UDC 62:681.5

ARE WE READY FOR AI: CHALLENGES, RISKS AND RESPONSIBILITIES

Kate Kasianova, Bachelor of Technical Sciences
V.N. Karazin Kharkiv National University
Kostantyn Yevhenovych Radoutskyi, Senior Lecturer
V.N. Karazin Kharkiv National University
Olena Lenska, Language Supervisor
V.N. Karazin Kharkiv National University

Keywords: artificial intelligence, AI regulation, AI risks, social impact, technological control, economic risks, algorithmic bias, job automation, data protectionethical technology.

The rapid development of artificial intelligence (AI) technologies poses a challenge to humanity in managing risk and responsibility. New AI systems create uncertain consequences on which the stability of society, the economy, and national security depend. The main challenges are related to the loss of control over technologies, their manipulative impact on people, and insufficient preparedness for new forms of threats. Technology brings both benefits and serious risks. The example of social media shows that even useful technologies can create dependencies, polarization, and misinformation. The situation with AI is similar, but the scale and speed of change are not comparable, and the consequences are deeper and more difficult to predict. These unaddressed challenges could lead to unpredictable disruptions across multiple sectors, such as governance, employment, and public health, highlighting the urgency of establishing effective regulatory frameworks [1] .

The main challenges and risks include social and psychological threats, economic risks, and the technological threat of loss of control. Social and psychological threats are related to the fact that algorithms hold users' attention, provoking addiction and polarization, and create deepfakes and fake voices, undermining trust and threatening democracy and security. These manipulations can erode public trust in governments, media, and institutions, paving the way for social unrest and political destabilization. Moreover, AI-powered recommendation algorithms reinforce echo chambers, further deepening societal divisions and making it harder to foster constructive dialogue. Economic threats manifest themselves in the replacement of jobs by AI systems, causing structural changes in the economy, and in the bias of algorithms, leading to discrimination in key decisions such as job selection or lending. The automation of routine tasks could benefit businesses by reducing operational costs, but it might also leave millions of workers unemployed, creating a significant challenge for social systems and labor markets. Additionally, biased algorithms may perpetuate inequalities by amplifying existing prejudices, resulting in unfair practices that harm vulnerable populations [2] .

The technological threat of loss of control relates to the possibility of creating AI that is capable of self-improvement, which creates the risk of technology spiraling out of control. This raises ethical and existential concerns, as autonomous systems might evolve beyond human understanding or control, posing threats not only to security but also to human agency. Furthermore, the emergence of manipulative AI systems that exploit human behavior through

subtle psychological cues endangers public order, political processes, and security, potentially leading to large-scale disruptions.

Regulatory principles such as data protection and algorithm transparency are important, but they alone do not eliminate all problems. Current regulatory measures have not kept pace with the rapid evolution of AI, and short-term solutions such as RLHF (reinforcement learning from human feedback) have only temporary effects. There is also a lack of international coordination, with fragmented regulations across countries, which can create loopholes and inconsistencies that are exploited by companies or malicious actors [1] .

Regulatory principles include privacy and data protection, accountability and transparency, and ethical use of technology. Privacy and data protection involve prohibiting the disclosure of personal information without user consent and restricting the use of biometric data. Stronger regulations are needed to prevent mass surveillance and unauthorized data collection, which could otherwise infringe on individual freedoms and human rights. The principle of responsibility and transparency requires the involvement of experts to resolve disputes, as well as the ability to disable personalization and limit the time of use of services, empowering users to regain control over their digital lives. Ethical use of technology implies that companies should be responsible for the behavior of their models, especially when these models influence critical decisions or societal processes. It also implies the creation of global institutions to coordinate the regulation and testing of technologies, ensuring that standards are consistent and aligned with human values [2] .

A comprehensive solution to the problem is to create global institutions to coordinate the development and deployment of AI, hold companies accountable for the use and consequences of their products, and collectively agree on the pace of AI deployment and testing to mitigate risks. These institutions should foster collaboration between governments, industry, and academia to promote responsible innovation while anticipating and addressing potential risks proactively. Policies must include strategies for workforce reskilling to address the employment shifts caused by automation, as well as mechanisms to monitor and manage the unintended consequences of AI [1] .

The challenge of AI is the need to balance development and control. These technologies can bring enormous benefits, such as advances in healthcare, education, and environmental protection, but without a responsible approach, they will create risks to security, the economy, and society. The goal of regulation is to ensure that AI is developed for the benefit of all people, preventing catastrophic consequences. It is essential to prioritize human-centric AI, where technology serves as a tool to enhance human capabilities and well-being, rather than undermining them. This requires collective action, foresight, and the willingness to adapt regulations as the technology evolves, ensuring that the future of AI aligns with the common good [2].

References

1. Tristan Harris, Aza Raskin. How AI Poses Catastrophic Risks to Society [Video] // YouTube. – 2023. Available at: https://www.youtube.com/watch?v=xoVJKj8lcNQ (Access date: 28.10.2024).

2. Девять главных этических проблем искусственного интеллекта // AIN.UA. – 2016. Available at: https://ain.ua/ru/2016/11/07/devyat-glavnyx-eticheskix-problem-iskusstvennogo-intellekta/ (Access date: 28.10.2024).