IRSTI 19.01.11 https://doi.org/10.26577/HJ.2024.v73.i3.4

M.S. Yessimova*, T.V. Shevyakova

Ablai Khan Kazakh University of International Relations
and World Languages, Kazakhstan, Almaty
*e-mail: yemanshuk@mail.ru

DEEP FAKES IN THE DIGITAL MEDIA AGE: OPPORTUNITIES AND THREATS

The article is devoted to an urgent problem today – the technology of deepfakes in the context of the information society. Thanks to modern technologies and digital tools, the rapid development of artificial intelligence (hereinafter – AI), the problem of deepfakes is becoming more widespread, as their influence on public opinion, political decisions and knowledge is only increasing. Deepfakes provide opportunities for spreading false information, exerting political influence, and even manipulating the image (using photos, videos and audio) of celebrities.

The aim of the article is to understand these technologies, identify the challenges and threats associated with them, and develop countermeasures.

The hypothesis of the study. In the era of digital media, deepfakes are powerful tool that can be used both to create positive innovations in various fields (for example, entertainment, education, marketing, etc.) and to spread disinformation and manipulation in the media space, which poses significant threats to public and information security.

The practical significance of the work lies in the fact, that research in the field of deepfakes can contribute to the development of tools and methods for detecting and analyzing this kind of content.

Research methods include general scientific, special scientific and specific scientific methods (content analysis, generalization, comparison, etc.).

The results of the study may be useful in further research of technologies deepfakes in the information space, as well as being useful in recognizing false AI-generated content.

Key words: deepfakes, AI, generated content, fake videos, digital environment.

М.С. Есімова*, Т.В. Шевякова

Абылай Хан атындағы қазақ халықаралық қатынастар
және әлем тілдері университеті, Қазақстан, Алматы қ.
*e-mail: yemanshuk@mail.ru

Цифрлық медиа дәуірдегі дипфейктер: мүмкіндіктер мен қауіптер

Бұл мақала бүгінгі күннің өзекті мәселесіне – ақпараттық қоғам контекстіндегі дипфейк технологиясына арналған. Заманауи технологиялар мен цифрлық құралдардың, жасанды интеллектің (бұдан әрі – АИ) қарқынды дамуының нәтижесінде терең фейктер проблемасы кең тарала бастады, өйткені олардың қоғамдық пікірге, саяси шешімдер мен білімге әсері тек күшейе түседі. Терең фейктер жалған ақпаратты таратуға, саяси ықпал етуге, тіпті атақты адамдардың имиджін (фото, видео және аудионы пайдалану) манипуляциялауға мүмкіндік береді.

Бұл мақаланың мақсаты – осы технологияларды түсіну, олармен байланысты сын-қатерлер мен қауіптерді анықтау және қарсы шараларды әзірлеу. Зерттеу гипотезасы: цифрлық медиа дәуірінде терең фейктер әртүрлі салаларда (мысалы, ойын-сауық, білім беру, маркетинг және т.б.) оң инновацияларды жасау үшін де, қоғамдық және ақпараттық қауіпсіздікке айтарлықтай қауіп төндіретін медиа кеңістікте жалған ақпарат пен манипуляцияларды тарату үшін де пайдалануға болатын қуатты құрал болып табылады.

Жұмыстың практикалық маңыздылығы – дипфейк зерттеулері осындай контентті анықтау және талдау құралдары мен әдістерін жасауға ықпал етуі мүмкін.

Зерттеу әдістері жалпы ғылыми, арнайы ғылыми және нақты ғылыми әдістерді (мазмұнды талдау, жалпылау, салыстыру және т.б.) қамтиды. Зерттеу нәтижелері технологияны одан әрі зерттеуде пайдалы болуы мүмкін ақпараттық кеңістіктегі дипфейктер, сондай-ақ жалған жасанды интеллектуалды мазмұнды тану кезінде пайдалы болуы керек.

Түйін сөздер: дипфейктер, АИ, жасалған мазмұн, жалған бейнелер, сандық орта.

М.С. Есимова*, Т.В. Шевякова
Казахский университет международных отношений и мировых языков
им. Абылай хана, Казахстан, г. Алматы
* e-mail: yemanshuk@mail.ru

**Дипфейки в эпоху цифровых медиа:
возможности и угрозы**

Данная статья посвящена актуальной на сегодняшний день проблеме – технологии дипфейков в контексте информационного общества. Благодаря современным технологиям и цифровым инструментам, стремительному развитию искусственного интеллекта (далее – ИИ) проблема дипфейков становится все более распространенной, поскольку их влияние на общественное мнение, политические решения и знания только усиливается. Дипфейки предоставляют возможности для распространения ложной информации, оказания политического влияния и даже манипулирования имиджем (использование фото, видео и аудио) знаменитостей.

