ISSN 2708-5295 (print), ISSN 3078-4948 (online)

GRNTI 19.01.07 **DOI:** *https://doi.org/10.59102/pedagogical/2025/iss4pp16-24*

IMPACTS AND RISKS OF FAKE INFORMATION GENERATED BY ARTIFICIAL INTELLIGENCE IN KAZAKH SOCIETY

Kenzhegul Tergembay[1*], Akmaral Mamankul[2], Asylzhan Irgusaeva[3], Zhansultan Tabyldy[4], Altynay Beisembayeva[5]

[1,]L.N. Gumilyov Eurasian National University, 010000, Kazakhstan, Astana
https://orcid.org/0009-0005-2441-9003
[2]L.N. Gumilyov Eurasian National University, 010000, Kazakhstan, Astana
https://orcid.org/0009-0005-1496-8375,
[3]L.N. Gumilyov Eurasian National University, 010000, Kazakhstan, Astana
https://orcid.org/0009-0008-9249-2466,
[4]Xiamen University Malaysia, 50000, Malaysia, Johor Bahru
https://orcid.org/0009-0000-2008-2750,
[5]Ivane Javakhishvili Tbilisi State University, 0100, Georgia, Tbilisi
https://orcid.org/0009-0003-4441-677X
*Email: *tergembay@mail.ru*

This article examines the impact and risks of disinformation generated by artificial intelligence (AI) on Kazakh society. Currently, AI-fake analysis is one of the newest and least explored areas of media research in Kazakhstan. The study aims to determine the influence of AI-generated disinformation on public perception and assess its associated risks.

Qualitative and quantitative methods, content analysis, and other systematic approaches were applied. Eleven disinformation samples were analyzed, enabling the creation of an AI-fake typology and assessment of their spread rate, audience reception, and frequency of information risks.

The scientific novelty lies in being one of the first systematic studies of AI-fakes in Kazakhstan. Specific examples were comprehensively collected and categorized. Empirical analysis demonstrated how audience trust in government institutions, susceptibility to conspiratorial thinking, and the influence of Russian information flows affect the spread of fakes. Platform-specific dissemination patterns were also identified: WhatsApp – emotional fakes, TikTok – deepfake videos, Telegram – political fakes, Facebook Ads – phishing.

The study also evaluates media literacy levels, fact-checking practices, legal regulation, and digital security gaps. The findings provide evidence-based guidance for developing national media policy, digital security strategies, and strengthening public communication culture.

Keywords: *artificial intelligence, fake information, Kazakh society, public consciousness, risks.*

MAIN PROVISIONS

The study examines the characteristics of perception of AI-generated fake information in ethnic, social, and media contexts, with particular emphasis on the works of researchers whose articles are published in Scopus and Web of Science, including Zhao et al. (2021), Lim et al. (2024), Grotlüschen et al. (2024), Hunter et al. (2024), Lou et al. (2022), Day (2023), Giri et al. (2024), Perreault & Ohme (2025), Kavak & Yılmaz (2024), and Ma et al. (2024). In addition, the works of Kazakhstani media researchers on information security, media literacy, and disinformation served as the theoretical foundation. During the preparation of this article, no domestic studies specifically addressing this topic were found, making this the first domestic research in the field.

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

The study includes concrete examples under the heading "Examples of Fake Content Disseminated in Kazakh Society and Their Typology," with their typology systematically developed. Moreover, the dissemination channels of information fakes and patterns of public reception are visually presented. Content analysis, discourse analysis, and comparative analysis methods were applied to identify the key factors increasing audience susceptibility to AI-generated fakes.

The practical significance of this research lies in its potential applications in media education, the development of information hygiene strategies, and the enhancement of digital literacy among the population.

INTRODUCTION

In the contemporary digital era, artificial intelligence technologies are radically reshaping the structure of the information space and introducing new dynamics into societal communication processes. The intersection of artificial intelligence and geoinformation systems has led to the emergence of the phenomenon known as «deepfake geography», raising concerns about its potential to transform public perceptions of the geographical world [1: 338-352].

