IRSTI 19.01.11 https://doi.org/10.26577/HJ20257848

**M.S. Yessimova[1*], L.S. Akhmetova[2], T.V. Shevyakova[1], A.N. Zhabayeva[1]**

[1]Ablai Khan Kazakh University of International Relations and World Languages, Almaty, Kazakhstan
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan
*e-mail: yemanshuk@mail.ru

# MEDIA LITERACY AND INTERNATIONAL STANDARDS OF JOURNALISM IN THE CONTEXT OF THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

This article is devoted to the study of media literacy and international standards of journalism in the era of AI. Nowadays, with the rapid introduction of artificial intelligence technologies, media literacy is becoming key to preserving the quality of journalism and protecting citizens from various manipulations of public opinion. International standards of journalism, based on the principles of reliability, fact-checking and transparency of information sources, are becoming a guideline for editorial offices operating in a digital environment. One of the most acute challenges in recent years is the spread of various kinds of fakes and deepfakes – synthetically created audio and video materials that can distort reality and undermine trust in the media. In this article, the authors explore not only the issues of media literacy in the age of AI, but also the problem of the criminalization of deepfakes in countries around the world. In our opinion, strengthening regulatory regulation, including criminalizing the creation and distribution of deepfakes, is considered a necessary measure to prevent their use in disinformation campaigns and criminal purposes. Taken together, these factors underscore the need for a systematic approach to improving media literacy and compliance with international journalism standards in the new technological reality.

The purpose of the work is to study the application of media literacy and international journalism standards in the context of the development of artificial intelligence.

The research methodology is based on a comprehensive and interdisciplinary approach using analysis, synthesis, comparative and systematic methods, as well as a competence-based approach, which allowed us to consider media literacy and international journalism standards as interrelated elements in the context of the introduction of artificial intelligence.

The scientific novelty of the research lies in a comprehensive analysis of media literacy and international journalism standards under the influence of artificial intelligence technologies. The authors comprehensively examined the role of media literacy as a tool to counter modern digital threats, including the creation and distribution of fakes and deepfakes. Special attention was paid to the issues of their legal assessment and criminalization, which made it possible to justify the need for regulatory mechanisms to protect the information space.

**Keywords:** media literacy, digital literacy, information literacy, international standards of journalism, deepfakes.

М.С. Есимова[1]*, Л.С. Ахметова[2], Т.В. Шевякова[1], А.Н. Жабаева[1]

[1]Абылай хан атындағы Қазақ халықаралық қатынастар
және әлем тілдері университеті, Алматы, Қазақстан
[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан
*e-mail: yemanshuk@mail.ru

## Жасанды интеллект дамуы жағдайында медиасауаттылық және журналистиканың халықаралық стандарттары

Бұл мақала медиасауаттылық пен журналистиканың халықаралық стандарттарын жасанды интеллект дәуірінде зерттеуге арналған. Қазіргі таңда жасанды интеллект технологияларының қарқынды енгізілу жағдайында медиасауаттылық журналистиканың сапасын сақтау және азаматтарды қоғамдық пікірді әртүрлі манипуляциялардан қорғау үшін маңызды мәнге ие болуда. Дәлдік, фактчекинг және ақпарат көздерінің ашықтығы қағидаттарына негізделген халықаралық журналистика стандарттары цифрлық ортада жұмыс істейтін редакциялар үшін басты бағдар болады. Соңғы уақытта аса өзекті сын-қатерлердің бірі – әртүрлі фейктер мен дипфейктердің таралуы. Олар шындықты бұрмалап, медиаға деген сенімді әлсіретететін синтетикалық аудио және бейнематериалдар болып табылады. Бұл мақалада авторлар жасанды

гі медиасауаттылық мәселелерін ғана емес, сондай-ақ әлем елдеріндегі дипфейктерді қылмыстандыру проблемасын да қарастырады. Біздің ойымызша, дипфейктерді жасау мен таратуды қылмыстық жауапкершілікке тартуды қоса алғанда, нормативтік реттеуді күшейту – оларды дезинформациялық науқандарда және қылмыстық мақсатта пайдаланудың алдын алу үшін қажетті шара болып саналады. Жалпы алғанда, бұл факторлар жаңа технологиялық шындық жағдайында медиасауаттылықты арттырудың және халықаралық журналистика стандарттарын сақтаудың жүйелі тәсілін енгізу қажеттілігін айқындайды.

**Жұмыстың мақсаты** – жасанды интеллекттің дамуы жағдайында медиасауаттылық пен халықаралық журналистика стандарттарын қолдану мәселелерін зерттеу.

**Зерттеу әдіснамасы** талдау, синтез, салыстырмалы және жүйелі әдістерді, сондай-ақ құзыреттілік тәсілді пайдалана отырып кешенді және пәнаралық тұрғыда негізделген. Бұл медиасауаттылық пен халықаралық журналистика стандарттарын жасанды интеллектті енгізу жағдайында өзара байланысты элементтер ретінде қарастыруға мүмкіндік берді.

**Зерттеудің ғылыми жаңалығы** жасанды интеллект технологияларының ықпалы жағдайында медиасауаттылық пен халықаралық журналистика стандарттарына кешенді талдаудан тұрады. Авторлар медиасауаттылықтың заманауи цифрлық қатерлерге, соның ішінде фейктер мен дипфейктерді жасау мен таратуға қарсы тұру құралы ретіндегі рөлін жан-жақты қарастырды. Ерекше назар олардың құқықтық бағасына және қылмыстандыру мәселелеріне аударылды, бұл ақпараттық кеңістікті қорғаудың нормативтік тетіктерінің қажеттілігін негіздеуге мүмкіндік берді.

**Түйін сөздер:** медиасауаттылық, цифрлық сауаттылық, ақпараттық сауаттылық, халықаралық журналистика стандарттары, дипфейктер.