Целью данной статьи является понимание этих технологий, выявление связанных с ними вызовов и угроз и разработка мер противодействия. **Гипотеза исследования: в** эпоху цифровых медиа дипфейки представляют собой мощный инструмент, который может использоваться как для создания положительных инноваций в различных областях (например, сферы развлечений, образования, маркетинга и т.д.), так и для распространения дезинформации и манипуляций в медиапространстве, что представляет значительные угрозы для общественной и информационной безопасности.

Практическая значимость работы заключается в том, что исследования в области дипфейков могут способствовать разработке инструментов и методов обнаружения и анализа такого рода контента.

Методы исследования включают общенаучные, специально-научные и конкретно-научные методы (контент-анализ, обобщение, сравнение и др.).

Результаты исследования могут быть полезны при дальнейших исследованиях технологий дипфейков в информационном пространстве, а также быть полезными при распознавании ложного, сгенерированного ИИ контента.

Ключевые слова: дипфейки, ИИ, сгенерированный контент, ложные видео, цифровая среда.

Introduction

Relevance of the research topic. Digital ethics refers to the moral principles and standards that define the ethical behavior and responsibilities of individuals, organizations, and governments in the digital environment. This includes considerations of fairness, transparency, confidentiality, and security of the impact of technology on individuals and society as a whole. In today's interconnected world, where digital technologies such as artificial intelligence, big data analytics, etc. Ubiquitous, digital ethics play a crucial role in how we interact with technology and with each other. The new digital reality of recent years, in most technological solutions of which artificial intelligence (AI) is the core, has significantly distorted the human perception of the difference between reality and fiction. One of the main "erasers" of such a facet is deepfake technology, which can be described as an automated technique, i.e. machine synthesis of audiovisual digital content (images, audio, video, and even text) in order to create modified and at the same time the most realistic content (Kapitanov A., 2022, https://russiancouncil.ru/analytics).

Deepfake technology can be used in some cases for entertainment or artistic purposes, as well as for malicious purposes. For example, fake videos of political leaders, celebrities, or other public figures can lead to the dissemination of misleading news or manipulation of people. That's why deepfake technology is also causing legal and ethical problems and attracting more and more attention. Today, international standards of journalism play an important role in the fight against disinformation and fake news, as they allow us to establish rules of ethics and professionalism that help journalists identify and prevent the spread of fake information. Deepfakes, that is, fake videos or images created using artificial intelligence technologies, are a serious threat to public trust in the media. Modern media should follow the principles of honesty, accuracy and reliability of information, which are enshrined in ethical norms and international standards of journalism. Compliance with these standards helps to strengthen trust in the media and prevent the spread of misinformation. In our opinion, journalists should be especially attentive to the content they publish and carefully check the sources of information in order to avoid misinformation and help society distinguish real news

from fake news. The problem of deepfakes is an increasingly serious threat to society in the context of modern digital media space. With the advent of deep learning and artificial intelligence technologies, the ability to create high-quality fake videos and audio materials has become more accessible, which can lead to the spread of misinformation, manipulation of public opinion, violation of privacy and trust. Today, fake news can be perceived not only from a negative, but also from a positive side. The topic of deepfakes in the era of digital media is extremely relevant and important in the modern information society, because with the rapid development of technologies for creating and distributing content, it has become possible to easily and quickly create realistic videos, photos and audio recordings that can be easily mistaken for reality. Thus, the study of deepfakes and the development of methods for detecting and combating them is becoming increasingly necessary in the context of the modern information space. We believe that the active coverage of this problem in scientific publications contributes to raising public awareness and contributes to the development of effective strategies for maintaining information security. The purpose of the study. The purpose of this study is to study the technology of creating deepfakes and their impact on the information environment.

Research objectives:

- to analyze the literature on this topic;
- to study the basic technologies for creating deepfakes;
- identify deepfake recognition methods;
- develop recommendations for recognizing deepfakes.

The object and subject of the study. The object of research is deepfakes, that is, fake media content created using technology, special programs and artificial intelligence. The subject of the study is the causes of the appearance of deepfakes, methods of their creation, as well as ways to protect against them in the modern media space.

Research methods. To achieve the aim of the study, a set of general scientific, special scientific and specific scientific methods were applied. Also, during the research, such methods as analysis, generalization, deductive, inductive, comparative, system-structural, etc. were used. The practical significance of the study.

Practical significance. This research has practical value for various stakeholders, including journalists, various media organizations and a wide audience, as it helps to raise awareness, improve ethical standards and develop new methods to combat deepfakes.

Literature review

Recently, technologies for creating fake photos and videos have become widespread, which, using a computer image synthesis technique based on artificial intelligence, transfer facial features from a human image to a target photo (video recording) with a high degree of verisimilitude. Despite the fact that these images (videos) are fictitious, they can be posted on the Internet as allegedly real. We are talking about technologies called deepfake, this term comes from English words deep learning (Rus. – глубинное обучение) and fake (Dobrobaba M.B., 2022).