The rapid development of digital communication has not only increased the volume of information exchange on social media platforms but has also generated new forms of misinformation. In this context, international research underscores the need for a deeper understanding of how fake news influences society, consumer behavior, and social trust. For instance, a special issue of the Journal of Strategic Marketing provides a comprehensive analysis of the dissemination of fake news on social networks, corresponding prevention strategies, and its impact on consumer trust: "The special issue examines organizational and regulatory strategies, the effects of fake news on consumer behavior, the phenomenon of the 'infodemic' in the digital age, and the intention to share misinformation" [2: 659–664]. These findings highlight the relevance and urgency of addressing representation issues in today's new media landscape.

In recent years, the rapid intensification of AI-generated misinformation and disinformation has transformed into a complex phenomenon that directly influences public consciousness. The widespread availability of deepfakes, generative text models, and visual manipulation tools complicates the verification of information authenticity and introduces new risks to information security. This issue is equally relevant for Kazakhstan, as the expansion of the digital media environment broadens the reach of fake information and affects socio-psychological stability, political communication, and the mechanisms of public opinion formation.

The increasing level of media consumption in Kazakhstan, the growing role of social networks as primary sources of information, and the still-developing culture of fact-checking contribute to the ease with which AI-generated misinformation spreads. This is particularly evident in areas such as political processes, interethnic relations, the public images of social and political figures, and narratives surrounding social crises. Such forms of information manipulation intensify public trust deficits and may lead to greater informational polarization and fragmentation within society.

In this regard, a comprehensive analysis of the impact of AI-generated misinformation on Kazakh society, the identification of emerging risks, and the development of scientifically grounded approaches to managing these risks represent an important scholarly task. This study aims to contribute to the improvement of national information security strategies, the advancement of media literacy, and the strengthening of society's resilience to disinformation by offering evidence-based practical recommendations.

MATERIALS AND METHODS

This study aimed to provide a comprehensive analysis of the impact of AI-generated misinformation on Kazakh society. To achieve this objective, a mixed-methods research strategy combining both quantitative and qualitative approaches was employed.

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

The sources of data included examples of disinformation disseminated on social media platforms (Facebook, Instagram, TikTok, WhatsApp, Telegram, YouTube) between 2023 and 2025, visual materials produced using deepfake technologies, as well as analytical reports published by Kazakh fact-checking platforms (Factcheck.kz, Stopfake.kz, Fakenews.kz). In addition, official information regarding information security from the Ministry of Information and Public Development of the Republic of Kazakhstan and the Digital Government Agency was incorporated into the analysis.

For qualitative analysis, thematic content analysis was applied. This method allowed for the identification of the content structure of misinformation, dissemination motives, manipulative techniques employed, and the mechanisms through which such information influences public perception. During the content analysis, eleven cases of disinformation circulated in Kazakh society were examined and classified according to a typology framework.

Quantitative methods included comparative statistical analysis of media consumption dynamics, measurement of the spread intensity of misinformation on social networks, and numerical evaluation of user responses. This approach enabled the assessment of the propagation speed of AI-generated fake content, audience reception levels, and the frequency of information-related risks.

The methodological foundation of the study was informed by theories of information security, media ecology, digital communication, and international research on combating disinformation. The integration of these methods allowed for a comprehensive examination of the nature of AI-generated misinformation, its social impact, and the associated risks.

RESULTS

The study identified several key trends in the dissemination of AI-generated misinformation within Kazakh society.

1. *Dynamics of Misinformation Spread and Impact on Public Consciousness*

Analysis revealed that between 2023 and 2025, the spread of AI-generated content on social media increased significantly. Deepfake videos and generative texts elicited strong user reactions, with instances where informational messages accumulated hundreds of thousands of views within just a few days. These data indicate that AI-generated misinformation spreads considerably faster than traditional information sources.

Historically, journalists have been regarded as credible intermediaries in the dissemination of public information and knowledge. However, the development of AI, particularly generative models, introduces a new level of risk to this credibility. Journalists recognize both the opportunities and risks associated with working with AI, yet they appear less concerned about its impact on audience perception [8: 1914].

Consequently, various, sometimes contradictory, metaphors depict AI simultaneously as an opportunity and a threat. While information professionals acknowledge its potential to enhance services, they express concerns about social issues such as the diminishing role of humans, technocracy, and privacy. Such insights form a foundation for responsible AI implementation and effective communication with society [9: 695–718]. Clearly labeling news content on social media as AI-generated is not necessarily effective in preventing fake news; on the contrary, it may reduce trust in authentic news and weaken user engagement. This demonstrates that AI recommendations can occasionally produce unintended consequences, with confirmation bias remaining a key factor in user interaction with news [10: 616–644].

