Experts on the Malicious Use of Artificial Intelligence and Challenges to International Psychological Security

Report by Evgeny Pashentsev

Edition by the International Center for Social and Political Studies and Consulting
December 2021, Moscow

Experts on the Malicious Use of Artificial Intelligence and Challenges to International Psychological Security

Report by Evgeny Pashentsev

Edition of the International Center for Social and Political Studies and Consulting

December 2021, Moscow

УДК 004.8:340
ББК 32.813
Р30

Pashentsev, Evgeny

Experts on the Malicious Use of Artificial Intelligence and Challenges to International Psychological Security. Edition of the International Center for Social and Political Studies and Consulting. – Moscow: LLC «SAM Polygraphist», 2021. – 62 pp.

ISBN 978-5-00166-528-1

Funding: The reported study was funded by RFBR and VASS, project number N 21-514-92001.

The present publication is a result of the implementation of the research project titled "Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia," funded by the Russian Foundation for Basic Research (RFBR) and the Vietnam Academy of Social Sciences (VASS). The responses received from a targeted survey of nineteen experts from ten countries and their subsequent analysis aim to highlight the range of the most serious threats to international psychological security (IPS) through malicious use of artificial intelligence (MUAI) and determine how dangerous these threats are, which measures should be used to neutralize them, and what the prospects for international cooperation in this area are. This publication attempts to determine whether MUAI will increase the threat level of IPS by 2030. The publication pays special attention to the situation in Northeast Asia (NEA), where the practice of MUAI is based on a combination of a high level of development of AI technologies in leading countries and a complex of acute disagreements in the region.

Cover image: Shutterstock.

Signed to print 30.11.2021. Digital printing. Order № 105196.

Printed in the printing house «OneBook.ru» LLC «SAM Polygraphist».

109316, Moscow, Volgogradsky Avenue, Building 42, Bldg. 5, Technopolis Moscow.

www.onebook.ru

Contents

Main Abbreviations

AI – artificial intelligence

MUAI – malicious use of artificial intelligence

NEA – Northeast Asia

IPS – international psychological security

Introduction

Evgeny Pashentsev

Artificial intelligence (AI) technologies, despite their high significance for social development, raise threats to international psychological security (IPS) to a new level. There is a growing danger of AI being used to destabilize economies, political situations, and international relations through targeted, high-tech, psychological impacts on people's consciousness. Meanwhile, crisis phenomena are rapidly increasing in frequency, number, and severity worldwide.

There is no need to explain here why, in 2020, the Doomsday Clock was set to 100 seconds to midnight for the first time in history and remains unchanged in 2021 (Defcon Level Warning System, 2021). Nor is there a need to explain why the UN Secretary General is serving as a megaphone for scientists, warning bluntly that failure to slow global warming will lead to more costly disasters and more human suffering in the years ahead (Dennis, 2021). And there is no need to explain why the growth in the world's billionaires' fortunes from 8 to 13 trillion dollars in the crisis year of 2020 (Dolan, Wang & Peterson-Withorn, 2021)—against the backdrop of record economic decline in recent decades, hundreds of millions of newly unemployed people, and, according to the UN, the growth in the number of hungry people in the world from 690 million in 2019 (Kretchmer, 2020) to 811 million in 2020 (World Health Organization, 2021)—does not contribute to solving these and other acute problems of our time.

Economic problems, the degradation of democratic institutions, social polarization, internal political and interstate conflicts against the backdrop of the ongoing COVID-19 pandemic, all under the conditions of rapid AI development, create extremely favorable ground for the malicious use of AI (MUAI). MUAI is an intentional antisocial action, whether in explicit or implicit form. Antisocial circles (from individual criminals and criminal organizations to corrupt elements in government to financial and commercial structures to the media to terrorists and neo-fascists) are already increasingly taking advantage of this situation, favorable to their purposes.

The manipulation of the public consciousness is especially destructive in historical moments of crisis. The inhumanity of fascism became apparent after the death of 50 million in the flames of the Second World War. However, the technology of manipulating the public consciousness, with the appropriate funding from certain corporate structures, ensured Hitler's victory in the Reichstag elections in 1933—a distant year, but highly instructive for those alive today. It is hardly by accident that, today, the governments and parliamentarians of the USA, China, Russia, India, EU countries, and other states and associations to varying degrees and in different ways show growing concern about the threat of high-tech disinformation on the Internet and the role of leading media platforms that actively use AI technologies. The question is clear: *can humanity collectively find a way out of an increasingly difficult situation with a quantitatively and, increasingly, qualitatively higher level of manipulation of the public consciousness*?

In 2019, evidence of organized social media manipulation campaigns was found. These took place in 70 countries, up from 48 countries in 2018 and 28 countries in 2017 (CloudFlare, 2020). In each country, at least one political party or government agency had used social media to shape public attitudes domestically (Ibidem). Bots today have convincingly authentic online profiles and advanced conversational skills, and can appear to be legitimate users embedded in human networks. Some automated accounts are also partially managed by humans, using profiles known as "cyborgs" or "sock puppets" (Samuels & Akhtar, 2019).

The problem of the relationship between MUAI and IPS was first systemically raised by the author in a speech at a round table at the Ministry of Foreign Affairs of Russia in November 2018

(ICSPSC, 2018; Pashentsev, 2018). The topic was then developed in several publications, of both single authorship and co-authorship with colleagues (Averkin, Bazarkina, Pantserev & Pashentsev, 2019; Bazarkina, Dam, Pashentsev, Phan & Matiashova, 2021; Bazarkina & Pashentsev, 2019 and 2020; Pashentsev, 2019a, b and c; Pashentsev, 2020a and b; Pashentsev, 2021; Pashentsev & Bazarkina, 2021). The author considers it necessary, especially in modern international circumstances and taking into account the topic of this study, to focus on threats to IPS through MUAI, which, in real life, is in constant feedback and a position of mutual influence with the psychological security (PS) problem at the individual, group, and national levels.

As noted by the author in a recent study, new threats to agenda-setting and political stability are arising from the advantages of offensive and defensive psychological operations using AI. These advantages are increasingly associated with quantitative and qualitative departures from the traditional mechanisms of producing, delivering, and managing information; new possibilities for having psychological impacts on people; and the waging of psychological warfare. In particular, these advantages may include: (1) the volume of information that can be generated, (2) the speed at which information can be generated and distributed, (3) the believability of information, (4) the strength of the intellectual and emotional impacts that can be created, (5) the analytical data-processing capabilities that are available, (6) the use of predictive analytics resources based on AI, (7) the methods of persuasion that can be used, and (8) new capabilities for integration in the decision-making process. Based on a qualitative and rather approximate assessment of the data available from primary and secondary open access sources, the author draws the preliminary conclusion that advantages 1 and 2 have already been achieved, whereas advantages 3–8 are in the developmental stage at the operational level (Pashentsev, 2021, p. 143).

