HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

MORAL THREATS IN THE AGE OF ARTIFICIAL INTELLIGENCE AND MECHANISMS FOR THEIR PREVENTION

Ruziboev Dilshod Azam ugli
University of Information Technologies and Management
Acting Associate Professor, Department of History
Rasulova Sevinch Majid qizi
2nd-year student, National Idea, Fundamentals of Spirituality and Legal Education program
https://doi.org/10.5281/zenodo.23067087

Annotatsiya. Mazkur maqolada sun'iy intellekt texnologiyalarining jadal rivojlanishi sharoitida yuzaga kelayotgan ma'naviy tahdidlarning mazmuni, shakllari, ijtimoiy ong va yoshlar dunyoqarashiga ta'siri hamda ularning oldini olishning institutsional, huquqiy, ta'limiy va ma'rifiy mexanizmlari tahlil qilinadi. Tadqiqotda sun'iy intellektning tarixiy taraqqiyot bosqichlari, generativ sun'iy intellektning ommalashuvi, deepfake, dezinformatsiya, manipulyativ kontent, raqamli qaramlik, shaxsiy ma'lumotlardan foydalanish, algoritmik stereotiplar va insonning mustaqil fikrlash qobiliyatiga ta'sir kabi masalalarga alohida e'tibor qaratilgan. Jahon iqtisodiy forumi, UNESCO, International AI Safety Report hamda O'zbekiston Respublikasining sun'iy intellektni rivojlantirishga oid strategik hujjatlaridagi ma'lumotlar asosida zamonaviy tahdidlarning ko'lami yoritilgan. Maqolada ma'naviy tahdidlarni faqat texnologik cheklovlar orqali emas, balki media va raqamli savodxonlik, tanqidiy tafakkur, axloqiy tarbiya, huquqiy madaniyat, ota-ona va ta'lim muassasalarining hamkorligi hamda inson markazli sun'iy intellekt siyosati orqali bartaraf etish zarurligi asoslanadi.

Kalit so'zlar: sun'iy intellekt, generativ sun'iy intellekt, ma'naviy tahdid, raqamli makon, dezinformatsiya, deepfake, media savodxonlik, raqamli savodxonlik, axloqiy qadriyatlar, tanqidiy tafakkur, yoshlar, axborot xavfsizligi, raqamli madaniyat.

Abstract. This article analyzes the content and forms of moral threats emerging amid the rapid development of artificial intelligence technologies, their impact on public consciousness and the worldview of young people, as well as institutional, legal, educational, and awareness-raising mechanisms for preventing such threats. Particular attention is paid to the historical stages of artificial intelligence development, the widespread adoption of generative artificial intelligence, deepfakes, disinformation, manipulative content, digital addiction, the use of personal data, algorithmic stereotypes, and the impact of artificial intelligence on individuals' ability to think independently. Based on data from the World Economic Forum, UNESCO, the International AI Safety Report, and strategic documents of the Republic of Uzbekistan concerning the development of artificial intelligence, the article highlights the scale of contemporary threats. It substantiates the need to address moral threats not merely through technological restrictions, but also through media and digital literacy, critical thinking, ethical education, legal awareness, cooperation between parents and educational institutions, and a human-centered approach to artificial intelligence policy.

Keywords: artificial intelligence, generative artificial intelligence, moral threat, digital environment, disinformation, deepfake, media literacy, digital literacy, ethical values, critical thinking, youth, information security, digital culture.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

Аннотация. В статье анализируются содержание и формы духовно-нравственных угроз, возникающих в условиях стремительного развития технологий искусственного интеллекта, их влияние на общественное сознание и мировоззрение молодежи, а также институциональные, правовые, образовательные и просветительские механизмы их предупреждения. Особое внимание уделяется историческим этапам развития искусственного интеллекта, распространению генеративного искусственного интеллекта, дипфейкам, дезинформации, манипулятивному контенту, цифровой зависимости, использованию персональных данных, алгоритмическим стереотипам и влиянию искусственного интеллекта на способность человека к самостоятельному мышлению. На основе данных Всемирного экономического форума, ЮНЕСКО, International AI Safety Report, а также стратегических документов Республики Узбекистан в сфере развития искусственного интеллекта раскрываются масштабы современных угроз. Обосновывается необходимость противодействия духовно-нравственным угрозам не только посредством технологических ограничений, но и через развитие медиаграмотности и цифровой грамотности, критического мышления, нравственного воспитания, правовой культуры, сотрудничество родителей и образовательных учреждений, а также реализацию человекоцентричной политики в сфере искусственного интеллекта.