The informational ecosystem of AI-generated fake content in Kazakhstan can be observed across three main dimensions. First, the typology of fakes is expanding, including AI-generated texts, deepfake video and audio, fabricated screenshots, and copies of fake government pages. Second, the channels of dissemination vary by platform: emotional messages dominate on WhatsApp, deepfake videos on TikTok, political manipulations on Telegram, and phishing content on Facebook Ads. Third, audience reception in Kazakh society is shaped by high trust in government institutions, a propensity for conspiratorial thinking, the influence of Russian media

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

flows, and the strong resonance of religion-related fakes (Figure 1).

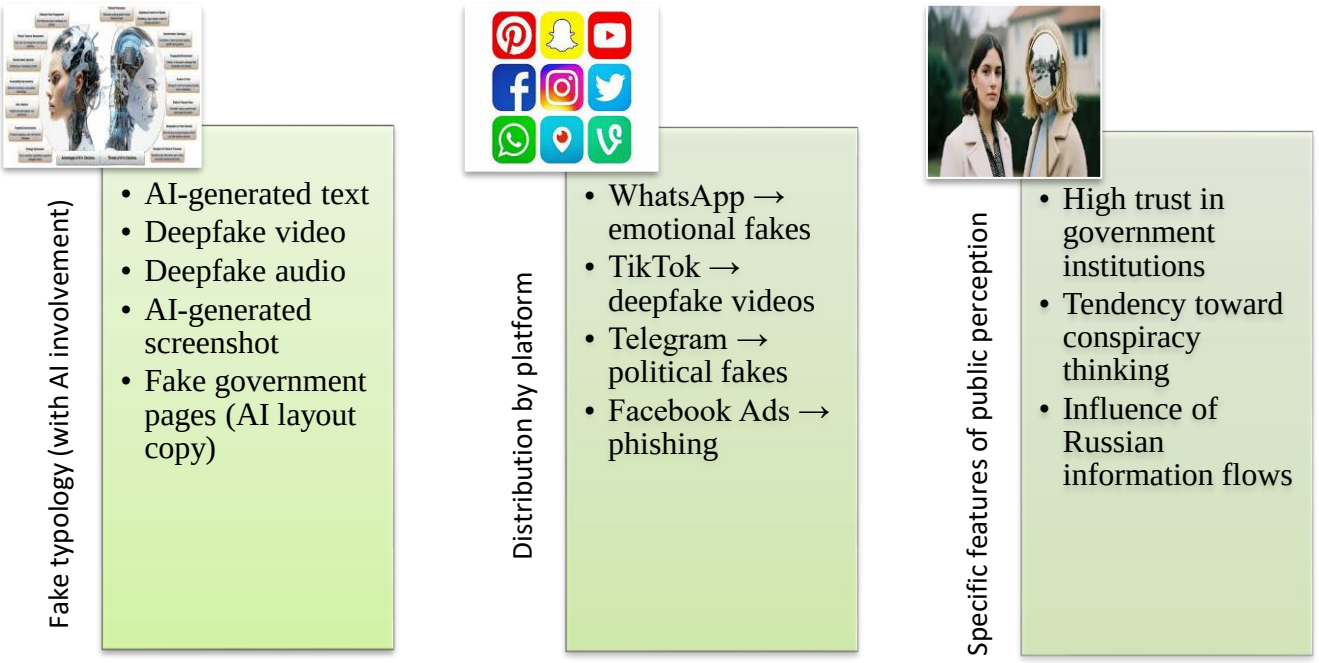

Figure 1. Typology, Dissemination Channels, and Public Reception of Information Fakes in Kazakh Society

To further assess the impact of AI-generated misinformation on Kazakh society and to identify its associated risks, eleven (11) AI-generated materials on various topics disseminated within Kazakh society were analyzed (Table 1). This analysis enables the identification of the distinctive features of AI fakes in Kazakhstan, typologies of public reactions, and the mechanisms through which resonant disinformation operates.