It should be noted that MUAI threats are growing in Northeast Asia (NEA). The region has not developed its own security system that would cover all countries of the region and serve all interests. Negative psychological impacts associated with various aspects of national and international development are increasingly affecting the sociopolitical situation and interstate relations in NEA. Recently, the pace of development of AI technologies there, especially in China, Japan, and South Korea, has sharply increased, which, despite the progressiveness of the achievement, also poses new challenges to IPS in the region, which require a timely response from state and non-state and national and international structures and institutions.

Meanwhile, the current systemic analysis of the MUAI and IPS problem within the framework of international cooperation leaves much to be desired and is fragmentary within the framework of MUAI research—not counting the efforts of the international group of specialists founded in 2019 to study the threats to IPS through MUAI, the Research MUAI group. The members of this group have published over 40 articles in indexed international academic journals on the topic of this study (Pashentsev, 2019c).

The above circumstances prompted the author to conduct a targeted expert survey as part of the implementation of the research project “Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia,” funded by the Russian Foundation for Basic Research (RFBR) and the Vietnam Academy of Social Sciences (VASS). The assessments given by nineteen experts from ten countries[1] obtained as a result of the expert survey and their subsequent analysis aim to highlight the most serious threats to IPS through MUAI and determine how dangerous these threats are to society, which measures should be taken to neutralize them, and what the prospects for international cooperation in NEA are. This survey attempts to determine whether MUAI will increase the level of threat to IPS by 2030. The experts paid special attention to the situation in NEA,

[1] Fifteen experts from Belarus, Cuba, France, Poland, Romania, Russia, the United Kingdom, the USA, and Vietnam have agreed to have their answers published.

where the practice of MUAI is based on a combination of a high level of development of AI technologies in leading countries and a complex of acute disagreements in the region.

The structure of this publication is designed in such a way that the reader can first get acquainted with the experts' answers to the questions posed, and then with their analysis.

The author expresses his gratitude to the RFBR, which has made this research possible; the experts, who have devoted their valuable time to preparing answers to the questionnaire; and the author's colleagues in the research project "Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia" Prof. Darya Bazarkina, leading researcher at the Institute of Europe of the Russian Academy of Sciences (Moscow), Dr. Nieet Dam, lecturer at HSE University (Moscow), Yuri Kolotaev and Ekaterina Mikhalevich, doctoral students at Saint Petersburg State University, and master's degree student Darya Matiashova for their help in forming an international knowledge base of experts on the topic of the survey. The author is also grateful to the authors' colleagues at the International Center for Social and Political Studies and Consulting (ICSPSC), which allowed this publication to come to fruition.

November 29th 2021

References

Averkin, A., Bazarkina, D., Pantserev, K., & Pashentsev, E. (2019). Artificial Intelligence in the Context of Psychological Security: Theoretical and Practical Implications. *Proceedings Of The 2019 Conference Of The International Fuzzy Systems Association And The European Society For Fuzzy Logic And Technology (EUSFLAT 2019)*, *1*, 101-107. doi: 10.2991/eusflat-19.2019.16

Bazarkina, D., Dam, V., Pashentsev, E., Phan, K., & Matiashova, D. (2021). The Political Situation in the Northeast Asia and Threats of Malicious Use of Artificial Intelligence: Challenges to Psychological Security. *Sotsialno-Gumanitarniye Znaniya (Social And Humanitarian Knowledge)*, *4*, 212-234. doi: 10.34823/SGZ.2021.4.51655

Bazarkina, D., & Pashentsev, E. (2019). Artificial Intelligence and New Threats to International Psychological Security. *Russia In Global Affairs*, *17*(1). doi: 10.31278/1810-6374-2019-17-1-147-170

Bazarkina, D., & Pashentsev, E. (2020). Malicious Use of Artificial Intelligence. New Psychological Security Risks in BRICS Countries. *Russia In Global Affairs*, *18*(4), 154-177. doi: 10.31278/1810-6374-2020-18-4-154-177

CloudFlare. (2020). How to Manage Good Bots. Good Bots vs. Bad Bots. Retrieved 5 November 2021, from https://www.cloudflare.com/learning/bots/how-to-manage-good-bots/

Defcon Level Warning System. (2021). Current Doomsday Clock Official Time Today. Retrieved 5 November 2021, from https://www.defconlevel.com/doomsday-clock.php#:~:text=January%2023%2C%202020%20to%202021%20%28Current%29%3A%20Doomsday%20Clock,2021.%20Click%20the%20change%20reasons%20to%20see%20why

Dennis, B. (2021). The U.N. chief's relentless, frustrating pursuit to bring the world together on climate change. Retrieved 5 November 2021, from https://www.washingtonpost.com/climate-environment/2021/10/25/antonio-guterres-climate-change/

Dolan, K., Wang, J., & Peterson-Withorn, C. (2021). The Forbes World's Billionaires list. Retrieved 5 November 2021, from https://www.forbes.com/billionaires/

ICSPSC. (2018). Prof. Evgeny Pashentsev spoke on Artificial Intelligence and Issues of National and International Psychological Security at the round table at the Ministry of Foreign Affairs of the Russian Federation. Retrieved 14 November 2021, from https://www.academia.edu/37933317

Kretchmer, H. (2020). Global hunger fell for decades, but it's rising again. Retrieved 5 November 2021, from https://www.weforum.org/agenda/2020/07/global-hunger-rising-food-agriculture-organization-report/

Pashentsev, E. (2018). Artificial Intelligence and Issues of National and International Psychological Security. Retrieved 14 November 2021, from https://www.alainet.org/en/articulo/196926

Pashentsev, E. (2019a). Destabilization of Unstable Dynamic Social Equilibriums through High-Tech Strategic Psychological Warfare. In N. Van der Waag-Cowling & L. Leenen (eds.), *Proceedings of the 14th International Conference on Cyber Warfare and Security ICCWS 2019 Hosted By Stellenbosch University and the CSIR, South Africa, 28 February – 1 March 2019* (pp. 322–328). Reading, UK: Academic Conferences and Publishing International Limited.