Ключевые слова: искусственный интеллект, генеративный искусственный интеллект, духовно-нравственная угроза, цифровое пространство, дезинформация, дипфейк, медиаграмотность, цифровая грамотность, нравственные ценности, критическое мышление, молодежь, информационная безопасность, цифровая культура.

INTRODUCTION

In the first quarter of the twenty-first century, the rapid development of digital technologies is fundamentally changing the ways in which humanity obtains information, creates knowledge, communicates, and organizes social relations. Artificial intelligence occupies a special place in this process, penetrating ever more deeply into science, education, medicine, industry, public administration, the economy, the mass media, and many areas of everyday life. In particular, the widespread adoption of generative artificial intelligence systems has sharply expanded the capacity to automatically produce text, images, audio, video, and software code. While such capabilities increase the efficiency of human activity, they also give rise to new social, ethical, and moral problems in the information environment. Although the development of artificial intelligence is a historically new phenomenon, its intellectual roots go back to the mid-twentieth century[1]. In 1950, Alan Turing discussed whether machines could think, and in 1956, at the Dartmouth Conference, the term “artificial intelligence” entered wide use as the name of a scientific field. Over the following decades, expert systems, machine learning, neural networks, and deep learning technologies developed. In the second half of the 2010s, generative models advanced on the basis of big data and high computing power, and in the 2020s their opening to mass users sharply transformed the role of artificial intelligence in society. The social scale of this transformation is enormous. The World Economic Forum’s Global Risks Report 2025 notes that roughly two-thirds of the world’s population, or about 5.5 billion people, use the internet, and that more than five billion people are present on social media[2]. The report emphasizes that generative artificial intelligence is making it cheaper to produce false or misleading content in the form of text, images, voice, and video, and is enabling its dissemination on a large scale. Under these conditions, the issue of moral security acquires particular

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

scientific and practical significance. As the speed and scale of information grow, so do the demands placed on a person's ability to evaluate it critically, to distinguish truth from manipulation, and to differentiate ethically acceptable content from unacceptable content. Therefore, in the age of artificial intelligence, the concept of a moral threat cannot be limited to harmful information or extremist materials alone. It encompasses a broad set of digital influences that may negatively affect an individual's values, worldview, independent thinking, social responsibility, ethical decision-making, and real social relationships[].

METHODOLOGY

The methodological basis of the study consists of the principles of historicism and systematicity, comparative analysis, content analysis, the generalization of statistical data, and an institutional approach. The development of artificial intelligence was examined as a historical process, and its evolution from early theoretical ideas to the stage of modern generative models was analyzed. Moral threats, in turn, were considered on the basis of the interrelation of technological, informational, social, psychological, and ethical factors. The study draws on UNESCO's guidance on the use of generative artificial intelligence in education and research, the World Economic Forum's Global Risks Report 2025, data from the International AI Safety Report 2025, and the strategic documents of the Republic of Uzbekistan on the development of artificial intelligence through 2030. In regulating artificial intelligence, UNESCO places particular emphasis on a human-centered approach and on the principles of data privacy, safety, equity, and meaningful use. In this study, the concept of a "moral threat" is interpreted as a set of informational and communicative influences that may harm the ethical values, social responsibility, historical and cultural memory, independent thinking, and social cohesion of an individual or society. Artificial intelligence itself is not assessed as a moral threat; the risk is determined by how the technology is used, the content it produces, the purpose of its use, and the presence or absence of human oversight[3].

RESULTS

The results of the study show that moral threats in the age of artificial intelligence can be viewed along several interrelated lines. The first and most rapidly spreading line is associated with disinformation and manipulative content. Because generative artificial intelligence makes it possible to create large volumes of text, images, audio, and video in a short time, it is sharply reducing the cost and time of producing information. Whereas producing a fake video once required a professional studio, editing software, and considerable technical expertise, today's generative tools automate parts of this process. The International AI Safety Report 2025 notes that general-purpose AI systems can be used for harmful fake content, voice impersonation, extortion, psychological manipulation, and damage to a person's reputation. At the same time, the report stresses that reliable statistical data for measuring the exact scale of such incidents are still insufficient. Deepfake technologies are a distinct manifestation of this problem. With a deepfake, a real person's face, voice, or movements can be artificially altered to show them saying things they never said or doing things they never did. The social danger of this technology is that it can damage not only the reputation of an individual but also trust in digital information as a whole[4]. The International AI Safety Report identifies deepfakes and other AI-generated fake content as tools for manipulating individuals, extortion, psychological pressure, and reputational harm. The World Economic Forum's 2025 report identifies misinformation and disinformation as among the most significant global risks in the near term. According to the report, over the two-year horizon to 2027, misinformation and disinformation rank as the top risk, and it is

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

emphasized that AI-generated content is making it increasingly difficult to distinguish authentic material from synthetic material. The report also notes that only 40 percent of respondents across 47 countries said they trust most news sources. The moral significance of this figure is considerable. Trust is one of the forms of social capital necessary for the stable functioning of society. When citizens do not trust sources of information, establishing the truth becomes more difficult, opposing perceptions between social groups intensify, and reaching consensus around shared values becomes more complicated. For this reason, the main danger of fake content created with artificial intelligence lies not only in the spread of "false information" but also in the erosion of people's confidence in their own ability to establish the truth[5].