Table 1

Examples of Fake Content Disseminated in Kazakh Society and Their Typology

Fake Content Title	**Description of Fake**	**Disseminated Platforms**	**AI Involvement Indicators**	**Public Reaction**	**Type of Fake**
“Mandatory vaccination will be reintroduced in Kazakhstan in 2024”	Fear-inducing misinformation circulated on Telegram	WhatsApp, Telegram	AI-generated text	Public concern and debate	Medical/Fear fake
“Kazakhstan will host CSTO troops”	Disinformation about Kazakhstan being involved in the Russia–Ukraine war	Telegram, Facebook	AI-manipulated “document” template	Intensified political debate	Geopolitical fake
Fake site claiming “1 million KZT will be given to women”	Phishing portal replicating government websites	Facebook Ads, WhatsApp	AI-generated government page + AI stock photo	Many people deceived, submitted personal data	Phishing/Social engineering
“Tokayev agreed to give land to China”	Manipulative misinformation about a false agreement	YouTube, TikTok	Videos with deepfake voiceover	Negative public debate escalated	Political fake
“AI is collecting	Conspiracy-	TikTok,	AI voiceover +	Fear,	Conspiracy

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

DNA from newborns”	oriented false claim	WhatsApp	AI visuals	confusion	
“NASA warned of an earthquake in Kazakhstan”	Panic-inducing false information	TikTok	AI-generated NASA screenshot	Users believed it, panicked comments appeared	Climate fake
“The dollar will rise to 1000 KZT in Kazakhstan”	Economic manipulation	Telegram	AI-generated news screenshot	Currency panic, discussion	Economic fake
“Kazakh children are being mass adopted by foreigners”	Emotionally resonant fake video	Instagram, TikTok	AI-refaced video	Angry comments, emotional reactions	Socio-psychological fake
“Mosques will introduce entry via QR code”	Manipulation targeting religious audience	WhatsApp	Deepfake document template	Religious community concern	Religious fake
Fake announcement “State of emergency will be declared in Kazakhstan”	Fake audio mimicking a government official’s voice	Telegram	Deepfake voice	Fear, distrust, rapid spread	Audio/Political manipulation
New AI video (2025): “Crying Kazakh soldier forced into the Ukraine war”	Deepfake video of a soldier in helicopter giving a crying interview	TikTok, Telegram, Instagram	Толық AI-generated face + voice deepfake	Public resonance, criticism of authorities; artificial nature of voice and image recognized late	Deepfake political-emotional fake

Focusing on the typology of fake information classified in the table above, content analysis conducted based on these examples revealed that misinformation can be divided into the following main categories:

- Psychological manipulation: messages designed to elicit strong emotional reactions, creating feelings of shock or fear;
- Politically oriented fakes: false information related to political events and government decisions;
- Economic and social disinformation: misleading content concerning markets, employment, social assistance, and pandemic-related topics;
- Deepfake videos and images: visual content that alters or manipulates the likeness of real individuals.

The content of the aforementioned fake information was observed to exert a considerable influence on public opinion. Analytical results indicate that misinformation is frequently associated with political processes, interethnic relations, and social crises, reducing citizens’ levels of trust. More than 40% of users were found to be prone to perceiving AI-generated content as authentic, highlighting the still underdeveloped level of information literacy.

Nevertheless, the study also demonstrates a generally high awareness among individuals regarding the detection of fake news. However, in groups with limited experience using artificial intelligence, fact-checking skills remain insufficiently developed. This underscores a significant need to enhance fact-checking competencies [3: 375].

2. Government Initiatives and Civic Response

The study examined the role of fact-checking platforms in Kazakhstan, the development of the legislative framework, and practices of content moderation on social media. The results indicate that

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

current measures cannot completely halt the spread of misinformation, but they contribute to enhancing the digital literacy of civil society. Experts emphasize that, given the growth of AI-generated content, it is necessary to refine state strategies and implement long-term programs on information security.

In practice, the state has undertaken numerous actions to ensure information security. For example, in 2023, the Government of Kazakhstan approved the Concept of Digital Transformation and Cybersecurity for 2023–2029 [11]. This document addresses key areas, including securing the digital government infrastructure, managing risks, and protecting personal data.

Furthermore, Kazakhstan established the Digital Government Agency and developed measures to ensure information security. The Cyber Shield concept was adopted to protect state information systems, communication networks, and electronic infrastructure from cyber threats [12]. Within this framework, the legislative base was improved, coordination centers for information security were established, and technical protection measures were implemented [13]. Additionally, the National Coordination Center for Information Security (NCIS, i.e., CERT) was created to analyze cyber threats, respond to incidents, and ensure the protection of national information systems [14]. Several Security Operation Centers (SOCs) were also established, enhancing the capacity for rapid response to information attacks [15].