Pashentsev, E. (2019b). Malicious Use of Artificial Intelligence: Challenging International Psychological Security. In P. Griffiths & M. Nowshade Kabir (eds.), *Proceedings of the European Conference on the Impact of AI and Robotics 31 October -1 November 2019 at EM-Normandie Business School, Oxford* (pp. 238–245). Reading, UK: Academic Conferences and Publishing International Limited.

Pashentsev, E. (2019c). The Work of an International Group of Experts on Threats for International Psychological Security (IPS) by Malicious Use of Artificial Intelligence (MUAI). Retrieved 5 November 2021, from http://globalstratcom.ru/wp-content/uploads/2019/10/Новость-2-АНГЛ.pdf

Pashentsev, E. (2020a). AI and Terrorist Threats: The New Dimension for Strategic Psychological Warfare. In D. Bazarkina, E. Pashentsev & G. Simons (eds.), *Terrorism and Advanced Technologies in Psychological Warfare: New Risks, New Opportunities to Counter the Terrorist Threat* (1st ed., pp. 83–115). New York: Nova Science Publishers.

Pashentsev, E. (2020b). Malicious Use of Deepfakes and Political Stability. *Abstracts Of Papers Presented At The European Conference On The Impact Of Artificial Intelligence And Robotics ECIAIR 2020*, 82.

Pashentsev, E. (2021). The Malicious Use of Artificial Intelligence through Agenda Setting: Challenges to Political Stability. In F. Matos (ed.), *Proceedings of the 3rd European Conference on the Impact of Artificial Intelligence and Robotics ECIAIR 2021. A Virtual Conference Hosted by ISCTE Business School, Instituto Universitário de Lisboa, Portugal. 18–19 November 2021* (1st ed., pp. 138-144). Reading, UK: Academic Conferences International Limited.

Pashentsev, E., & Bazarkina, D. (2021). The Malicious Use of Artificial Intelligence against Government and Political Institutions in the Psychological Area. In D. Bielicki, *Regulating Artificial Intelligence in Industry* (1st ed., pp. 36–52). London and New York: Routledge.

Samuels, E., & Akhtar, M. (2019). Are 'Bots' Manipulating the 2020 Conversation? Here's What's Changed Since 2016. Comments. Retrieved 5 November 2021, from https://www.washingtonpost.com/politics/2019/11/20/are-bots-manipulating-conversation-heres-whats-changed-since/#comments-wrapper

World Health Organization. (2021). UN report: Pandemic year marked by spike in world hunger. Retrieved 5 November 2021, from https://www.who.int/news/item/12-07-2021-un-report-pandemic-year-marked-by-spike-in-world-hunger

Questions and Answers by the Experts

1. What threats to psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for the modern world? Why?

Vian Bakir and Andrew McStay

A potential and growing threat to the psychological security of individuals, groups, and nations is the use of AI to data-mine the body for the purposes of profiling: this adds a new, and more invasive, layer of surveillance to the already existing surveillance of our sentiments and emotions online. Termed "emotional AI," this is a weak form of AI in that it is designed to gauge and react to human emotions through text, voice, computer vision, and biometric sensing. It is a potential threat because of its fast adoption across the world; its methodological problems and the drive toward ever-increasing data surveillance (e.g., of users' lives in broader contexts) to fix these problems; context-specific public acceptance or rejection of this technology; regulation that is unprepared for its rollout, but that is starting to recognize its dangers in the wrong hands; and already-evidenced examples of worst-case scenarios in dominant forms of emotional AI.

- The emotional AI sector is growing fast across the world, and has been rolled out across multiple domains in society including transport, education, border control, health, workplaces, insurance, communication, and entertainment. Social media is already a dominant form of emotional AI (as it is designed to deploy AI to be maximally engaging to users, thereby making emotional expression viral on social media), but there are many other emergent forms.

- It promises much (to be able to read the emotional state of humans via a variety of inputs) but, to date, the methodology underpinning the commercial deployment of the technology is weak, leading to claims of accuracy being suspect. Despite these suspect claims, this does not appear to be slowing its rollout across the world.

- The emotional AI industry is aware of its methodological weaknesses; thus, industry leaders such as Microsoft are turning to methods that make greater use of social contextual awareness that surveil even more data. The methodological fix, then, is likely to lead to intensified surveillance.

- National surveys conducted by the Emotional AI Lab show that the UK public is not against deployment of emotional AI in certain contexts (e.g., for greater personalization and safety in cars) but rejects it in areas where there can be abuse of power (e.g., for the purposes of microtargeting in political campaigns or the surveillance of workers in the workplace).

- Regulation is not yet ready for the increasing use of emotional AI and how to distinguish between services of potentially pro-social value from less desirable applications. Nonetheless, the EU (which has the world's strictest privacy protections) published its draft AI regulations in 2021 to form the basis of its legislative program over the next few years, wherein the use of AI for manipulative purposes is banned, and emotional AI is seen as being of either 'limited' or 'high' risk.

- Worst-case scenarios of the malicious use of emotional AI as it applies to psychological security include the use of dominant forms of emotional AI (social media) to spread false, emotive, divisive narratives at moments of importance to the civic body, such as elections, referenda, and communal and religious gatherings where people from different social strata mix. Scholarship has already documented such activities across the world.

Raynel Batista Tellez

- *Automation of social engineering advertisement practices*: Users' online personal information is used to automatically generate custom malicious websites, emails, and links they would be likely to click on, convinced by chatbots that humans may trust to be another "real" person in a video chat.

- *Robot users or fake people*: Some forms of robots, such as drones or chatbots for customer service, imitate human behavior and record massive amounts of data beyond human limitations, simulating natural human language and behaviors through cognitive automation techniques.

- *Automation of hacking*: AI is used to improve target selection and prioritization for hacking purposes, evade detection, and creatively respond to changes in the target's behavior. Autonomous software has long been able to exploit vulnerabilities in systems, but more sophisticated AI hacking tools may exhibit much better performance both compared to what has historically been possible and, recently, compared to humans.

- *Fake news reports using deepfake technology*: Highly realistic videos are made of state leaders appearing to make inappropriate comments they never actually made, but that viewers usually trust.

- *Automating influence campaigns*: AI-enabled analysis of social networks is leveraged to identify key influencers, who can then be approached with malicious offers or targeted with disinformation.

