The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford University Press.
4. O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown.
5. Kadirli, M. (2021). "Artificial Intelligence and Information Security Issues." Uzbekistan Journal of Information Technologies 3(2), 45-58.
6. Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking.
7. European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
8. Harari, Y. N. (2018). 21 Lessons for the 21st Century. Spiegel & Grau.
9. DiResta, R. (2020). "The Supply Side of Disinformation." Stanford Internet Observatory.

10. Chesney, B., & Citron, D. (2019). "Deepfakes and the New Disinformation War." Foreign Affairs.
11. Nguyen, C. T. (2020). "Echo Chambers and Epistemic Bubbles." Episteme 17(2), 141-161.
12. Mirzayev, A. (2023). "The Role of Artificial Intelligence in the Digital Economy." Scientific Collection of Tashkent State University of Economics.
13. UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
14. IEEE. (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.
15. Lanier, J. (2018). Ten Arguments for Deleting Your Social Media Accounts Right Now. Henry Holt and Co.
16. Goodfellow, I., et al. (2014). "Generative Adversarial Nets." Advances in Neural Information Processing Systems.
17. Vosoughi, S., Roy, D., & Aral, S. (2018). "The spread of true and false news online." Science 359(6380), 1146-1151.
18. Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.
19. Sunstein, C. R. (2017). Republic: Divided Democracy in the Age of Social Media. Princeton University Press.
20. Schwab, K. (2016). The Fourth Industrial Revolution. World Economic Forum.