Scientists all over the world are engaged in the study of the problem of deepfakes. Let's consider different approaches to the interpretation of this term.

«Fake news includes both misinformation and denial of veracity, with the latter two differing in their degree of veracity and intent. Disinformation reflects the dissemination of false information that is known to be false by the person disseminating the information, and thus reflects lies and malicious intent. Disinformation refers to the dissemination of false information that is perceived by the sender as true, thereby reflecting a lie and lack of malice» (E. Aimeur, S. Amri, G. Brassard, 2023; N. Nour, J. Gelfand, 2022).

Fake videos or deepfakes are audiovisual fakes deliberately created to suggest that someone did or said something that never happened (Chesney R. & Citron D. K.; Nelson A.& Lewis, J. A., 2019). It should be noted the exponential leap caused by these manipulation mechanisms and their ability to dramatically and shockingly distort reality. Added to this is the possibility of rapid and widespread distribution and the fact that they can be used by users who do not necessarily own a wide range of technologies. As a result, deep forgeries are becoming more realistic and resistant to detection. Deepfakes are synthetic videos that look very similar to real ones. (Vaccari, C., Chadwick, A., 2020). As for the history of the appearance of this phenomenon, it is known that deepfakes appeared a long time ago, so one of the most famous photographs of former US President A. Lincoln, published after his death in 1865, is a distant ancestor of the modern deepfake. In the picture, the head of the American president is superimposed on a photorealistic engraving of 1852

depicting the Vice President of the United States, J. Calhoun, which was made on the basis of a previously painted portrait in oil (R. Chawla, 2019).

In December 2023, the Council and the European Parliament reached an interim agreement on what will be the first Artificial Intelligence Law within the EU to guarantee safe and transparent AI. The ultimate goal is to promote the introduction of reliable human-centered AI and protect health, safety, fundamental rights and democracy from its harmful effects (European Parliament, 2023). Among the methods banned by AI for its intrusive and discriminatory use, MEPs named the non-selective extraction of facial images from the Internet or images from CCTV cameras to create facial recognition databases. Fines for non-compliance with the aforementioned law range from 35 million euros, or 7% of the total turnover, to 7.5 million fines. The use of AI in the dissemination of content through social networks and, in particular, in its application to audiovisual materials is of particular relevance; Although this new technology opens up opportunities in many fields such as artistic, audiovisual and graphic creativity, it also presents a challenge for citizens who need to create content (Francisco Javier & Ruiz del Olmo, 2014: 2). Researchers Shahid et al. note that most users lack the skills and desire to detect fake videos and are unaware of the risks and harms associated with this type of forgery. In particular, in countries such as India, where the named authors conduct their research, the fact is emphasized that even when users know that this is a fake video, they prefer not to take any action and sometimes voluntarily share videos that do not correspond to reality, but which contribute to their worldview (Shahid, F., Kamath, S., Sidotam, A., Jiang, V., Batino, A. & Vashistha, A., 2022). Recently, there has been a change in the trend in the use and consumption of online discourses, preferring or relying on audiovisual discourse. The popularity of video platforms such as TikTok, Instagram, or YouTube encourages video consumption at the expense of text. Thus, the growth of the video format is not only unstoppable, it has already taken its place in the Internet space – this is evidenced by the Digital News Report, which says how in 2023 There has been a gradual decrease in activity on traditional platforms such as Facebook, while it has increased on TikTok and other networks based almost entirely on the video format. Video news consumption is also growing in absolutely all markets (Neuman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K., 2023).