- *Automated disinformation campaigns*: Individuals are targeted in swing districts with personalized messages to affect their voting behavior.

Robert Borkowski

I consider shifting the mood of the masses to be the greatest threat. New artificial intelligence tools allow much more effective, and invisible, manipulation of social moods and attitudes, which carries a significant threat in political life. I also do not rule out the possibility that new generations of terrorists will reach for the MUAI arsenal.

Anna Bychkova

- The aggravation of the problem of dependence upon users' information for the purpose of improving artificial intelligence algorithms that adapt to users' interests.

- The ability of a well-trained artificial intelligence algorithm to generate content that is perceived as having been created by a person and evokes certain emotions.

- The disunity of users on social networks due to the "echo chamber" and the "information bubble" ("filter bubble").

The producers of information is at the same time its consumers; they receive physiological and emotional pleasure, seeking a response to the content, which is expressed by the number of views, likes, and comments. Artificial intelligence algorithms are configured to promote the most radical views because it is such content that receives the greatest response on social networks. The creation of algorithms, like Generative Pre-trained Transformer 2 (GPT-2), that automatically write comments on a given topic and that are emotionally colored and can be perceived as having been written by a person (Expressive Text to Speech, IBM Watson Tone Analyzer) can be used to pursue political, economic, and extremist goals. Together, this provokes the polarization of public opinions, creates obstacles to reaching compromises, and, as a result, can be used to create an atmosphere of hatred and hostility.

Matthew Crosston

I do not agree entirely with the question asked as is.

Svetlana S. Gorokhova

The malicious use of AI technologies can cover the entire range of threats aimed at infringing upon the psychological security of an individual, different social groups, and society as a whole. Such an impact can affect emotional perceptions of surrounding reality and deform cognitive perceptions of information, even if the individual or group is aware of the intentional distortion. In addition, under the influence of these processes, the deformation of human consciousness as a whole is quite likely. Moreover, the vector of deformational changes will directly depend upon the goal of the attackers, and may encompass the entire spectrum of social deviation, starting with the development of religious (sectarian) fanaticism, and ending with the formation of extremist views. These become the most significant threats in connection with the emerging opportunities for the malicious use of AI technologies for selfish and criminal purposes.

Nguyen Quoc Hung

The use of AI for criminal and terrorist activities can be considered the most dangerous threat. The rapid progress in the field of AI increases the risk of this technology being applied to carry out automated attacks by criminals. The malicious use of AI poses threats to digital and political security, allowing perpetrators to carry out large-scale and potentially lethal attacks. The cost of conducting attacks may be lower when AI is used to perform tasks that would otherwise require human participation

Pavel Karasev

The use of information and communications technologies (ICT) by some leading countries as a means of achieving foreign and domestic policy objectives is a confirmed fact. In general, an increase in the use of ICTs for malicious information-based influence is a threat in itself. Technologies for the preparation and distribution of content are constantly improving, for example in terms of information targeting, user profile analysis, “fake news,” and the employment of opinion leaders to replicate this news. Skillfully crafted and carefully targeted information can have great effects on the opinions and perceptions of any population. The use of artificial intelligence (AI) for these tasks (namely the analysis of big data and the creation of texts) allows for a near-instant response and adaptation to the changes in the current situation and subtle, effective and, more importantly, routine manipulations of social behavior. The operations are carried out according to tailored scenarios and at the pace necessary for the earliest possible introduction of the given narratives into the minds of opinion leaders, denying the targeted side a prompt reaction to the information attack. This threat is the most relevant for the modern world, and its implications call for thorough consideration and study.

Alexander Raikov

Invulnerable and latent information attacks are the most dangerous threats in the modern psychological climate. An artificial intelligence (AI) system can target the psychological aspects in citizens’ lives. For example, the medical care system can seem unfair in certain regions of a country. This can be linked to various psychological factors that figure in this field, for example the influence

of local personalities. If a region has a major medical center with a charismatic chairperson of medicine who has trained generations of residents to practice in a particular way and those residents mostly end up working within a certain radius from the center, then people will see the effect of the chairperson in their area. People are affected by the situation in which they live. And when people from a neighboring region understand that they do not have such a medical canter with a charismatic chairperson, psychological stress may arise, increasing the likelihood of protest. A special AI system can reveal such a situation by analyzing big data and generating special information about unfair inequality of regions for malicious dissemination of this information in the "unfairly offended" region.

Marina Reshetnikova

The urgency of addressing psychological security problems has consistently increased over the past decade. The use of artificial intelligence (AI) has played a primary role in this, with its targeted high-tech impact on the consciousness of citizens. The development of mathematical support for various types of programs such as deepfake software or mobile applications such as Jinri Toutiao has only intensified the negative impact on IPS in general. An essential role in increasing the awareness of malicious use of AI is the creation of various groups of chatbots under state patronage. The tasks of these groups are very different—from destabilizing the situation within a particular group of civilians to influencing the decisions of international organizations.

Last but not least is the role that state structures play in the growth of problems for IPS. Not long ago, the meme about "Big Brother" following everyone was more of a joke, but then, in the COVID-19 period, the situation concerning "Big Brother" technologies—a specific set of modern technologies based on AI, capable of following any mass of people arbitrarily, broadly, and penetratingly—significantly escalated. Under the guise of their slogans being about care and the population's health, governments often solve their internal political problems. Moreover, this has not only been noticed by countries under authoritarian regimes, but also those with liberal governments. There is no need to look deeply for examples of this. There was a scandal in the southern provinces of Italy, where face scanners were installed in shopping centers under the guise of being temperature-measuring devices. Alternatively, another scandal surrounded the Israeli company NSO, whose Pegasus software could be used to spy on citizens. Attention has been drawn to the fact that this software is available in the public domain. Anyone can, completely legally, under contract with NSO, purchase it. At the center of this scandal was the murder in Istanbul of the well-known Saudi opposition journalist Jamal Khashoggi. The Saudi intelligence services were monitoring him with the help of the Israeli Pegasus software. This case caused a massive international scandal. Under the influence of the public, the Israeli Ministry of Defense raised the discussion of revoking the NSO corporation's license to further develop this software.