М.С. Есимова[1]*, Л.С. Ахметова[2], Т.В. Шевякова[1], А.Н. Жабаева[1]

[1]Казахский университет международных отношений и мировых языков
имени Абылай хана, Алматы, Казахстан
[2]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан
*e-mail: yemanshuk@mail.ru

## Медиаграмотность и международные стандарты журналистики в условиях развития искусственного интеллекта

Данная статья посвящена исследованию вопросов медиаграмотности и международных стандартов журналистики в эпоху ИИ. В наши дни в условиях стремительного внедрения технологий искусственного интеллекта медиаграмотность приобретает ключевое значение для сохранения качества журналистики и защиты граждан от различных манипуляций общественным мнением. Международные стандарты журналистики, основанные на принципах достоверности, проверки фактов и прозрачности источников информации, становятся ориентиром для современных редакций, работающих в цифровой среде. Одним из наиболее острых вызовов в последнее время является распространение различного рода фейков и дипфейков – синтетически созданных аудио- и видеоматериалов, способных искажать действительность и подрывать доверие к медиа. В данной статье авторами исследуются не просто вопросы медиаграмотности в эпоху ИИ, но также проблема криминализации дипфейков в различных странах мира. На наш взгляд, усиление нормативного регулирования, включая криминализацию создания и распространения дипфейков, рассматривается как необходимая мера для предотвращения их использования в дезинформационных кампаниях и преступных целях. В совокупности эти факторы подчеркивают необходимость системного подхода к повышению медиаграмотности и соблюдению международных стандартов журналистики в новой технологической реальности.

Цель работы – изучить вопросы применения медиаграмотности и международных стандартов журналистики в условиях развития искусственного интеллекта.

Методология исследования основана на комплексном и междисциплинарном подходе с применением анализа, синтеза, сравнительного и системного методов, а также компетентностного подхода, что позволило рассмотреть медиаграмотность и международные стандарты журналистики как взаимосвязанные элементы в условиях внедрения искусственного интеллекта.

**Научная новизна исследования** заключается в комплексном анализе медиаграмотности и международных стандартов журналистики в условиях влияния технологий искусственного интеллекта. Авторами всесторонне была рассмотрена роль медиаграмотности как инструмента противодействия современным цифровым угрозам, включая создание и распространение фейков и дипфейков. Особое внимание было уделено вопросам их правовой оценки и криминализации, что позволило обосновать необходимость нормативных механизмов защиты информационного пространства.

**Ключевые слова:** медиаграмотность, цифровая грамотность, информационная грамотность, международные стандарты журналистики, дипфейки.

## Introduction

In the era of fake news, media literacy is becoming an essential skill of modern people, as the ability to distinguish truthful information from misinformation not only makes them better able to understand information flows, but also makes people more aware and critically thinking. Modern technologies are developing extremely rapidly today, and in particular, technologies and types of artificial intelligence, the appearance of which in people's daily lives puts humanity in front of the need to solve new problems. One of these tasks is related to the consumption and distribution of fake news, as well as fake images and videos, the so-called deepfakes. All this is an even more difficult task, since now the number of formats has increased, and the speed of content creation is growing rapidly. To top it off, I would like to note that nowadays it is becoming increasingly difficult to determine which, for example, article was generated by AI, and which was written directly by a human. Therefore, it is extremely important not only to work on reducing digital inequality, but also to pay attention to digital and information media literacy. It's important to focus on how we interact with different platforms and fully understand how they work, especially in the age of AI. It should be noted that at the moment AI acts as an excellent assistant and a very convenient tool in journalism, although creative solutions are still in the hands of media workers themselves. Nevertheless, the amount of generated video, audio, and text content is only increasing, and under these conditions, the importance of media literacy and the ability to apply international journalism standards is increasing, which makes the research topic extremely relevant. Thus, the authors of the article believe that the study of this topic will make it possible to identify effective approaches to preserving the quality of media content and the responsibility of journalists in the era of widespread introduction and development of AI.

The purpose of the study is to determine the role of media literacy and international journalism standards in ensuring the quality and reliability of information in the context of using artificial intelligence technologies. To achieve this purpose the next objectives were set: conduct a review of the scientific literature; study different approaches to the interpretation of the terms media literacy, digital literacy, information literacy; analyze modern challenges to journalism in connection with the wide dissemination of AI; to evaluate the impact of AI on compliance with ethical standards in journalism, propose recommendations on the integration of media literacy and international standards into the professional practice of journalists.

The object of the research is media literacy and international standards of journalism as tools for ensuring the reliability of information in the context of the development of AI. The subject of the study is the influence of media literacy and international journalism standards on the quality and reliability of media content in the age of AI.

The core research question seeks to investigate the impact of artificial intelligence technology development and deepfake proliferation on compliance with international standards of journalism, and to determine the systemic measures-such as enhanced media literacy and the criminalization of deepfakes-required to safeguard the information space against disinformation and manipulation.

The research results can be used in the creation of educational programs on media literacy, the development of recommendations for journalists and editors on compliance with international standards in the digital environment, improving the quality of media content and building audience confidence, as well as in the formation of strategies for adapting media organizations to technological changes and the use of artificial intelligence in accordance with ethical standards.

## The materials and methods

Hundreds of books and scientific articles have been written on media literacy, and in many countries this discipline is taught in schools and universities. The research materials include works by domestic and foreign authors on media literacy, ethics of journalism, and the impact of artificial intelligence technologies on the modern media space. In turn, the methodology of this research is comprehensive and interdisciplinary. First of all, it should be noted that the work is based on a system of general scientific methods – analysis and synthesis. A systematic approach was actively applied, which made it possible to consider media literacy and international journalism standards as interrelated elements of a unified professional and educational system. In addition, the work uses a competence-based approach that allows assessing media literacy as a set of knowledge, skills and attitudes necessary for journalists and the audience to perceive and disseminate information in a high-quality manner. This is especially important

in conditions when the introduction of artificial intelligence technologies is changing the structure of information flows and formats of media interaction with society. The comparative analysis includes regulatory documents from the EU (DSA, DMA, Artificial Intelligence Act, 2024), the USA (NDAA 2020; Malicious Deepfake Prohibition Act, 2021), Great Britain (legislative amendments 2020-2025) and Kazakhstan (Civil Code of the Republic of Kazakhstan, art. 145; Law of the Republic of Kazakhstan "On Artificial Intelligence", 2025). This approach made it possible to identify differences in law enforcement practice and identify key trends in the development of AI regulation.