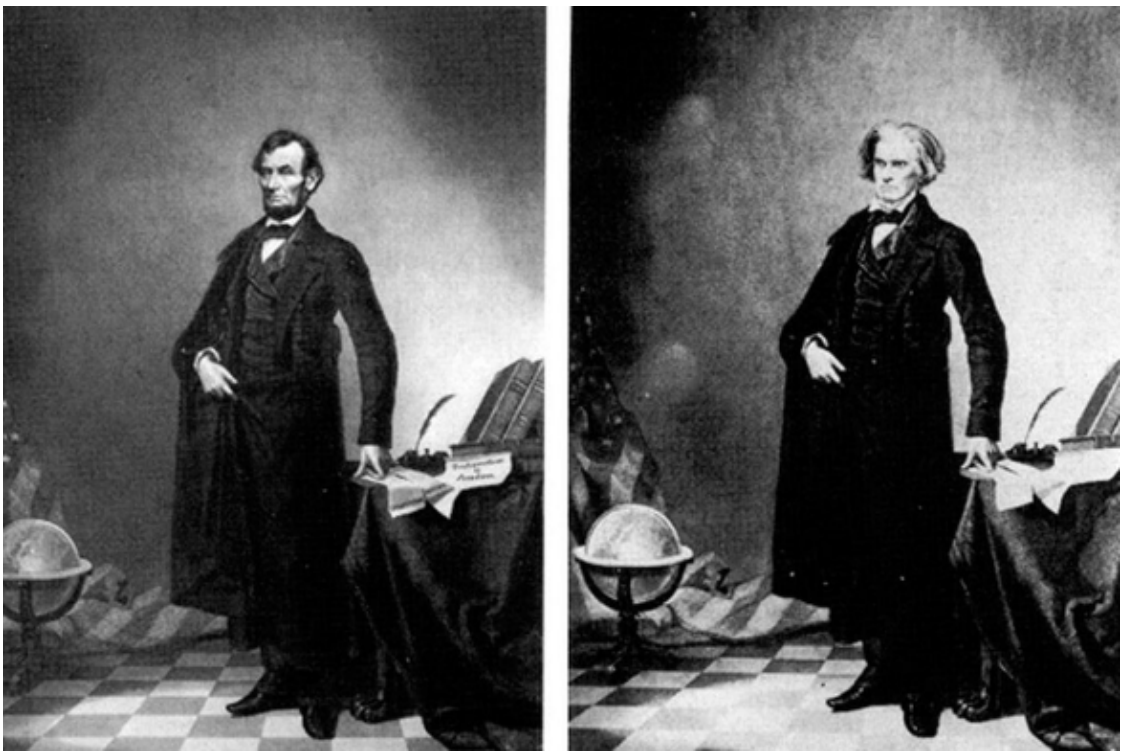

Figure 1 – An example of 19th century deepfake technologies

Discussion and results

A review of the literature allows us to narrow down the available interpretations of deepfakes, thus it can be summarized that deepfake is a form of artificial intelligence (AI) that generates video, images and, often, audio that can change or create images of people, events and objects in reality. AI can create or manipulate images, videos, and audio that look and sound realistic. This is no longer hypothetical, as more and more fake images and videos of famous actors, politicians and influential people are circulating on the Internet, not excluding Joe Biden, US President Tom Cruise, actors Keanu Reeves or deceased stars such as Freddie Mercury.

Table 1 – Types of deep fakes

Type of deep fake	Description	Examples of using deep fakes
Facial deep fakes	Replacing or changing a face in an image or video.	Fake videos with famous people, various jokes and memes.
Voice deep fakes	Generating or changing a voice to mimic another person's speech.	Fake phone calls, using voices in celebrities.
Text deep fakes	Creating a text that mimics the writing style of another person.	Fake news, fake correspondence.
Deep action fakes	Replacing or changing the movements of people in the video.	Creating unrealistic videos, for example, videos with accidents, catastrophes
Deep fake scenes	Generate new scenes based on real videos or images.	Movies, commercials.
Deep fake Images	Completely artificially created images that look like real ones.	Creating fake profiles on social networks, news.

Perhaps the first serious scandal with deep forgery broke out in 2020 around the falsified Extinction Rebellion Belgium video, in which Belgian Prime Minister Sophie Wilm Als spoke about the relationship between the coronavirus pandemic and climate change. The artificial intelligence used to create the video was able to manipulate a person's facial expression, facial expressions and voice. The fake video was taken seriously by many, it caused a wave of outrage and praise, emphasizing that the "post-truth era" is unregulated and temporary, in which the authenticity of media content, along with the role of a media gatekeeper, has come to naught, and there is nothing definite, even if we see it with our own eyes. It is extremely difficult to detect and verify explicit videos or fake photos, especially considering that they are mainly published on social media platforms, which are inherently limited in moderation and fact-checking capabilities. Deepfake technology is becoming more advanced every day and is available to an increasing number of people who do not require a high level of programming or video editing experience. In theory, this technology can be used for harmless purposes such as entertainment, education, research or the arts, but it can also pose a serious danger to the individual and society, such as spreading false information, influencing public opinion, violating people's privacy rights, damaging reputation, etc. For example, on some Russian websites, advertisements for weight loss pills often flashed using photos of singer Rosa Rymbaeva, producer Bayan Alaguzova. "Photos of the People's Artist of the Kazakh SSR, singer Roza Rymbaeva appeared on advertising banners on various Russian websites. In the pop-up photos, the singer allegedly advertises a weight loss product. According to the artist's son Ali Okapov, the singer herself is currently resting, but her family already knows about the appearance of such photos on the Internet. He noted that these are "fake" photos, and Rymbaeva does not actually advertise weight loss products." (https://tengrinews.kz/show/reklama-sredstva-pohudeniya-rozoy-ryimbaevoy-poyavilas-238670/).