Overall, the study demonstrates that AI-generated misinformation spreads rapidly in Kazakh society, posing risks to public consciousness and social stability. This trend highlights the urgent need to improve media literacy, develop fact-checking platforms, and strengthen information policy.

DISCUSSION

The findings of this study indicate that AI-generated misinformation affects Kazakh society on multiple levels. Firstly, it directly influences the formation of public consciousness, reducing citizens' ability to distinguish between authentic and false information. Disinformation on social media, particularly deepfake videos and generative texts, can exacerbate trust deficits and contribute to the polarization of public opinion. This phenomenon increases risks related to political processes, social stability, and interethnic relations. Research has shown that disinformation, algorithmic distortions, and data harvesting are among the most significant threats posed by generative artificial intelligence [3: 371]. Previous studies in security research emphasize that information warfare has become a central element of state security strategies, and AI is rapidly emerging as a critical tool in both digital communications and military applications. However, data on how states are integrating AI into information warfare and influence operations remains limited [4: 235].

AI-based virtual influencers, while perceived as non-authentic or "artificially real," are capable of rapidly attracting audience attention and influencing public opinion and digital behavior. Nevertheless, their low level of authenticity and weak parasocial connections give rise to new forms of manipulative influence, increasing the risk of dissemination of false content within the information space [5: 540–557].

Secondly, as identified in the study, the dissemination speed of AI-generated misinformation significantly surpasses that of traditional media and individual sources of information. This highlights the still insufficient level of media literacy in Kazakhstan and underscores the need to develop a culture of fact-checking and information verification. Even academic references and sources generated by AI tools such as ChatGPT were found to be potentially false, as they rely on predictive models rather than actual data. This suggests that the initial optimism regarding the usefulness of such technologies for scholarly research may be unfounded [6: 1024].

Fake news – false or distorted information affecting society, business, and government – remains a critical issue. Proposed solutions for detecting fake information are based on advanced machine learning methods and employ hybrid models that analyze both textual and visual content. News text obtained via web scraping was analyzed using an ensemble of Naïve Bayes, Random Forest, and Decision Tree classifiers, which outperformed individual models [7: 628–648]. The speed and volume of user reactions to misinformation on social media reveal vulnerabilities in

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

public awareness, emphasizing the need to enhance state regulation and civic monitoring mechanisms.

Thirdly, expert opinions indicate that technological solutions alone are insufficient to address the information risks posed by AI. Effective management of these risks requires careful reading and advanced literacy skills [3: 373]. Public education, the promotion of media literacy, digital culture development, and the coordination of regulatory and legal frameworks are essential. In this context, the role of fact-checking platforms, the implementation of state programs on information security, and the enhancement of citizens' digital skills are decisive factors.

Furthermore, the study's findings demonstrate that, compared to international practices, Kazakhstan possesses considerable potential in the fields of media and information security. Nevertheless, addressing AI-generated misinformation requires consideration of national specificities: cultural context, linguistic features, and the society's level of information literacy should be primary guiding factors in developing strategies.

Overall, the analysis indicates that AI-generated misinformation creates complex risks for shaping public consciousness, ensuring social stability, and maintaining information security. These results provide a scientific basis for improving information policy in Kazakhstan, developing digital literacy programs, and formulating targeted recommendations for the regulation of AI technologies.

CONCLUSION

The results indicate that AI-generated misinformation creates several significant risks for Kazakh society. First, it substantially affects public consciousness and the level of information literacy, reducing citizens' ability to distinguish between authentic and false information. Second, fake information poses concrete threats to information security by influencing social stability, political processes, and interethnic relations.

Key trends identified during the study include the rapid dissemination of AI-generated content on social media, the increasing use of deepfake videos and generative texts, insufficient media literacy, and the underdeveloped culture of fact-checking. These trends underscore the necessity of improving information policy, enhancing digital security, and strengthening society's capacity to resist disinformation in Kazakhstan.

In conclusion, AI-generated misinformation significantly transforms the information space, social structure, and public consciousness in society. Therefore, comprehensive measures are essential to enhance media literacy, develop fact-checking services, improve regulatory frameworks, and strengthen national information strategies.