In 2018, the manual *"*Media literacy: a practical textbook on media and information literacy for higher education institutions*"* was published. Its goal was *"to enhance the ability of the populations of Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan to recognize false messages and disinformation, to enable them to make informed decisions, and to promote an understanding of the role and responsibility of the media in a democratic society"* (Shturkhetsky, 2018). In Kazakhstan, one of the most in-depth studies can be called the publication "Media literacy and media education: a handbook» (Akhmetova et al., 2019). According to the authors of the publication, "in almost every section of the handbook, the reader will find both abstract philosophical and sociological categories such as "language" or "rhetoric", as well as special terms that reveal the "secrets" of media professions and are used in the field of media (by journalists, editors, advertisers, planners, etc.). For the attention of readers, the handbook contains the most common words in this profession. In addition, definitions of some of the most important words and phrases of an international and political nature, as well as civil society terms, are given. All these terms are necessary, first of all, for schoolchildren and young people in order to be in demand in society and in the labor market." (Akhmetova et al., 2019: 3). It is also worth noting the textbook on media literacy for university lecturers in Kazakhstan, prepared by Internews in Kazakhstan, with Dr. Gulnara Asanbayeva, Candidate of Philosophical Sciences, Associate Professor, and Internews Regional Media Literacy Consultant, serving as the responsible editor. *(*Asanbayeva, 2019). The electronic textbook is divided into modules and is also supplemented with video lessons on YouTube. This edition provides the basic concepts and principles of media and information literacy, examines the features of digital literacy, and describes the role of critical thinking in the evaluation and assessment of information, including methods, technologies, and much more. In addition, the edition contains many engaging practical assignments that will be useful in teaching media literacy in higher education institutions.

Another important publication to mention is the *"Guide to media and information literacy"* (Orazymbetova, 2022). This guide systematizes knowledge of media and information literacy, taking into account the accessibility of information for people with disabilities. The purpose of the guide is to provide readers, including librarians, with an understanding of the theoretical foundations for the formation and development of media literacy knowledge, as well as the role of the school library in fostering students' information culture.

There are multiple interpretations of media literacy, information literacy, and digital literacy. The use of the term "media literacy" abroad began in the 1970s; however, for a long time, it coexisted alongside such concepts as functional literacy, audiovisual literacy, and information literacy (Houk. et al., 1974).

The concept of information literacy became established in academic discourse in the 1970s, when foreign researchers first emphasized the importance of acquiring a new set of skills, including the ability to recognize and use information. (Burchinal, 1976).

It is also important to mention the concept of "visual literacy" – initially understood as the necessity for a modern individual to perceive content delivered through various mass media. The term was introduced in 1968 by John Debes, the founder of the International Visual Literacy Association. (Bamford, 2003)

At the same time, taking into account the close interrelation between media literacy and information literacy, UNESCO decided to merge the two concepts into a single notion – media and information literacy. (FEZ declaration on media and information literacy, 2011)

Regarding interaction with the information environment at a new stage, researcher Bondarenko writes that "This is no longer just some kind of improvement of groups of human relations (1), environmental development (2), internal development and reflection (3). This is still the same (somewhat veiled) advertisement of the consumer society, where independent thinking is reduced to walking along already prepared paths. There is no place for creative professions in this society – only those who are amenable to this type of optimization. We are not

creating new texts – there is a long-term rewrite of the existing one ahead. Yes, traditional journalism, writers and poets, and artists can step aside, giving way to actions such as recreate & reproduce – in direct translation, to recreate and reproduce – but not to creativity, which is simply not provided for in this model. An artist becomes a person who can develop the right prompts (queries), and creativity boils down to tracking the work of the neural network and stopping at the stage when the results of the work at least partially begin to bring satisfaction" (Bondarenko E.A., 2023). The article "Digital literacy in the age of artificial intelligence: challenges and strategies for media education" (Boldyreva T.V. et al., 2025) considers that the introduction of AI technologies radically changes educational processes, and one of the main tasks of modern Russian education is to create a secure digital environment. Digital technologies, of course, offer many advantages: they help to do interactive tasks, automate the routine work of teachers, and simulate different situations, but they also have disadvantages, such as an increase in the number of deepfakes and false information.

Given the variety of definitions of digital, information, and media literacy, we have decided to combine the main interpretations, as well as their differences and examples, in the following table.

**Table 1 –** Interpretations and differences between digital literacy, information literacy, and media literacy

| Criteria | Digital literacy | Information literacy | Media literacy |
|---|---|---|---|
| Definition | The ability to effectively use digital technologies to search, process and create information. | The ability to search, evaluate, analyze, interpret, and use information. | The ability to analyze, critically perceive and create media content (including mass media, social networks, advertising). |
| Key skills | Use of computers, smartphones, Internet resources, and software. | Search, selection and evaluation of information from various sources; understanding the context and significance of information. | Assessment of the reliability of media information, understanding ways of manipulation and disinformation in the media, creation and dissemination of content. |
| Examples | Using e-mail, social media, and creating digital documents. | Search for scientific articles on the Internet, analyze news, evaluate the reliability of data. | Critical perception of news on social media, creation of a blog or a YouTube video. |
| The main focus | Technological competence and work with digital devices. | Processing and using information to make decisions and solve problems. | Analyzing media content, critically recognizing its impact, and creating your own media content. |
| Application examples | Working with online banking, using online courses, participating in webinars. | Preparation of scientific research, solving educational tasks, analysis of statistical data. | Analyzing and creating YouTube videos, understanding the impact of advertising and news on the Internet. |
| Connection with critical thinking | Importance in the context of security and efficiency of working with digital devices. | It is required for accurate assessment and use of information, avoiding false data. | The main focus is on distinguishing between media manipulation and one's own opinion. |
| The purpose of the training | Training in the effective use of digital technologies and Internet resources. | Training in methods of working with information to gain knowledge and make informed decisions. | Teaching critical perception of media information and safe interaction with the media. |

Source: Compiled by the authors

As you can see, this table highlights how digital, information, and media literacy are aimed at developing certain user skills: from working with technology (digital) and information processing (informational) to critical perception and media content creation (media).