Artificial intelligence will do a lot in the coming years, and today we should expect a very serious coexistence of man and machine. Despite all the advantages of AI development, today, unfortunately, neural networks have also become a tool for scammers. "Every third citizen of Kazakhstan sees on YouTube invitations to participate in foundations on behalf of the President of Kazakhstan Kassym-Jomart Tokayev, our deputies. The worst thing is that

many people believe this and, at least, follow the link. Scammers also use the "faces" of famous Kazakhstani bloggers, who allegedly share the secrets of their success and attach links to dubious sites. Another method of fraud using neural networks is a voice message from a relative asking them to send money to the specified account. It is very difficult to distinguish the truth from a fake here. By the way, scammers take the voice for such messages from spoken videos on social networks and voice messages in personal dialogues and shared chats. Creating a fake voice message costs a fraudster half a dollar. At the same time, neural networks can create conflict situations in social networks (https://mk-kz.kz/social/2024/02/14/).

Artificial intelligence today calls into question the credibility of everyone and everything, and will also affect what people believe to be true. How can you protect yourself? When we consider how artificial intelligence can affect authenticity or what is considered to be true, it is worth considering these methods separately, and we need not only to assess the current situation, but also to be prepared for expected situations in the coming months and years. Dangerous technologies are readily available and it is obvious that artificial intelligence with generative language models is practically capable of creating perfect text today, and it can also do this individually for each victim or company, automatically and continuously changing the text material used. Today, we interact with most people and even companies through emails, messaging programs, and chats. It's worth checking how many types of messaging or chat apps are installed on our devices. Therefore, it is important to understand that if attackers want to copy or use the authority of a person or company, they can do so even using a text message sent by artificial intelligence. Today, you can even communicate on behalf of a service provider or a dedicated manager with any number of people, using artificial intelligence, tirelessly, for 7-24 hours. When the first programs for creating deepfakes became publicly available in 2017, they were used mainly for entertainment – surely you have at least once tried to age yourself with a filter or checked how you would look like the characters of Harry Potter or Soviet comedies. Nicolas Cage is one of the most popular characters in comic dipfakes. So, the author of the YouTube channel "Battletoads Inc." used a face-swapping system to replace Cage's face with all the participants of the New Year's TV show (https://vg-times.ru/news/73934).

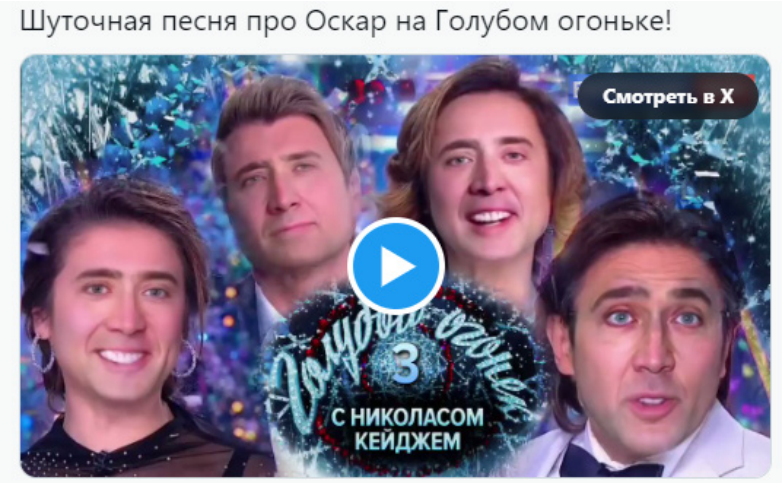

Figure 2 – Deepfake using the image of Nicolas Cage

But it quickly became clear that, like any other technology, deepfakes can be used for more destructive purposes. Let's look at programs for creating deepfakes that use machine learning and artificial intelligence technologies to create realistic fakes of video materials, images and audio recordings. These programs can vary in complexity of use and functionality.

Table 2 – The best neural networks for creating deepfakes

Application	Description
DeepFaceLab	Description: One of the most popular and powerful tools for creating deepfakes. It is used mainly for video. Features: Allows you to create and edit deepfakes with high precision. It has many settings to improve the image quality. Platform: windows.
FaceSwap	Description: An open source project that allows users to share faces on videos or images. Functions: Supports multithreaded processing and provides various algorithms to improve the result. Platform: Windows, macOS, Linux.
Zao	Description: A mobile application that has gone viral due to its ability to quickly and efficiently replace faces on video. Features: Easy to use and instant results. The user uploads his photo, and the app inserts his face into the selected video clips. Platform: iOS, Android.
Deep Art Effects	Description: An application that uses AI to style images and create deepfakes. Features: Applies various artistic styles to photos and videos, giving the impression that they were created by famous artists. Platform: Windows, macOS, Android, iOS.
Reface	Description: An application for mobile devices that allows you to share faces in photos and videos. Features: Quick creation of deepfakes, a variety of available templates, the ability to share the results on social networks. Platform: iOS, Android.
MyHeritage's Deep Nostalgia	Description: An application that allows you to animate old photos, creating the effect of animated portraits. Features: Animation of faces in photos, image quality improvement.