Vitali Romanovski

In my opinion, accelerated digitalization and the digital divide are among the most relevant threats. The pandemic has underscored the need for the flawless, uninterrupted operation of data processing infrastructure for the effective functioning of social and economic systems. The overall dependence of public services, business processes, and personal well-being on digital infrastructures and their security has become more apparent. As a result, individuals' and communities' physical surroundings and psychological safety have become more sensitive and vulnerable to malicious cyber- and cognitive attacks. At the same time, growing digitalization and digital interdependence have highlighted the cyber vulnerabilities of regional economies. The digital divide has raised the issue of unequal distribution of profits between digital "haves" and "have-nots." Competition for

technological dominance has gained momentum and exacerbated military–political contradictions between states and socioeconomic contradictions within societies. Malicious actors will likely continue to capitalize on these trends. It is increasingly feasible that AI technologies will appear in their toolbox shortly.

Sergey A. Sebekin

The possibilities of using AI for malicious purposes will increase in future. There have already been cases of computer programs managing to convince a person that they are talking with a real person. It is all about when advanced narrow AI will be able not only to convince a person of its reality, but also to have a person act at the pleasure of the person using the AI for malicious activities. This has become more important now that AI is actively used for psychological brainwashing or social engineering, and, most likely, will be used for antisocial purposes on a wider scale in the near future.

Pierre-Emmanuel Thomann

The most relevant threat to the future is the geopolitical threat; that is, the use of artificial intelligence to impose a unipolar world and prevent the emergence of a more multipolar world. The malicious use of AI to enhance state-sponsored acts of terrorism is also a relevant threat.

Marius Vacarelu

The main problems can be described by the Latin expression: *bellum omnium contra omnes*. It refers to the increased wish to confront other people, putting all other things aside.

2. How much does the malicious use of artificial intelligence increase the level of threat to international psychological security today?

	Bakir V. and McStay A.	Batista Tellez R.	Borkowski R.	Bychkova A.	Crosston M.	Gorokhova S.	Hung N. Q.	Karasev P.	Raikov A.	Reshetnikova M.	Romanovski V.	Sebekin S.	Thomann P.-E.	Vacarelu M.	Expert from Belgium	Expert from Russia	Expert from Vietnam #2	Expert from Vietnam #3
Strongly		V				V		V		V					V			
Noticeably	V			V			V		V			V		V		V	V	V
Only slightly			V		V						V		V					
Not at all																		

Vian Bakir and Andrew McStay

On the basis of the question specifying “today,” we suggest the answer “noticeably.” Biometric forms of emotional AI are still emergent and being trialed by governments across the world. However, in geographic areas that have few limitations on the state’s surveillance of populations, the

trials are already raising human rights concerns, with some calling for such technologies to be totally banned. Scholars have also documented efforts by international and national actors to use social media (themselves a form of emotional AI) to sow emotive, false, and divisive narratives among populations during elections and referenda, honing and targeting these to specific audiences (e.g., this has been documented in Brazil, the USA, Spain, the UK, and Nigeria). However, the question of their real-world influence (i.e., whether these efforts have actually tipped elections) is not yet proven (and may never be, given the complexity of disentangling what makes someone vote a certain way, or vote at all).

3. How much will the malicious use of artificial intelligence increase the level of threat to international psychological security by 2030?

	Bakir V. and McStay A.	Batista Tellez R.	Borkowski R.	Bychkova A.	Crosston M.	Gorokhova S.	Hung N. Q.	Karasev P.	Raikov A.	Reshetnikova M.	Romanovski V.	Sebekin S.	Thomann P.-E.	Vacarelu M.	Expert from Belgium	Expert from Russia	Expert from Vietnam #2	Expert from Vietnam #3
Strongly	V	V		V		V	V	V	V	V					V			
Noticeably			V		V						V	V	V	V		V	V	V
Only slightly																		
Not at all																		

Vian Bakir and Andrew McStay

The malicious use of dominant forms of emotional AI (e.g., using social media to spread divisive and deceptive narratives) will only increase in the very many areas of the world that have minimal regulations on social media, low levels of digital literacy, and pre-existing social tensions that can be stoked. Even if people are resistant to being influenced by such profiling systems, the attempts to influence people will spread. Mere awareness of these attempts could be enough to erode trust in the democratic process and the legitimacy of electoral results. For instance, President Trump's repeated false claims of voter fraud led to an attempt by disaffected Trump supporters in the USA to overturn the election of Joe Biden in the Capitol Hill attempted insurrection of 2021.

4. What measures (political, legal, technical or other) do you consider to be important to neutralize the threat to international psychological security caused by the malicious use of artificial intelligence?

Vian Bakir and Andrew McStay

Technical means are important. For instance, we need to be able to automatically detect false and hateful content online in order to assess it, flag it as problematic (e.g., having users make their

own decisions), or remove it. (Whether these content moderation decisions should be performed by the technology platforms or the government has no easy answer, as both could lead to undue censorship and abuse of platform or governmental power.) Automation is necessary because of the sheer scale and speed of the circulation of hateful, deceptive information. However, humans must always be in the loop when such decisions are made due to the nuanced nature of deceptive and hateful content and the importance of the human right of freedom of speech.

Legal means are important. For instance, the EU draft of AI regulations (2021) usefully considers the levels of risk that various AI technologies may pose, and presents a scale of prohibited, high risk, and low risk activities. Each of these risk categories impose specific obligations on those developing and deploying AI in society (rather than placing the burden on citizens to first prove harm). This more precautionary approach seems like a good start. For instance, it proposes that prohibited AI include that which "deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behavior in a manner that causes or is likely to cause that person or another person physical or psychological harm" (Title II Article 5). Arguably, this could include the use of social media to spread divisive and false narratives to swing an election (e.g., voter suppression strategies for people of color in the USA). Also prohibited is "the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behavior of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm" (Title II Article 5). This could cover the targeting of people on social media according to their psychological vulnerabilities (e.g., propensity to believe conspiracy theories). Title IV also imposes transparency obligations on "AI systems" that are used to detect emotions or determine association with (social) categories based on biometric data: they "must take account of the specific risks of manipulation they pose".

Education is also important. People's digital literacy needs to increase for them to understand what profiling is, how it may be achieved, and to what ends it may be used. For instance, the UK's 2016 "Brexit" referendum saw "dark ads" (online advertisements only seen by targeted recipients) being discussed in public for the first time, but three years later, by the time of the UK's 2019 General Election, many were still unaware of these techniques. YouGov survey research commissioned by the non-governmental organization, Open Rights Group, showed that although 54% of the UK population was aware of how political parties target or tailor adverts based on the analysis of their personal data (political microtargeting), almost a third (31%) were not very aware or not aware at all. Only 44% of the national sample were very or fairly aware of "dark ads," with a similar figure (41%) of the sample being not very or at all aware (Open Rights Group 2020). That there is still relatively low awareness after several years of public discourse on this issue is alarming in that it shows that a significant proportion of the electorate are unaware of how parties may attempt to manipulate them. Greater transparency of AI systems would be a positive first step, but educators and communicators (e.g., journalists and filmmakers) are needed to relay the social significance of these systems and the perils of the lack of transparency to audiences.