**Results and discussion**

It's no secret that "fake news" is considered one of the main threats to humanity today and for good reason – technology is developing so rapidly that it can deceive even the most experienced users. That

is why today media and digital literacy should help users classify and systematize hundreds of manifestations of fake news – from sensational journalism, pseudoscience to the extremely dangerous phenomenon of deepfake, which, with the help of artificial intelligence and the most modern image generation technologies, makes it possible for fake news to be voiced by a real person. Should we now trust our eyes and ears in the global information environment? The simple answer is not to trust only one channel through which we receive information, but to seek confirmation from other sources. In an era when every Internet user can act as a source of information, the fundamental rule of journalism is to verify information from two or more sources.

The challenges faced by modern journalism are growing, as in addition to its informational and educational function, it has taken on the function of fact-checking and created platforms to refute fake news. Politicians also have a growing moral responsibility to not legitimize this type of online content, which draws strength from human fears and prejudices. Major online platforms have already created mechanisms to block fake news, but the fight against them is uneven. They spread many times faster than real news information, which links to real people, links to official sources and verifies the authenticity of the published information, and behind this information are real journalists, verified facts and events. (Dimov et al., 2020).

Artificial intelligence has also made it possible to quickly and easily create and spread disinformation without requiring deep technical knowledge. Tools such as XAI's Grok, DeepFaceLab, or China's Zao app make it easy to manipulate photos, videos, and audio, allowing hoaxes to spread faster. Although most of the tools have not yet been fully developed, as technology evolves, the impact of disinformation may also increase. In combination with the process of spreading fake news and the need to increase the level of media literacy of the population, there is a problem of universal penetration of artificial intelligence (AI) into the media environment. Using it makes it easier and faster to generate pseudo-content and fake news in the same format as news. (Zafer Can Ugurhan et al., 2021). In fact, ChatGPT, created by OpenAI, is a natural language processing system trained using an extensive collection of conversations that mimic human ones. (Vidal, et al., 2019).

So, how can we deal with disinformation, fake news, and deepfakes today? In our opinion, it has long been necessary to adopt appropriate legislative norms to regulate AI in the country's information space and protect people in order to counter fakes and deepfakes. In the European Union, for example, there are already special laws such as the DSA (The digital services act package, 2020) or the Artificial intelligence act (EU special law, 2024), designed to regulate the field of AI. Exposing hoaxes and deepfakes is very important because there are different interests behind them, and they make people no longer believe everything they see and read. Today, it is especially important to strengthen the trust of citizens through various tools, such as the artificial intelligence-based WhatsApp chatbot for direct communication with users or Narrative Panel, a monitoring system that helps fact checkers and journalists understand how lies spread, identify disinformation trends and effectively counter them.

Today, researchers and technology company founders are jointly working on ways to track and label AI-generated content. By employing various methods and cooperating with news organizations, they aim to prevent further erosion of the public's ability to distinguish between truth and falsehood.

1. Manufacturers such as Sony, Nikon, and Canon have begun developing methods for embedding special "metadata" that identifies when and by whom a photograph was taken directly at the moment of image creation.

2. Some companies, including Reality Defender and Deep Media, have developed tools that detect deepfakes based on the underlying technology used by AI-powered image generators. ["V nogu s dipfeykami: primenenie tekhnologii i eticheskie aspekty"] (2024).

In addition, special algorithms can be used to track AI-generated content by analyzing it for signs of AI usage (for example, a distinctive style of text generation or specific fragments of text). Specialized systems are capable of detecting "AI traces", such as characteristic patterns in writing, unusual features in images or videos that are typical of generative models.

Deepfakes represent a serious challenge faced by humanity as a whole. In the media environment, there are numerous solutions capable, for instance, of recreating the face of a targeted individual so convincingly that the deepfake resembles the person even more than their real appearance. At the same time, the number of attacks using such technology in Kazakhstan has increased several folds, and forecasts suggest that this figure will only continue to rise. Of course, some deepfakes can be detected with the naked eye. For example, such images may

contain various defects: absence of wrinkles, unnaturally smooth skin, lack of moles, or mismatched hairstyles. However, one should not rely solely on this – countering deepfakes requires AI models as a counterforce, ideally in combination.

Yet even if all implemented methods prove effective and all major technology companies fully adopt them, people will still need to approach online content critically.

The use of deep synthesis technology creates grounds for classifying such acts, for example, as defamation under the qualifying criterion of using information and telecommunication networks, which is applicable to the public dissemination of disinformation. (Efremova M. A. et al, 2024) Definitely, theoretically, it is possible to criminalize deepfakes by adding relevant provisions to civil, administrative and criminal legislation.

However, the situation does not look so clear from the point of view of bringing a criminal to justice, since a cybercriminal, as is often the case, publishes malicious content in the name of personal security, in a foreign jurisdiction or through an anonymous network. In this context, current trends in legal regulation around the world are shifting towards self-censorship or self-regulation on digital platforms, including the media. Ideally, they should do everything possible to detect and remove deepfakes and other potentially dangerous content. However, in practice, the situation is different.

The government, represented by the regulator, defines the criteria for "correct" and "incorrect" content, and the digital platform accepts them for guidance and enforcement purposes. Failure to comply with this requirement may result in heavy fines or termination of operations within the national jurisdiction.

Criminalization of deepfakes is an important issue in the context of combating manipulation, falsification and the spread of disinformation. Deepfakes, or fake images, videos, and audio created using artificial intelligence technologies, can be used for deception, slander, manipulation, or even blackmail. Given their growing popularity and technological accessibility, many countries have begun to take measures to combat them at the legislative level.

Let's highlight the main problems associated with deepfakes, including:

1. Manipulation of public opinion, for example, deepfakes can be used to create fake news and videos, which can influence elections, public protests, and even interstate relations.

2. Defamation and reputation theft – using deepfakes to create false or offensive materials can harm the reputation of individuals, leading to possible lawsuits.

3. Blackmail and threats – the creation of compromising fake videos in which, for example, a person appears to be involved in criminal activity or indecent acts, used for blackmail purposes.

4. Fraud – it's no secret that deepfakes can be used to create fakes that lead to financial losses, for example, by creating fake audio recordings or videos with trusted personalities.