From the table, you can see that deepfake creation programs provide users with the ability to create amazingly realistic videos and images, despite the fact that their use raises important ethical and legal issues. Users should be aware of the need to use these technologies responsibly to avoid potential abuse and negative consequences. Next, let's look at examples of the use of deep fakes in the modern media environment.

Table 3 – Examples of using deepfakes in the modern media space

Example	Description	Using
Entertainment applications	Applications that allow users to replace faces with videos, such as FaceApp and Reface.	Personal entertainment, creating humorous videos.
Political campaigns	Deepfakes used to create fake videos featuring politicians.	Misinformation, manipulation of public opinion.
Media and film industry	Restoration of the appearance of actors who cannot participate in the filming.	Creating special effects, restoring old frames, replacing actors.
Pornographic deepfakes	Unauthorized videos with framed faces of celebrities or individuals.	Cyberbullying, blackmail, violation of privacy.
Educational projects	Creating historical reconstructions using deep fakes for educational purposes.	Historical reconstruction, educational materials.
Fraud and extortion	Using deepfakes to create false videos in order to extort money or information.	Fraud, social engineering (Social engineering is a method of obtaining information that involves the use of incompetence, unprofessionalism, negligence, fear or greed of a person. During the attack, the fraudster establishes contact with the carrier of the necessary information, disposes to himself, promises benefits or misleads, trying to obtain information that he can use for blackmail, theft of money or other assets. Social engineering involves both the use of a computer and a phone, mail correspondence, SMS, etc.)
Marketing and advertising	Using deepfakes in advertising campaigns to create unique content.	Increased audience engagement, creative advertising solutions.

1. An analysis of the Internet content shows that deepfakes are used in a variety of fields, for example: 1. The sphere of politics. Nowadays, cases of using deepfakes to discredit election candidates by creating fake videos where they make dubious statements have become more frequent.

2. The sphere of cinema. One of the first actors to be "resurrected" on the movie screen was the famous actor Bruce Lee. But computer graphics did not exist in those years, at least in the sense that there is in the world today. Then, in 1973, filmmakers used mirrors and photo cuts to "generate" the right shots and complement the story with dialogue. In the movie "Rogue One: Star Wars. Stories" (2016) used a deepfake to recreate the image of actress Carrie Fisher as the young Princess Leia. To create a virtually rejuvenated Leia, the Organs used CGI technology and Norwegian actress Ingvild Deila, combining her game with images of a young Carrie Fisher. The movie "Rogue One: Star Wars. Stories" was released in 2016 – the same year that 60-year-old Fisher passed away. A representative of the ILM studio responsible for digitalization claimed that the actress saw this version of herself and fully approved.

3. The porn industry. According to Bloomberg, during the boom of artificial intelligence, the number of pornographic deepfakes created without consent has increased dramatically – the number of videos has increased nine times since 2019. In May 2023 alone, about 150,000 videos appeared on 30 sites, which received a total of 3.8 billion views. Most of these sites offer libraries of fake programs with celebrity faces transplanted onto the bodies of porn actors (https://adpass.ru/porno-s-bolshim-intellektom).

4. The field of education. With the help of a deepfake, you can conduct online lectures for students. Students will be able to choose for themselves what their teacher will be – with the appearance of Cristiano Ronaldo or Einstein. The neural network allows you to create animations with historical characters for educational projects, such as a video with a live image of Albert Einstein explaining the theory of relativity.

5. Financial sector: Fraudsters have become more active in using deepfakes in the financial technology sector. Today, the neural network has learned to imitate the real voice of a person giving instructions to do something, and these are completely new risks. Most of all, banks fear that scammers will be able to use deepfakes to pass voice authentication, used to verify customers and provide them with access to their accounts. These examples demonstrate both positive and negative uses of deepfake technology in the modern media space. "Today, technology is increasingly being used not only to create fake news, but also in fraudulent schemes. In 2023, cases of deepfake fraud increased 31 times worldwide. With the help of deepfakes, attackers extort money from ordinary people and corporations, create pornography, and use technology to politically discreditation" (https://factcheck.kz/metodika-fch/kak-raspoznat-dipfeyk-video).

How to recognize a deepfake? Deep forgery can seem so plausible that it is quite easy to believe everything that is shown or told to us. Despite the impressive capabilities of deepfakes, with a little attention and common sense, you can avoid the pitfalls of this kind of technology. What people should pay attention to:

1. The sharpness of the voice. If there is a face on the video but no sound, there is the first reason for suspicion. The deepfake software uses audio recordings pre-recorded using a mobile phone or computer, which can be distorted with shades other than the original voice. It is important to be wary if the video lasts for several seconds.