Raynel Batista Tellez

First, to promote effective international cooperation, MUAI elevates threats to IPS to a qualitatively new level, which requires an adequate assessment and reaction from society. The comprehension of the new, cross-cultural threats of MUAI lead to the formulation of large-scale strategies to protect sovereignty; enforce regional roles for building consensus, engagement, and international collaboration; and force experts to acknowledge the ethical implications of surveillance, persuasion, and physical target identification for regional equilibrium.

A cross-cultural approach of MUAI supports the idea of cultural competency as a mechanism of social influence and establishes the distribution of power from the perspective of security as a sociocultural phenomenon. AI is highly likely to have different social impacts on regional geopolitical balance, depending on people's cultural settings traced by customs, values, and behaviors.

"Cultural knowledge" refers to individuals knowing about some cultural characteristics, history, values, beliefs, and behaviors of other ethnic or cultural groups.

A strategy to neutralize the threat to international psychological security caused by the malicious use of AI must also include a cross-cultural approach, defining three levels of interaction:

- *Cultural awareness* is the stage of understanding other groups. It involves being open to the idea of changing cultural attitudes.

- *Cultural sensitivity* is knowing that differences exist between cultures, but not assigning values to the differences (better or worse, right or wrong). Clashes at this level can easily occur, especially if a custom or belief in question goes against the idea of multiculturalism. Internal conflict (intrapersonal, interpersonal, and organizational) is likely to occur at times over this issue. Conflict will not always be easy to manage, but it can be made easier if everyone is mindful of the organizational goals.

- *Cultural competence* brings together the previous stages—and adds operational effectiveness. A culturally competent organization has the capacity to bring into its system many different behaviors, attitudes, and policies and work effectively in cross-cultural settings to produce better outcomes.

Robert Borkowski

Much depends on the sound and honest policy of state authorities on counteracting MUAI threats, although counteracting, for example, the spread of fake news is extremely difficult. Moreover, politicians must have a strong will to take an honest approach to counteracting this threat and have a real understanding of the threat, which, unfortunately, is not visible in many countries. Social education and the development of society's awareness of threats and rational attitudes are very important.

Anna Bychkova

Today, we are witnessing the desire of both the state and the technocratic sector to replace the human factor in the search for and the neutralization and removal of certain content with technologies based on machine learning. Artificial intelligence becomes both regulator and judge: the program must train on the 'big data' it is provided according to programmed rules, so that later it is the program that decides the "fate of the content": delete, redo, or leave unchanged. Thus, technologies, in fact, not only determine which content will be in demand by users at that moment and in the near future, but also represent an analogue of the "state regulator," which can, based on the "conclusions" of artificial intelligence, determine the future fate of media resources. A comprehensive approach is needed based on the analysis of existing threats and forecasts of the main trends: a combination of political, legal, and technological measures, as well as measures in the field of educational policy aimed at improving media literacy, preventing information dependence.

Matthew Crosston

Other: education of people to understand how much the malicious use of AI capitalizes on a fundamental failure of people to discern information and analyze context, rather than confirmating MUAI as a true explicit threat on par with kinetic weapons.

Svetlana S. Gorokhova

At our current stage of historical development, we cannot and should not repeat the mistakes that were made by humanity in less enlightened eras; therefore, even at these early stages of the widespread introduction of new technologies into our lives, it is necessary to lay down appropriate norms in the legal field that not only provide the possibility of imposing retrospective responsibility on perpetrators, but also consider the prospects of establishing historical (prospective) responsibility for those who are engaged in the development and implementation of potentially dangerous and fundamentally new technologies that were simply impossible before. This long-term responsibility can principally be imposed by developing relevant regulatory legal acts concerning rules, duties, and prohibitions related to the general prevention of possible harm that, in the future, may be caused to citizens in the process of their interaction with the newly introduced and new technologies equipped with AI. This, prospective responsibility, can, first of all, be expressed in the inclusion in the relevant normative legal acts of rules, duties and prohibitions related to the general prevention of possible harm, which in the future may be caused to citizens during the interaction with the introduced new technologies, equipped with AI.

Nguyen Quoc Hung

- Effectively combating the actions of hostile forces and criminals who violate information security.
- Concentrating resources to create and gradually develop the information technology industry, especially the information security (cybersecurity) industry in Vietnam.

Pavel Karasev

One priority should be the creation of a monitoring system and the timely identification of signs of information influence operations. Taking into account the fact that the cognitive capabilities of any one individual are insufficient for analyzing and comprehending a huge volume of information in the global media sphere, it would be necessary to use AI technologies to develop this system. Another major task should be ensuring the capacity to provide a timely response to signs of upcoming information and political operations, including the refutation of fake news. It is important to realize that the challenge of countering information operations cannot be solved by only technical or legal means. Alternative narratives are needed, and their creation requires the convergence of disciplines from different branches of science – humanitarian, social, technical, and natural. To build accurate models that can form the foundation for machine learning, it is necessary to translate the current achievements of psychology, sociology, political science, and other humanities into the language of mathematics and computer programs.

Alexander Raikov

Political international collaboration on this topic is very important. I think that a special agreement has to be created and approved. The disparity in ethical codes can be analyzed and

adapted to make this agreement. Special security technologies are used to neutralize the threat to international psychological security. The new results in scientific studies in the field of hybrid, strong, general, and super AI must be taken into account while creating the method and tools for this neutralization. In a modern, multi-level economic system, the variety of approaches and management models and the variety of feedback leads to the corresponding management systems having unique responses to changing the conditions and factors of development and security. We cannot manage security if we do not have information about events or the knowledge, including implicit and hidden knowledge, that allow us to analyze and interpret events and make adequate decisions. AI systems can detect such events and therefore be used to maliciously correct feedback, causing irreparable economic and psychological damage to countries. Ironically, AI systems are the only technological measure for neutralizing the threat to international psychological security caused by the malicious use of AI.