5. Legislation and criminalization of deepfakes.

What modern methods of dealing with deepfakes exist today? First of all, technological solutions such as the development of technologies for detecting deepfakes can be attributed here, and AI programs and algorithms such as Deepware Scanner or Microsoft Video Authenticator can analyze content and identify whether it was created using deepfake technologies.

Next, educational initiatives should be mentioned. Since the dangers of deepfakes are associated with a lack of understanding of their impact, an important aspect of the fight is to raise awareness and train users to identify content manipulation.

Finally, we would highlight platforms and social networks. Social networks such as Facebook, Twitter, and YouTube are implementing policies for labeling and removing deepfakes, which helps reduce their spread. These companies can also collaborate with governments to develop anti-fraud methods.

Given the global nature of the deepfakes problem, consideration of international experience is an important element of analysis, and comparing different regulatory models allows us to identify trends and assess the applicability of foreign approaches for Kazakhstan. In this regard, we will provide an overview of key legal measures in other countries.

A comparative analysis shows that the legislative authorities pay the most attention to issues of protection against the spread of disinformation, protection of privacy and prevention of cybercrime. This table demonstrates the diversity of strategies, but a common desire to create ethical and legal barriers against the unfair use of deepfakes. It should be emphasized that now the legislation of the Republic of Kazakhstan contains a regulatory framework that can and should be used to counteract the creation and dissemination of deepfake content.

**Table 2 –** Legal measures against deepfakes: an international comparative review

| Country/Region | Laws and Measures | Key provisions |
|---|---|---|
| USA (Federal Laws) | H.R.2395 (2021)<br><br>The National Defense Authorization Act (FY 2020) | Obliges online platforms to disclose information that the content was created by AI.<br><br>In 2020, US President Donald Trump signed the first law in the country's history aimed at countering counterfeiting (The National Defense Authorization Act for Fiscal Year 2020). The implementation of this law involved the preparation of research in the field of exerting undue influence on the domestic political processes of the country through the dissemination of false information, including deepfakes. In addition, competitions (Deepfakes Prize Competition) were organized in order to develop the most effective ways to counter deepfakes. The remuneration for the winners of the competition was set at 5 million US dollars.In 2021, the Malicious Deepfake Prohibition Act was passed in the United States, which makes it illegal to create and distribute deepfakes for the purpose of causing harm (Efremova M. A. et al, 2024). |
| USA (States) | California laws, etc. | Some states, such as California, have introduced their own laws criminalizing the creation and distribution of deepfakes, especially in the context of pornography that involves a person without consent (so-called "revenge porn" deepfakes). |
| European Union | Digital Services Act (DSA) and Digital Markets Act (DMA) (Abramova, 2023)<br>General Data Protection Regulation (GDPR) (2018) | Obligations for large platforms to combat illegal content (including deepfakes); requirements for transparency of algorithms.<br><br>Although this law primarily regulates data protection, it can also be used to protect individuals from unwanted manipulation of their images or personal data in the context of deepfakes. |
| Great Britain | Amendments to legislation (2020), new measures (2025) | The UK is considering criminalizing deepfakes in order to ensure security and prevent malicious activity. In 2020, legislative changes were proposed to increase penalties for the creation and distribution of pornographic deepfakes, as well as to protect against the use of such videos for the purpose of defamation or manipulation. The UK officially makes it a criminal offence to create and distribute intimate deepfakes. From 2025, the authors of such fakes may be imprisoned for up to two years. These measures have become part of a new strategy to combat online violence and protect victims from digital humiliation. |
| Australia | National measures against deepfakes | Punishments for using deepfakes to harm individuals or organizations; a ban on using the Chinese AI startup DeepSeek on government devices. |
| Kazakhstan | Civil Code of the Republic of Kazakhstan, Article 145 (2024)<br><br>The Law of the Republic of Kazakhstan "On Artificial Intelligence" (2025) | In Kazakhstan's legislation, there already exist the concepts of deception and abuse. These can be used to hold individuals accountable for the use of deepfakes, for example, for fraudulent purposes. There is also a law prohibiting the use of images of a specific person, including video recordings, without their consent. Article 145 of the Civil Code of the Republic of Kazakhstan strictly regulates the issue of using images, stipulating that no one has the right to use the image of any person without their consent, and in the event of death – without the consent of their heirs (Grazhdanskii Kodeks Respubliki Kazakhstan, 2024).<br><br>Recently signed by the President of the Republic of Kazakhstan K-J. Tokayev's Law "On Artificial Intelligence" creates a legal framework for regulating multimedia and synthetic content, which makes the possibility of legislative control over deepfake technologies even more realistic. The law provides for:<br>definition of the concepts "artificial intelligence system", "model", "synthetic result", etc.;<br>mandatory labeling of goods, works and services created with the help of AI – which can also apply to media / information content.<br>Distribution of responsibility between owners, developers and users of AI systems, depending on the degree of their participation and potential risk ("Ob iskusstvennom intellekte" (2025) |

Source: Compiled by the authors

Based on table 2, we want to highlight the key trends in the development of regulation.:

1. Strengthening the responsibility of digital platforms. In all jurisdictions, there is a shift from self-regulation to the establishment of legally significant responsibilities for large online platforms, including requirements for transparency of algorithms, content moderation mechanisms, and risk management (EU – DSA/DMA; UK – Online Safety Updates; USA – sectoral standards).

2. A risk-based approach to AI. AI regulation is based on the principle of risk differentiation: from prohibitions of high-risk systems to requirements for safety, testing and documentation (EU AIA; similar approaches are gradually being formed in the UK and Kazakhstan).

3. Fight against manipulative and malicious content. There is increasing attention to deepfake technologies, media manipulation, threats to national security and information integrity (USA – deepfake initiatives; EU – prohibition of certain manipulative practices; Kazakhstan – regulation of personal data and fake materials within the framework of information security).

4. Improving the standards of personal data protection. The trend towards harmonization and stricter requirements for the handling of data, including biometric and behavioral data, is observed both in the EU (GDPR + new acts), as well as in the UK and Kazakhstan (updating laws on personal data).

5. Institutionalization of AI ethics and algorithmic transparency – establishes the obligation of ethical standards, audit of algorithms, independent supervision and external impact assessment mechanisms (EU – mandatory audit; UK – ethical guidelines; Kazakhstan – formation of a state supervision system within the framework of the law "On Artificial Intelligence").