2. Paying attention to non-verbal gestures. Fake news creators usually use templates, which are then overlaid with different faces based on the number of hours spent on each creative. If there are failures in the synchronization of sound with the movement of lips, eyes or head, or even inconsistencies in speech, it is more than likely that we are dealing with a deep forgery.

3. The presence of imbalances. Synthetic faces are almost indistinguishable from real ones, and this can cause a false sense of authenticity. However, there are certain natural movements that the algorithms have not yet been able to reproduce. If the main character's face is disproportionate to the size of his head and body, or his facial expressions do not match the poses he takes, perhaps we see a deepfake.

4. Paying attention to the mouth and eyes. It is possible to determine whether a person is a fake if he blinks involuntarily or unnaturally, or if he does it several times. Another aspect of reality that is difficult to reflect is the color of people's lips and the shape of their tongues and teeth when they talk. You should also pay attention to the winking of the person in the video. To determine if a video is a deep fake, you can focus on the number of blinks. The average person blinks almost 10 times a minute. However, in some deep fake videos, this number can sig-

nificantly exceed or fall below 10. Of course, many people may not notice this when watching a video, so by paying attention to a person's blinking reflex, you can determine whether the video is a deep fake, thereby verifying its authenticity.

5. The use of specialized tools. Today, technologies are developing so rapidly that for every malicious technology or software there is always a counteraction, regarding deepfakes, several applications and programs can be named designed to detect deepfakes, for example:

Deepware Scanner is a special application for detecting deepfakes. Reality Defender is an online tool for detecting deepfakes in images and videos.

Deepfake-o-meter is a service for analyzing video content for deepfakes.

The artificial intelligence industry, led by companies such as OpenAI, Google and Meta, is under increasing pressure as the global public demands responsibility for the content created by its products. Experts expect the industry to prevent users from creating misleading malicious materials and offer a way to track their origin and distribution. Today, calls to trace the origin of content using artificial intelligence are becoming increasingly desperate. In recent months, audio and video materials have already influenced political campaigns and voting in Slovakia, Taiwan and India. The new OpenAI forgery detector can help to cope with the problem, although, in our opinion, it will not solve it completely.

Conclusion

The rapid development of artificial intelligence and what it can and can do today is causing a lot of discussion. All these innovations often delight people, but they can go so far that people cannot always understand what is real and what is the product of artificial intelligence. We found out that Deepfakes are fake or forged audio recordings, videos, or visual effects, usually created using neural networks. For example, a person's video can be changed very realistically, making it look like he is saying something completely different from what he is saying in the video. In the same way, you can use various visual effects to make it look as if a person is doing some action that he did not do, or said something. All this can be done very realistically using multiple images or videos. In modern times, audio, photo and video content generation technologies have reached a very high level. At the same time, even a very well-prepared deepfake can (for now) be distinguished using specialized technologies. More recently, deepfakes were made for entertainment or to annoy someone, but now they are beginning to be used in the interests of various companies – television, cinematographic, etc. Actors no longer even need to personally voice the characters of films or cartoons – technologies allow synthesizing any words and phrases allegedly uttered by the actor himself (tonality, pitch of voice, etc. is synthesized very accurately). The relationship between deepfakes and international standards of journalism is an important aspect in the context of the spread of falsifications and disinformation in modern media. And although, in general, deepfakes pose more of a threat (ethically and legally) and they are more negative in nature, this technology should also have positive aspects. For example, deepfakes can be created to raise awareness of social problems existing in the world and to disseminate medical information on a large scale, since it can be adapted in any language, which will expand access to data.

In the field of medicine, research can be conducted that does not require real patients. In turn, the use of fake images will allow artificial intelligence programs to be trained to recognize more anomalies, which will eventually lead to greater accuracy. In education, this can be an important aid in learning, since teachers, for example, have the opportunity to use videos with historical discourse created in order to make lessons more attractive to students, but warning that its incorrect execution can damage knowledge if it is implemented. Using incorrect information. This tool will also have artistic purposes. For example, museums could hold video projections about various deceased artists, and thus one could hear the statements they made during their lifetime, as if they were speaking today. However, such use would entail discussion of broader issues such as copyright and the threat of plagiarism to existing works. Deepfake is a tool that is constantly evolving and is available to many people who, in certain cases, may use it for other purposes, so legal and ethical regulation of this technology is necessary. As part of the study, we found that deepfake, for example, can be used to use a fake identity or to steal money or information through access to databases. In general, we came to the conclusion that in most cases the use of this technology is inappropriate: 96% of deepfake videos found on the Internet contain unacceptable and unauthorized content. Anyway, deepfakes, as a form of information manipulation, contradict the basic principles of international journalism standards.