The second significant threat is the possible decline of independent thinking and intellectual activity. Generative AI offers capabilities for writing texts, translating, drawing conclusions, generating ideas, solving problems, and summarizing information. These capabilities create great efficiency in education and research. However, if a person relies excessively on ready-made answers, the skills of analysis, searching for evidence, comparing sources, and drawing independent conclusions may not develop sufficiently. Therefore, the moral criterion for using artificial intelligence is not the question of "whether or not to use it," but rather "how and to what extent to use it." UNESCO notes that generative AI is developing very rapidly in education systems and that in many countries national regulatory mechanisms have not managed to adapt to these changes. The organization recommends an approach that is human-centered, age-appropriate, protective of data privacy, and pedagogically responsible. This approach is also important from the standpoint of moral education, since an educational institution is not merely a provider of knowledge but an institution that shapes the learner's attitude toward information and intellectual responsibility.

The third threat is related to the erosion of ethical norms in the digital space. Anonymity, virtual identities, and the ease of content production may weaken the sense of accountability among some users. Insults, cyberbullying, defamation, the unlawful dissemination of personal data, manipulative advertising, and false identities affect not only an individual's rights but also their moral integrity. Artificial intelligence, in turn, may expand the possibilities for automating such content or spreading it on a large scale.

The fourth threat is the possible distortion of historical and cultural memory. Generative systems produce answers on the basis of large volumes of existing data, but their output is not always historically accurate. AI systems may "invent" non-existent sources, cite incorrect dates or names, or conflate different historical events. This is especially serious in fields where the accuracy of sources is essential, such as history, culture, national heritage, and law. Therefore, verifying AI-generated historical or social content against independent sources is becoming an essential scholarly skill of the digital age.

The fifth threat is the reinforcement of algorithmic stereotypes and one-sided views. Artificial intelligence systems are trained on data created by people. If the initial data contain stereotypes, discrimination, or one-sided perceptions of particular groups, such distortions may also appear in the model's outputs. The World Economic Forum report notes that algorithmic bias may be amplified in connection with human and societal biases.

The sixth line concerns moral issues related to the right to privacy. Artificial intelligence systems are capable of processing large volumes of data. When a person's photograph, voice, correspondence, search history, or other digital traces are misused, this can harm their private life. For

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

this reason, data privacy is not only a technical or legal matter but also an ethical one. UNESCO likewise identifies the protection of data privacy as one of the key directions of generative AI policy[6].

The seventh threat is the possible reduction of real interpersonal communication and excessive attachment to technological tools. Artificial intelligence can be used as a conversation partner, adviser, translator, teacher, or creative collaborator. However, a person's social experience does not consist solely of exchanging information. Empathy, cooperation, responsibility, forgiveness, mutual trust, and social experience are formed through real human relationships. Therefore, technological conveniences should not fully replace interpersonal relationships.

In Uzbekistan, too, the development of artificial intelligence is becoming one of the promising directions of state policy. The strategic document of the Republic of Uzbekistan on the development of artificial intelligence through 2030 sets out tasks aimed at introducing artificial intelligence technologies into various sectors, increasing the country's technological potential, and developing the relevant infrastructure. Alongside the use of artificial intelligence, this process also requires that its ethical, legal, and moral consequences be taken into account.

DISCUSSION

Preventing moral threats in the age of artificial intelligence is not a problem that can be solved by a single technical filter or prohibition. New technologies are constantly evolving, and a protective mechanism on one platform may not work on another. Therefore, the main focus should be on building strong information immunity within the individual.

The first mechanism is strengthening media and digital literacy. A modern user should be able to check the source, author, and time of creation of any information before accepting it, as well as its consistency with other independent sources and the authenticity of visual material. It is advisable for educational programs to develop practical skills addressing questions such as "How should AI-generated material be verified?", "How can a deepfake be detected?", and "How can the reliability of a source be assessed?"