On November 10, 2023, the international organization Reporters Without Borders published the Charter of Paris on Artificial Intelligence and Journalism, a regulatory and ethical document containing ten principles aimed at forming ethical guidelines for the media sphere when using artificial intelligence technologies. This set of regulations defines the fundamental principles that ensure the protection of the quality of news content and the reliability of information in the context of the rapid development and integration of AI into journalistic practice. "The social role of journalism is to act as a reliable intermediary for society and individual citizens, and this is the cornerstone of democracy. Artificial intelligence systems can help in this, but only if they are used transparently, fairly and responsibly. By reaffirming these principles, we stand for independent journalism and commit to providing reliable news in the age of AI," reads the preamble to the document, timed to coincide with the Peace Forum that started on the same day in Paris (Volkov, 2023).

To illustrate the real threats and assess the readiness of the legal system of the Republic of Kazakhstan to regulate the phenomena associated with the use of deepfake technologies in Kazakhstan, an analysis of two cases was conducted.

**Case 1. Fake "interview" of Akim of the North Kazakhstan region**

**Case description.** In April 2025, a video "interview" by Gauez Nurmukhambetov, akim of the North Kazakhstan region, appeared on the Internet, in which he allegedly urged residents to take part in the "Immortal Regiment" campaign. However, the akimat has officially stated that this video is a fake generated by AI.

Later, a second similar video appeared: an "interview" with the mayor, allegedly dedicated to the gasification of the region. The akimat noted that it was also fake, and announced the filing of an appeal to law enforcement agencies.

**Content analysis:**

- the video was accompanied by a message stating that it was allegedly provided by the akim's press service, which increased the audience's trust. (Politsiya nachala proverku posle poyavleniya v seti "intervyu" s akimom SKO – https://tengrinews.kz/kazakhstan_news/politsiya-nachala-proverku-poyavleniya-seti-intervyu-akimom-568061/);
- the press service of the akimat of the north Kazakhstan region confirmed that the video was created using artificial intelligence;
- the police have launched an investigation into the dissemination of false information.

**Implications and significance for the regulation of deepfake technologies:**

- this case clearly shows that deepfake technologies are used to create pseudo-reports and interviews with public figures, which can manipulate public opinion;
- the lack of prior verification and labeling of such videos contributes to their dissemination and misleads citizens;
- this case justifies the need to institutionalize the verification of AI-synthesized content at the level of government agencies and the media – this is important to prevent destabilization and manipulation through deepfake.

**Case 2. Fake videos with Akims**

In April 2025, Bairaqmedia.kz reported an increase in the spread of fake videos involving Akims (regional leaders). (Feikovye video s uchastiem akimov rasprostraniaiut v Seti – https://bairaqmedia.kz/ru/article/feikovye-video-s-uchastiem-akimov-rasprostraniaiut-v-seti.html). According to media reports, attackers use AI to create videos where akims allegedly speak out on politically sensitive topics or call for participation in various actions. It is also noted that such videos are distributed through social networks and messengers with the support of a network of bots, and the National Cyber Incident Response Center (TSARKA CERT) of Kazakhstan has already warned about the growth of such threats. This example clearly illustrates that deepfake is used as a tool of political disinformation and manipulation, especially at the regional level. The mass distribution of videos with fake statements by officials can undermine confidence in the authorities, provoke protest sentiments and destabilize the socio-political situation.

Thus, timely detection of fake materials largely depends on the media literacy of the audience and the correct application of AI protocols, which causes:

1. The need to apply the norms of the new AI Law. Case analysis shows that at the time of the spread of fake videos in Kazakhstan, there were no specialized regulations governing labeling, verification and responsibility for the creation of synthetic media content. It was incidents such as fake videos involving akims that demonstrated the vulnerabilities of the legal field and the need for systemic regulation of deepfake technologies. The Law of the Republic of Kazakhstan 'On Artificial Intelligence,' signed by President K.-J. Tokayev on November 17, 2025, established the basic principles for AI regulation, including requirements for transparency, developer responsibility, and the prohibition of manipulative technologies. However, in the context of countering disinformation, its norms need to be further elaborated. In particular, clearer mechanisms are needed for mandatory labeling of content created using AI, procedures for verifying the authenticity of digital materials (including photos and videos), as well as the establishment of separate legal liability for the dissemination of misleading synthetic images and videos (deepfake).

2. The need to develop detection technologies. Since visual and audio-generative deepfakes are increasingly used in fraudulent and disinformation schemes (including cases with officials), it is important that the media and government agencies have access to tools for automatic detection of such fakes (deepfake detectors, video surveillance).

3. Increasing the level of media literacy. The systematic dissemination of deepfake highlights the importance of educating citizens, journalists, and government officials in critical evaluation of digital materials. Training programs and explanatory work should include information about the typical features of synthetic videos, methods of their verification, and rules for responsible dissemination of information in a digital environment.

As part of the study, the authors would like to offer recommendations on the integration of media literacy and international standards into the professional practice of journalists.

**Table 3** – Recommendations on the integration of media literacy and international standards into the professional practice of journalists

| The direction of integration | Recommendations | Expected result |
|---|---|---|
| Professional training | Include mandatory courses in media literacy, fact-checking, and dealing with digital threats, including deepfakes, in journalism training programs. | Improving the competence of journalists in assessing the reliability of information and protecting against manipulation. |
| Ethical standards | Update editorial codes taking into account international standards (reliability, transparency of sources, balance of opinions) and challenges of AI. | Strengthening audience confidence in the media and reducing the spread of misinformation. |
| Technological training | The development of tools for detecting deepfakes and other synthetic AI materials, the introduction of automated content verification systems. | Prompt detection and prevention of the publication of falsified materials. |

*Continuation of the table*

| The direction of integration | Recommendations | Expected result |
|---|---|---|
| Legal regulation | Familiarization of journalists with the legislative norms on combating disinformation, including the criminalization of the creation and distribution of deepfakes, as well as the Law of the Republic of Kazakhstan "On Artificial Intelligence" (2025) regulating transparency, responsibility of developers, and prohibiting manipulative AI technologies. | Compliance with legal norms and prevention of media involvement in illegal information operations. |
| Classroom work | Development of educational projects and headings aimed at improving the media literacy of the audience. | Formation of society's critical approach skills in the perception of information. |
| International cooperation | Participation in online and offline projects for the exchange of experience in the field of media literacy and journalism standards. | Enriching professional practices and implementing the best global approaches. |

Source: Compiled by the authors

The recommendations presented reflect the need for a systematic approach to integrating media literacy and international journalism standards into professional practice. They cover key aspects ranging from educational and ethical to technological and legal, which is especially relevant in the context of the proliferation of artificial intelligence technologies and related threats such as deepfakes. The implementation of these measures will not only improve the quality of journalistic work, but also strengthen the audience's trust in the media, as well as create legal and organizational mechanisms to protect the information space from manipulation and disinformation.