The creation and dissemination of disinformation using deepfakes can seriously undermine public

confidence in the media, as well as lead to negative social consequences. It is important for journalists and the media to adhere to the principles of international standards of journalism when working with information, especially in the context of deepfakes. This includes the obligation to verify the sources of information, the importance of the reliability and objectivity of the material, as well as providing specific facts for a correct understanding of events.

The fight against deepfakes requires joint efforts on the part of journalists, the media, governments and civil society. International journalism standards can serve as a basis for developing strategies to combat disinformation and protect society from the harmful effects of deepfakes. It is also important to educate journalists and the public about deepfake recognition methods in order to reduce the likelihood of their impact on public consciousness.

References

Aimeur E., Amri S., Brassard G., Fake news, disinformation, and misinformation: a review, Social Network Analysis and Mining 13 (2023)

Chawla R., Deepfakes: How a pervert shook the world, International Journal of Advance Research and Development, Vol 4, 2019

Chesney, R. y Citron, D. K. (2018). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3213954

Dobrobaba M. B. Dipfejki kak ugroza pravam cheloveka//Lex Russica, 2022

Francisco Javier & Ruiz del Olmo. Vídeos falsos y desinformación ante la ia: el deepfake como vehículo de la posverdad// Revista de Ciencias de la Comunicación e Información. 2024. Vol. 29, 1-14

Nelson, A. y Lewis, J. A. (2019). Trust your eyes? Deepfakes policy brief. Center for Strategic and International Studies.

Neuman, N., Fletcher, R., Eddy, K., Robertson, C. T. y Nielsen, R. K. (2023). Reuters Institute Digital News Report 2023. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023

N. Nour, J. Gelfand, Deepfakes: a digital transformation leads to misinformation, The Gray Journal 18 (2) (2022) 85–94, https://doi.org/10.26069/greynet2022-000.471-gg.

Shahid, F., Kamath, S., Sidotam, A., Jiang, V., Batino, A. y Vashistha, A. (2022). "It Matches My Worldview": Examining Perceptions and Attitudes Around Fake Videos. En Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-15) Association for Computing Machinery.

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, 6(1).

Electronic resources

European Parliament (2023). MEPs ready to negotiate first-ever rules for safe and transparent AI. European Parliament. https://acortar.link/a9Os2l

Kak neĭroseti obmanyvayut kazahstancev – https://mk-kz.kz/social/2024/02/14/kak-neiroseti-obmanyvayut-kazakhstancev.html

Kapitanov A. Dipfejk: nevinnaya tekhnologiya dlya razvlecheniya ili ugroza sovremennomu obshchestvu? – https://russiancouncil.ru/analytics-and-comments/analytics/dipfeyk-nevinnaya-tekhnologiya-dlya-razvlecheniya-ili-ugroza-sovremennomu-obshchestvu/

Mariyam Ospanova. Kak raspoznat' dipfejk-video

- https://factcheck.kz/metodika-fch/kak-raspoznat-dipfeyk-video

Nikolas Kejdzh na «Golubom ogon'ke». Bloger pokazal krutoj dipfejk – https://vgtimes.ru/news/73934-nikolas-keydzh-na-golubom-ogonke-bloger-pokazal-krutoy-dipfeyk-video.html

Reklama sredstva dlya pohudeniya s Rozoj Rymbaevoj poyavilas' v Runete – https://tengrinews.kz/show/reklama-sredstva-pohudeniya-rozoy-ryimbaevoy-poyavilas-238670/

Zashchity ot pornograficheskih dipfejkov ne sushchestvuet – https://adpass.ru/porno-s-bolshim-intellektom-pochemu-s-seksualnymi-deepfakes-nikto-ne-boretsya/

Information about authors:

Yessimova Manshuk (corresponding author) – Master's degree, Senior Lecturer at the Department of International Communications of the Ablai Khan Kazakh University of International Relations and World Languages (Kazakhstan, Almaty, e-mail: yemanshuk@mail.ru);

Shevyakova Tatyana – Doctor of Philology, Professor of the Department of International Communications of the Ablai Khan Kazakh University of International Relations and World Languages (Kazakhstan, Almaty, e-mail: david-sec@mail.ru @mail.ru).

Авторлар туралы мәліметтер:

Есімова Мәншүк (корреспонденттік авторы) – магистр, Абылай хан атындағы қазақ халықаралық қатынастар және әлем тілдері университетінің халықаралық коммуникациялар кафедрасының аға оқытушысы (Қазақстан, Алматы қ., e-mail: yemanshuk@mail.ru)

Шевякова Татьяна – филология ғылымдарының докторы, Абылай хан атындағы қазақ халықаралық қатынастар және әлем тілдері университетінің халықаралық коммуникациялар кафедрасының профессоры (Казахстан, Алматы, e-mail: david-sec@mail.ru@mail.ru).

Келіп түсті: 29 сәуір 2024 жыл
Қабылданды: 8 шілде 2024 жыл