Marina Reshetnikova

In the current situation, it is unlikely to expect political, legal, technical, or other actions directed toward IPS from government agencies. The post-COVID-19 situation plays an important role in this. The fight against the pandemic today is undoubtedly the most important task facing the governments of almost all countries worldwide. However, this raises the question of whether they are using the pandemic to solve their domestic political problems. Moreover, here again, we return to the controversy about the exacerbation of “Big Brother” technologies. The only way to neutralize the threats posed by the malicious use of AI is to create international public organizations and associations dedicated to monitoring and controlling the usage of such technologies.

Vitali Romanovski

The active role of the national government, interstate cooperation, and private sector involvement are essential to developing strategies to counter AI technologies’ employment by malicious actors. It is important to enhance interagency collaboration and information exchange upon applying AI and other digital technologies in the national security sphere. Governmental entities should also develop policies to increase the population’s resilience to offensive cognitive operations from other states and non-state actors.

Sergey A. Sebekin

Since AI has been used to exert psychological influence, it would be logical to assume that the weak element is the person, not the technologies that they use and through which it is possible to influence others. It is important for people to be taught a high level of critical thinking so that they are less likely to be gullible and succumb to various AI-based provocations.

Pierre-Emmanuel Thomann

The promotion of a more multipolar world, with strong international cooperation platforms at different levels (local, regional, global) would help to neutralize the threat to international psychological security

Marius Vacarelu

A strong education to ethics, but we must admit that a complete neutralization is impossible.

5. How important is international cooperation in successfully countering the malicious use of artificial intelligence? On what international platforms (and why) is this cooperation the most effective? What are the existing obstacles to this cooperation?

Vian Bakir and Andrew McStay

International cooperation between all stakeholders is very important, but is probably not sufficient, and requires the support of strong regulation at supra-national levels. As a case in point, a Code of Practice of Disinformation was signed by dominant social media platforms between 2018 and 2020, and set a wide range of commitments. These include transparency in political advertising; demonetization of purveyors of disinformation; closure of fake accounts; the empowerment of users to report disinformation and to understand why they have been targeted by an advertisement; the empowerment of researchers by providing data; and the prioritization of authentic, accurate, and authoritative information to users while not preventing access to otherwise lawful content or messages solely because they are thought to be "false." However, disinformation remains prevalent online, and as a result, the EU may be moving towards a more assertive co-regulatory approach in its forthcoming Digital Services Act.

An international framework for tackling the malicious use of AI is needed, otherwise states are likely to impose their own solutions, which may well contravene important human rights such as freedom of expression. For instance, coercive responses of many governments seeking to tackle online disinformation have included arrests, Internet shutdowns, and legislation on fake news that stifles dissenting views.

Raynel Batista Tellez

Cultural competency gives international relations and the distribution of power the capacity to promote actors' cooperation and to create a sense of belonging and identity. The concept of power is central to international relations. Power is the production, in and through social relations, of effects that shape the capacities of actors to determine their circumstances and fate. However, the failure to develop alternative conceptualizations of power limits the ability of international relations scholars to understand how global outcomes are produced and how actors are enabled and constrained differentially to determine their fates. If technology and culture are, together, seen as a circle of influence or circle of sustainability, then, cross-cultural competency influences the global distribution of power. Therefore, digital technologies modify space, time, relationships, and types of communication that continue to coexist with the other areas inherent in a culture, and perceptions and understandings of AI are likely to be profoundly shaped by local cultural and social contexts.

Robert Borkowski

The greatest value in interpersonal contacts and in international relations is the exchange of ideas, comparing one's own situation with that of others and sharing experiences, thanks to which one can learn from others and their mistakes. It would be beneficial for the international community if a platform were organized for the exchange of ideas, organized by scientific centers in the form of an international congress on MUAI and the publication of the global MUAI Global Report. Maybe it should be done under the auspices of the United Nations. It should also contribute to the development of appropriate regulations in international law. International cooperation in this area has not developed for three reasons. First, the risks of MUAI are underestimated. Second, in international relations, the threats of MUAI are dominated by particularisms and, until something

spectacular and unfortunate happens, counteracting MUAI will be downplayed. Third, some states use MUAI themselves so they will not be interested in more stringent initiatives against it.

Anna Bychkova

When speaking about the role of international cooperation, it is necessary to determine its subjects. When talking about interstate cooperation, it is important to understand that the technological giants engaged in the development of artificial intelligence algorithms are quasi-states. They adopt generally binding rules of behavior for all users (legal norms of sorts), the violation of which incurs a penalty implemented by the algorithms (not always objective). Given that the spread of IT is of a cross-border nature, it is logical that states should unite to develop standards, principles, and restrictions aimed, for example, at protecting universal human rights. Such platform could be the UN Commission on Human Rights Council. The obstacles to such cooperation may be the lobbying efforts of IT corporations, which are sponsors of a huge number of NGOs that promote their interests in disguise. There is a need to protect the sovereignty of individual states, whose leaders may see threats in such an association.

Matthew Crosston

International cooperation is almost irrelevant in countering MUAI, as it operates at a sub-level far below where international laws, sanctions, and countermeasures could successfully operate.

Svetlana S. Gorokhova

It is difficult to overestimate the importance of international cooperation in successfully countering the malicious use of artificial intelligence. I believe that it would be most effective to use interaction at the highest level: the state level. However, it should be borne in mind that there are serious obstacles to such cooperation, caused primarily by the fact that the analysis of the state policy directions of the most technically developed countries clearly illustrates their intention to participate in and win the global race of achievements in the field of artificial intelligence. Any race involves, at best, competition, and, at worst, rivalry or even hostility. Of course, this is a significant obstacle to fruitful cooperation.

Pavel Karasev

International cooperation on countering psychological influence is necessary, not only due to the characteristics of the ICT environment (its transboundary nature, globality, and anonymity) but also out of the necessity to develop common approaches, especially taking into account the possible use of AI technologies. In addition, today, significant disagreements remain between individual states and groups of countries even on more general issues of security in the global information space. This makes broad international cooperation on these issues unfeasible. Work at the regional level is more effective. For example, the platforms of the Shanghai Cooperation Organization, BRICS, and Collective Security Treaty Organization have proven themselves to be effective in countering information security threats – at the regular meetings and summits of these associations, information security issues are discussed and a common point of view is developed regarding countering current threats emanating from the ICT environment, including the malicious use of AI technologies.