## Conclusion

The study showed that in the context of digital content, journalists should first pay attention to the following main points:

- it is necessary to realize that we live in an era of fraud and disinformation;
- it is recommended to check the displayed "resonant" content from several sources;
- in case of political content, you should wait for a public statement and/or double-check information from official sources;
- people who watch questionable videos should pay attention to the details, so if the image is too pixelated or blurred, and strange noises appear, this is a reason to wonder if it's fake. Unnatural movements often catch the eye in such videos, for example, facial expressions seem to "float", facial muscles twitch, and the person does not look the same as in real life;
- suspicion may also be aroused by other inconsistencies, for example, the voice and intonation do not match the usual ones, the physique seems different, the hairstyle or small facial features are strikingly different from the original.

This only strengthens our confidence that it is extremely important and necessary to combat such fakes today, and the question here is how to do it correctly. On the one hand, criminal liability for creating and distributing fakes can protect society from serious risks, and on the other, we must not forget about freedom of speech and the right of people to express their opinions. This means that laws must be flexible and change with new technologies, otherwise they simply will not keep up with new ways of deception and falsification.

Thus, media literacy and critical thinking are key tools in the fight against AI disinformation. People must be able to analyze what they see and hear in order to distinguish the truth from lies. Special programs such as Deepware and AI or Not can help identify AI traces, but they are not always available to the general public, require certain skills to use, and may give erroneous answers.

In this regard, the development of information literacy and the ability to visually recognize AI fakes is of great importance. So far, many fakes are not of high quality and can be identified by careful examination of the details.

The fight against disinformation created by neural networks is impossible without special laws and strategies to control AI, for example, the EU and Kazakhstan were among the first in the world to adopt laws on artificial intelligence.

The authors' analysis of cases of the spread of fake videos with the participation of akims of Kazakhstan demonstrates the real threats of manipulation of public opinion using deepfake technologies.

Synthetic content creates an impression of authenticity, undermines trust in the authorities and can destabilize the socio-political situation. The Law on Artificial Intelligence, adopted in November 2025, laid down the basic principles of AI regulation, including transparency, developer responsibility, and prohibition of manipulative technologies, but in the context of countering disinformation, it requires clarification of norms on labeling content, verifying its authenticity, and responsibility for the distribution of synthetic materials. It remains important to develop deepfake detection technologies and increase the media literacy of citizens and specialists for critical evaluation of digital materials.

The conducted research has shown that in the context of the rapid introduction of artificial intelligence technologies, the media space is undergoing qualitative changes affecting both the mechanisms of production and distribution of news content, as well as the nature of audience interaction with the media.

The analysis showed that artificial intelligence technologies can become both a useful tool and a source of certain risks for journalists. On the one hand, they can greatly facilitate work and increase its productivity, and on the other, it is these same technologies that sometimes contribute to the appearance of distorted or even fictitious information. Based on the above, we consider it necessary to combine the development of media literacy among the audience with the responsible observance of professional and ethical standards by journalists themselves. Only then can the negative effects of using new technologies in the age of AI be effectively countered.

The practical application of the research findings lies in the possibility of integrating the results into educational programs in journalism and communication, developing editorial policies for the media, and forming recommendations for the safe and responsible use of AI in the information environment.

Thus, in the context of the digital transformation of the media space, a combination of media literacy, ethical guidelines and critical thinking is of particular importance, which will not only increase society's resistance to disinformation, but also strengthen the role of journalism as a reliable source of objective information.

**References**


Abramova, A. (2023) The EU Digital Services Act and Digital Markets Act -The Impact on Tech Companies – https://sumsub.com/blog/the-eu-digital-services-act-and-digital-markets-act/

Artificial Intelligence Act. – https://en.wikipedia.org/wiki/Artificial_Intelligence_Act

Bondarenko, E. A. (2023) Iskusstvennyi intellekt i problemy mediaobrazovaniya [Artificial intelligence and the problems of media education] – https://mic.org.ru/vyp/37-2/iskusstvennyy-intellekt-i-problemy-mediaobrazovaniya/

Bamford, A (2021). The visual literacy white paper. – https://aperture.org/wp-content/uploads/2013/05/visual-literacy-wp.pdf

Boldyreva, T. V., Kvetkin, P. D. (2020) Tsifrovaya gramotnost' v epokhu iskusstvennogo intellekta: vyzovy i strategii dlya mediaobrazovaniya [Digital literacy in the age of artificial intelligence: challenges and strategies for media education] – https://elar.urfu.ru/bitstream/10995/146007/1/978-5-7996-4039-2_062.pdf

Burchinal, L.G. (1976). The Communication Revolution: America´s Third Century Challenge. College Station, TX: A&M University Library.