The second mechanism is developing critical thinking. One of the most important factors of moral security is a person's refusal to accept ready-made conclusions blindly. A pupil or student should be able to ask questions such as "Who created this information?", "What evidence is it based on?", "What do other sources say?", "Does the author have any interests at stake?", and "Could the AI's answer contain errors?"

The third mechanism is fostering a culture of responsible use of artificial intelligence in education. AI should be used not as a tool that thinks for the learner or completes assignments for them, but as a technology that supports learning. For example, a student may first analyze a problem independently and then use AI to test their own reasoning, look for opposing viewpoints, or edit a text. Such an approach combines the use of technology with independent thinking.

The fourth mechanism is legal regulation. Mechanisms of legal liability for deepfakes, the unlawful use of personal data, fraud, cyberbullying, and other harmful activities must be adapted to modern technological conditions. At the same time, regulation should not unjustifiably restrict people's rights to lawfully obtain information, to create, and to use technology. Thus, an effective legal policy must ensure a balance between security and innovation.

The fifth mechanism is labeling the origin of AI-generated content and developing authentication technologies. Showing which parts of content were created by a human and which with

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

the help of AI enables users to make informed decisions. However, the International AI Safety Report notes that such technical tools, including watermarking, are not an absolute solution and can in some cases be removed. Therefore, technical identification must work together with media literacy and legal mechanisms[7].

The sixth mechanism is cooperation among the family, educational institutions, and society. Minors in particular need special protection from the influence of information in the digital environment. Parents should not merely prohibit internet use but should teach their children to discuss the content they have viewed, verify its sources, and evaluate it from an ethical standpoint. Schools and higher education institutions, for their part, may develop digital culture as a separate subject or as an integrated component of other subjects.

The seventh mechanism is expanding reliable digital resources on national history and cultural heritage. If users can easily find high-quality and reliable content on history, literature, national values, or legal matters, their dependence on random and manipulative sources decreases. For this reason, it is important for state and scientific organizations to expand digital libraries, electronic encyclopedias, scientific databases, and verified educational resources.

The eighth mechanism is training specialists in AI ethics. In the future, it will not be enough for a programmer or engineer to possess only technical knowledge. They must also be knowledgeable about data privacy, algorithmic fairness, human rights, intellectual property, information security, and social consequences. Likewise, lawyers, educators, journalists, and civil servants also need to understand the capabilities and limitations of AI technologies.

The ninth mechanism is establishing a system of scientific research and continuous monitoring. Because AI technologies are changing very rapidly, today's risk model may change within a few years. Therefore, it is necessary to monitor moral threats, study digital behavior among young people, analyze the mechanisms by which deepfakes and disinformation spread, and test new methods of protection.

The tenth mechanism is a human-centered technology policy. UNESCO's approach is based on precisely this principle: AI should serve human interests and be applied in a safe, fair, and meaningful manner. This principle corresponds directly to the issue of moral security, since the development of technology should be judged not only by its speed or economic efficiency but also by its impact on human dignity and social well-being. When discussing the moral risks of artificial intelligence, it is also essential not to forget the technology's positive potential. AI can broaden access to education, improve communication tools for persons with disabilities, accelerate scientific research, and expand opportunities for translation and access to information. Therefore, the problem lies not in the existence of the technology but in its social and ethical governance. The International AI Safety Report likewise points to the need to manage risks while taking into account that general-purpose AI systems have many beneficial applications alongside the potential for harmful use[7].

CONCLUSION

The development of artificial intelligence represents the next major stage of technological progress in human history. With its help, many forms of information processing, knowledge creation, learning, production, and creative activity are being fundamentally transformed. However, as technological capabilities expand, so does the need for moral and ethical responsibility. The results of the study show that the main moral threats of the age of artificial intelligence may manifest themselves in areas such as disinformation and manipulative content, deepfakes, the misuse of personal data,

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

algorithmic stereotypes, digital dependence, the weakening of independent thinking, cyberbullying, the distortion of historical and cultural memory, and the decline of ethical responsibility in the digital space.

REFERENCES:

1. UNESCO. Guidance for Generative AI in Education and Research. Paris: UNESCO, 2023.
2. International AI Safety Report. International AI Safety Report 2025. 2025.
3. World Economic Forum. Global Risks Report 2025. Geneva: WEF, 2025.
4. Republic of Uzbekistan. Strategy for the Development of Artificial Intelligence Technologies in the Republic of Uzbekistan until 2030.
5. Turing A. M. Computing Machinery and Intelligence. Mind, 1950.
6. Floridi L. The Ethics of Artificial Intelligence. Oxford University Press.
7. UNESCO. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO, 2021.