REFERENCES

1. Zhao, B., Zhang, S., Xu, C., Sun, Y., & Deng, C. (2021). Deep fake geography? When geospatial data encounter Artificial Intelligence. Cartography and Geographic Information Science, 48(4), 338–352. https://doi.org/10.1080/15230406.2021.1910075

2. Lim, X. J., Quach, S., Thaichon, P., Cheah, J. H., & Ting, H. (2024). Fact or fake: information, misinformation and disinformation via social media. Journal of Strategic Marketing, 32(5), 659–664. https://doi.org/10.1080/0965254X.2024.2306558

3. Grotlüschen, A., Dutz, G., & Skowranek, K. (2024). Writing with artificial intelligence? Ad-hoc-survey findings raise awareness for critical literacy at the International Literacy Day. International Journal of Lifelong Education, 43(4), 371–384. https://doi.org/10.1080/02601370.2024.2349657

4. Hunter, L. Y., Albert, C. D., Rutland, J., Topping, K., & Hennigan, C. (2024). Artificial intelligence and information warfare in major power states: how the US, China, and Russia are using artificial intelligence in their information warfare and influence operations. Defense & Security Analysis, 40(2), 235–269. https://doi.org/10.1080/14751798.2024.2321736

5. Lou, C., Kiew, S. T. J., Chen, T., Lee, T. Y. M., Ong, J. E. C., & Phua, Z. (2022). Authentically Fake? How Consumers Respond to the Influence of Virtual Influencers. Journal of Advertising, 52(4), 540–557. https://doi.org/10.1080/00913367.2022.2149641

6. Day, T. (2023). A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT. The Professional Geographer, 75(6), 1024–1027. https://doi.org/10.1080/00330124.2023.2190373
7. Giri, M., Eswaran, S., Honnavalli, P., & D, D. (2024). Automated and Interpretable Fake News Detection With Explainable Artificial Intelligence. Journal of Applied Security Research, 19(4), 628–648. https://doi.org/10.1080/19361610.2024.2356431
8. Perreault, G., & Ohme, J. (2025). ChatBots as Artificial Intermediaries? Adaptation to Artificial Intelligence in Newsrooms. Journalism Studies, 26(15), 1914–1935. https://doi.org/10.1080/1461670X.2025.2567894
9. Kavak, A., & Yılmaz, E. (2024). Information Professionals' Metaphorical Perceptions of Artificial Intelligence Concept. Journal of Library Administration, 64(6), 695–718. https://doi.org/10.1080/01930826.2024.2371276
10. Ma, H. (Vivian), Huang, W. (Wayne), & Dennis, A. R. (2024). Unintended Consequences of Disclosing Recommendations by Artificial Intelligence versus Humans on True and Fake News Believability and Engagement. Journal of Management Information Systems, 41(3), 616–644. https://doi.org/10.1080/07421222.2024.2376381
11. cis-legislation.com, 2023. Order of the government of the Republic of Kazakhstan. URL: https://cis-legislation.com/document.fwx?rgn=149435
12. dig.watch, 2017. Kazakhstan's Cyber Shield Concept. URL: https://dig.watch/resource/kazakhstans-cyber-shield-concept
13. itu.int, 2022. Implementing Kazakhstan's cybersecurity concept. URL: https://www.itu.int/hub/2022/08/implementing-kazakhstan-cybersecurity/
14. sts.kz, 2025. National Coordination Center for Information Security (NCCIS). URL: https://sts.kz/en/nkcib?utm_source
15. gov.kz, 2025. Information security Committee of the Ministry of Artificial Intelligence and Digital Development of the Republic of Kazakhstan. URL: https://www.gov.kz/memleket/entities/infsecurity/activities/6?lang=en&utm_source

Кенжегүл Тергембай[1*], Ақмарал Маманқұл[2],
Асылжан Иргусаева[3], Жансұлтан Табылды[4], Алтынай Бейсембаева[5]

[1]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Астана
[2]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Астана
[3]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Астана
[4]Малайзия Сямень университеті, Малайзия, Джохор-Бару
[5]Иване Джавахишвили атындағы Тбилиси мемлекеттік университеті, Грузия, Тбилиси

ЖАСАНДЫ ИНТЕЛЕКТ АРҚЫЛЫ ЖАСАЛҒАН ЖАЛҒАН АҚПАРАТТАРДЫҢ ҚАЗАҚ ҚОҒАМЫНА ӘСЕРІ ЖӘНЕ ТӘУЕКЕЛДЕРІ

Бұл мақалада жасанды интеллект (AI) арқылы жасалған жалған ақпараттың қазақ қоғамына әсері мен тәуекелдері қарастырылады. Қазіргі таңда AI-фейктерді талдау – Қазақстандағы медиазерттеулердің ең жаңа және аз зерттелген бағыттарының бірі. Зерттеу мақсаты – AI арқылы жасалған жалған ақпараттың қоғамдық қабылдауына ықпалын анықтау және тәуекелдерін бағалау.