Deepfakes Accountability Act/H.R.2395, (2021), https://www.congress.gov/bill/117th-congress/house-bill/2395/text?utm_source=chatgpt.com

Efremova, M.A., Russkevich, E.A. (2024). Dipfeik (deepfake) i ugolovnyi zakon. [Deepfake and criminal law] // Vestnik Kazanskogo Yuridicheskogo Instituta MVD Rossii, Tom 15, No. 2. DOI: https://doi.org/10.37973/VESTNIKKUI-2024-56-13

Feikovye video s uchastiem akimov rasprostraniaiut v Seti – https://bairaqmedia.kz/ru/article/feikovye-video-s-uchastiem-akimov-rasprostraniaiut-v-seti.html

FEZ declaration on media and information literacy – https://milobs.pt/wp-content/uploads/2011/06/Fez-Declaration.pdf

General Data Protection Regulation (GDPR), (2018) – https://gdpr-info.eu/

Dimov P., Ivanov E. Falshivite novini. Razpoznavane na falshivite novini i dezinformatsiyata v săvremennata hibridna sreda za sigurnost. [Fake news] – https://rndc.bg/wp-content/uploads/2020/05/FakeNews-21-05-2020-pechat.pdf

Grazhdanskii Kodeks Respubliki Kazakhstan (Obshchaia chast') (s izmeneniiami i dopolneniiami po sostoianiiu na 01.09.2024 g) [The civil code of the republic of Kazakhstan (general part)].

Houk, A., Bogart, C. (1974). Media Literacy: Thinking About. Dayton, Ohio: Pflaum/Standart. https://unesdoc.unesco.org/ark:/48223/pf0000165633

Mediagramotnost' i mediaobrazovanie: spravochnik (2019) / sost.: L.S. Akhmetova, A.V. Verevkin, A.A. Kudabai, T.Yu. Lifanova, D.P. Shorohov. – Almaty: Qazaq universiteti. – 158 s.

Mediagramotnost': prakticheskoe uchebnoe posobie dlya vysshikh uchebnykh zavedeniy po mediynoy i informatsionnoy gramotnosti (avtor-sostavitel' – Sergey Shturkhetskiy, 2018) – https://mediasabak.org/media/3f1921fa-9f7b-46a6-b110-c08809510337.pdf

"Ob iskusstvennom intellekte", Zakon Respubliki Kazakhstan ot 17 noyabrya 2025 goda № 230-VIII ZRK

Politsiya nachala proverku posle poyavleniya v seti “intervyu“ s akimom SKO – https://tengrinews.kz/kazakhstan_news/politsiya-nachala-proverku-poyavleniya-seti-intervyu-akimom-568061/

Posobie po mediagramotnosti dlya prepodavateley vuzov vysshikh uchebnykh zavedeniy Kazakhstana. – Almaty: Internews, 2019. – 250 s.

Rukovodstvo po mediynoy i informatsionnoy gramotnosti, pod redaktsiey Zh. A. Orazymbetovoy. – Astana: Biblioteka Nazarbayev Universiteta, 2022. – 180 s.

The Digital Services Act package. – https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

The National Defense Authorization Act for Fiscal Year 2020 – https://www.congress.gov/116/plaws/publ92/PLAW-116publ92.pdf

Uchebnoe posobie po mediagramotnosti. – https://newreporter.org/uchebnoe-posobie-po-mediagramotnosti/

Vidal Ledo, M. J., Madruga González, A., Valdés Santiago, D. (2019). Inteligencia artificial en la docencia médica. Educación Médica Superior, 33(3).

V nogu s dipfeykami: primenenie tekhnologii i eticheskie aspekty. – https://habr.com/ru/companies/sberbank/articles/848772/

Volkov Grigorii. Eticheskii barometr dlia zhurnalistov v epokhu II. – https://novayagazeta.eu/articles/2023/11/13/eticheskii-barometr-dlia-zhurnalistov-v-epokhu-ii

Zafer Can Ugurhan, Y., Kumtepe, E.G., Kumtepe, A.T., Saykılı, A. (2021). From Media Literacy to New Media Literacy: A Lens into Open and Distance Learning Context. Turkish Online Journal of Distance Education, 21, 135-151.

***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 (Almaty, Kazakhstan, e-mail: yemanshuk@mail.ru);*

*Akhmetova Laila – Doctor of Historical Sciences, Professor of the UNESCO Department of Journalism and Communication of Al-Farabi Kazakh National University (Almaty, Kazakhstan, e-mail: laila_akhmetova@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 (Almaty, Kazakhstan, e-mail: david-sec@mail.ru);*

*Zhabayeva Aigerim – Senior Teacher of the Department of International Communications of the Ablai Khan Kazakh University of International Relations and World Languages (Almaty, Kazakhstan, e-mail: aigerim_miregia@mail.ru).*

***Авторлар туралы мәлімет:***

*Есімова Мәншүк Сыйлыққызы (корреспондент авторы) – магистр, Абылай хан атындағы Қазақ халықаралық қатынастар және әлем тілдері университетінің халықаралық коммуникациялар кафедрасының аға оқытушысы (Алматы, Қазақстан, e-mail: yemanshuk@mail.ru);*

*Ахметова Ләйлә Сейсембекқызы – тарих ғылымдарының докторы, әл-Фараби атындағы ҚазҰУ ЮНЕСКО журналистика және коммуникация кафедрасының профессоры (Алматы, Қазақстан, e-mail: laila_akhmetova@mail.ru);*

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

*Жабаева Айгерім Нұршайхықызы – Абылай хан атындағы Қазақ халықаралық қатынастар және әлем тілдері университетінің халықаралық коммуникациялар кафедрасының аға оқытушысы (Алматы, Қазақстан, e-mail: aigerim_miregia@mail.ru).*

***Сведения об авторах:***

*Есимова Маншук Сыйлыккызы (корреспондентный-автор) – магистр, старший преподаватель кафедры международных коммуникаций Казахского университета международных отношений и мировых языков имени Абылай хана (Алматы, Казахстан, e-mail: yemanshuk@mail.ru);*

*Ахметова Ләйлә Сейсембекқызы – доктор исторических наук, профессор кафедры ЮНЕСКО по журналистике и коммуникации Казахского национального университета имени аль-Фараби (Алматы, Казахстан, e-mail: laila_akhmetova@mail.ru);*

*Шевякова Татьяна Васильевна – доктор филологических наук, профессор кафедры международных коммуникаций Казахского университета международных отношений и мировых языков имени Абылай хана (Алматы, Казахстан, e-mail: david-sec@mail.ru);*

*Жабаева Айгерим Нуршайхиевна – старший преподаватель кафедры международных коммуникаций Казахского университета международных отношений и мировых языков имени Абылай хана (Алматы, Казахстан, e-mail: aigerim_miregia@mail.ru).*

*Received: September 19, 2025*
*Accepted: December 8, 2025*