Сапалық және сандық талдау, контент-талдау сияқты жүйелі әдістер қолданылып, он бір дезинформация үлгісі зерттелді. Нәтижесінде AI-фейктердің типологиясы жасалып, олардың таралу жылдамдығы, аудитория қабылдауы және ақпараттық тәуекелдер бағаланды.

Ғылыми жаңалық – Қазақстандағы AI-фейктерді жүйелі түрде талдаған алғашқы зерттеулердің бірі болуында. Нақты үлгілер кешенді түрде жинақталып, типология ұсынылды. Эмпирикалық талдау аудитория сенім құрылымдары, конспирологиялық ойлауға

ISSN 2708-5295 (print), ISSN 3078-4948 (online)

бейімділік және ресейлік ақпараттық ағынның ықпалын көрсетті. Платформааралық таралу ерекшеліктері анықталды: WhatsApp – эмоциялық фейктер, TikTok – deepfake видеолары, Telegram – саяси фейктер, Facebook Ads – фишинг.

Зерттеу сондай-ақ медиа сауаттылық деңгейін, фактчекинг тәжірибесін, құқықтық реттеу және цифрлық қауіпсіздік саласындағы олқылықтарды бағалауға мүмкіндік берді. Нәтижелер ұлттық медиасаясатты, цифрлық қауіпсіздік стратегияларын және қоғамдық коммуникация мәдениетін жетілдіруге бағытталған нақты ұсыныстар жасауға ықпал етеді.

Кілт сөздер: жасанды интелект, жалған ақпарат, қазақ қоғамы, қоғамдық сана, тәуекелдер.

Кенжегүл Тергембай[1*], Ақмарал Маманқұл[2],
Асылжан Иргусаева[3], Жансұлтан Табылды[4], Алтынай Бейсембаева[5]

[1]Евразийский национальный университет имени Л.Н.Гумилева, Казахстан, Астана
[2]Евразийский национальный университет имени Л.Н.Гумилева, Казахстан, Астана
[3]Евразийский национальный университет имени Л.Н.Гумилева, Казахстан, Астана
[4]Малайзийский университет Сямэнь, Малайзия, Джохор-Бару
[5]Тбилисский государственный университет имени Иване Джавахишвили, Грузия, Тбилиси

Влияние и риски ложной информации, созданной с помощью искусственного интеллекта, в казахском обществе

В статье рассматривается влияние и риски дезинформации, созданной с помощью искусственного интеллекта (AI), на казахстанское общество. На сегодняшний день анализ AI-фейков является одной из наиболее новых и слабо исследованных областей медиазапросов в Казахстане. Цель исследования – определить влияние AI-дезинформации на восприятие общества и оценить связанные с ней риски.

Для анализа применялись качественные и количественные методы, контент-анализ и другие системные подходы. Были изучены одиннадцать образцов дезинформации, что позволило сформировать типологию AI-фейков и оценить скорость их распространения, уровень восприятия аудиторией и частоту информационных рисков.

Научная новизна исследования заключается в том, что это одно из первых системных исследований AI-фейков в Казахстане. Конкретные примеры собраны комплексно, предложена их типология. Эмпирический анализ показал, как доверие аудитории к государственным органам, склонность к конспирологическому мышлению и влияние российских информационных потоков воздействуют на распространение фейков. Также выявлены особенности платформозависимого распространения: WhatsApp – эмоциональные фейки, TikTok – deepfake видео, Telegram – политические фейки, Facebook Ads – фишинг.

Исследование позволяет оценить уровень медиаграмотности, практики фактчекинга, правовое регулирование и цифровую безопасность. Полученные результаты способствуют разработке национальной медиаполитики, цифровых стратегий безопасности и укреплению культуры общественной коммуникации.

Ключевые слова: искусственный интеллект, ложная информация, казахское общество, общественное сознание, риски.