Alexander Raikov

International cooperation in countering the malicious use of artificial intelligence is crucial. However, meetings that take the typical format of allowing everyone to express their thoughts will not help. Typical meetings are divergent in nature. They generate many ideas, but they do not create synergies. What is needed is a specialized intellectual platform that will ensure the stable and purposeful convergence of the discussion process toward a strong result that will provide adequate opposition.

Marina Reshetnikova

It is the development of international cooperation that, perhaps, will ensure successful opposition to the malicious use of AI. It is of note that the legendary Edward Snowden, who, until recently, was living somewhere in the vastness of Russia, joined this confrontation. In his opinion, the expansion of AI to the extent of violating constitutional rights has gone too far and warrants opposition. It is hard to disagree with him. The first step that needs to be taken is to exert public pressure on government agencies to expel structures like NSO from the AI market, that has been successfully done in Israel.

The main threat to IPS caused by the malicious use of AI is the violation of constitutional rights, namely personal inviolability. This violation gives rise to a global psychosomatic disorder, leading to the destabilization of the world order. The formation of permanent physiological disturbance provokes the feeling of personal insecurity. Furthermore, this can cause not only internal political and ethnic unrest but also large-scale artificial disasters. In this situation, the salvation of humanity depends on the development of protective technologies. It is through them that people will technologically confront “Big Brother.” These technologies will have to allow a people to constantly simulate and model their “alternative personality” with other parameters of geolocation, appearance, and so on This is the only way to resist actors like NSO and their offspring, such as Pegasus.

Vitali Romanovski

International cooperation is of utmost importance in successfully countering MUAI. The establishment of appropriate international norms and standards relating to the application of emerging technologies, such as AI, will eventually have to be discussed at the level of the UN system’s actors, processes, and activities. Among these are the Secretary-General’s High-Level Panel on Digital Cooperation, the Open-Ended Working Group, and the Group of Governmental Experts. In addition, there is relevant ongoing research at the level of the United Nations Institute for Disarmament Research, the United Nations Interregional Crime and Justice Research Institute, and the United Nations Office on Drugs and Crime. However, multilateral institutions’ bureaucratic inertness and the growing distrust between the global powers are among the key obstacles to such cooperation.

Sergey A. Sebekin

International cooperation is necessary to solve any problem that is more or less global or interstate in nature – the same applies to the malicious use of AI for psychological influence. In the future, it will be important to create interstate commissions on the malicious use of AI, various kinds of advisory mechanisms and hotlines. In the near future, it will be important to start considering the issue of MUAI for psychological influence on such platforms as the Shanghai Cooperation

Organization, BRICS, and the United Nations. However, the escalating geopolitical confrontations and domestic political antagonisms in different countries cast doubt upon such a favorable development of events for society.

Pierre-Emmanuel Thomann

International cooperation to counter MUAI remains very limited as geopolitical rivalry between great powers is increasing. Ad hoc coalitions might be more successful than large international organizations.

Marius Vacarelu

The main obstacles are geopolitical interests and internal political competition. Because for many politicians "the ends justify the means", international cooperation will exist only between countries that do not compete for the same territories, resources or geopolitical positions.

6. Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for your country?

Vian Bakir and Andrew McStay

Surreptitious influencing via psychological manipulation on social media is a real threat in the UK. In the 2016 "Brexit" referendum on whether or not to leave the European Union, Cambridge Analytica offered one of the "leave" campaigns data-driven insights on voter segmentation and messaging based on psychographic clustering, persuadability, and partisanship, a sub-category of which was "voter suppression."

Raynel Batista Tellez

Since 2016, Cuba has been moving faster to introduce a digital transformation to public administration. There is a national taskforce for developing AI-driven technologies due to potential security threats from the malicious use of artificial intelligence technologies. The most relevant threats have been identified in order to develop a cybersecurity strategy:

- Automated disinformation campaigns
- Automated hacking
- Robot users or fake people

Robert Borkowski

In my opinion, the greatest threats may be the influencing of moods and social attitudes toward increased conflicts and social tensions, and the increasing political polarization in society. Generating mass hysteria, fear, aggression, and hostility in Polish society is, unfortunately, relatively easy to carry out.

Anna Bychkova

Since Russia is embedded in the global information space, threats to IPS are relevant to our country. One problem is the wide distribution and popularity among developers of open-source solutions. Most new research is published on the Internet, and the degree of detail in the documents is such that you can simply take fragments of programs and create complete solutions based on them. This is one of the most serious problems of using artificial intelligence not for the benefit of society, but for malicious purposes.

Matthew Crosston

Manipulation and exploitation of the generally high level of political ignorance and educational flaws that exist across social strata in American society.

Svetlana S. Gorokhova

In the Russian Federation, in my opinion, the most urgent threat to the information and psychological security of the population is, with the advent of AI technologies, the expanding opportunities to manipulate the minds of all citizens, and principally young people. Blind trust in any (even unverified) information obtained from the Internet combined with a natural distrust of information obtained from official (state) sources can have serious psychological consequences, expressed through the strengthening of legal and social nihilism, the emergence of radical views, the rooting of a depressive perception of surrounding reality, and the formation of the stereotype that “anywhere is better than here.”

Pavel Karasev

The greatest threat is posed by influence operations carried out by states to achieve goals of foreign and domestic policy. This threat is growing against the backdrop of the increasing intensity and number of such operations, enabled by the use of AI technologies. A particular danger is the use of deep content targeting for a more accurate information and psychological influence on certain groups of the population and individuals, aimed at the erosion of traditional Russian spiritual and moral values.

Alexander Raikov

My country is Russia. Russians themselves should not discuss and suggest the most relevant malicious uses of artificial intelligence to our country in an open media space because it gives our foes a chance to make the most dangerous malicious attacks.

Marina Reshetnikova

Russia is at the very beginning of the combat for IPS because AI technologies are only just entering our lives. However, even at this stage, we are faced with specific problems. One of them is gaining momentum in the chatbot attacks.

Vitali Romanovski

The critical threats for Belarus are AI employment for the data-poisoning of strategic ecosystems, such as public transport or power grids, and targeted cognitive operations against the population.

Sergey A. Sebekin

In my opinion, a potential threat to Russia is the use of deepfakes, as well as the inculcation in people's minds of any information that erodes political stability through the use of AI technologies. With the help of chat bots, for example, ideas about society, power and the current political situation can be artificially imposed, which is indicative of the psycho-emotional state of citizens.

Pierre-Emmanuel Thomann

The threat to national sovereignty is the most relevant

Marius Vacarelu

The main threat is AI use by political parties – for example, one of them used Cambridge Analytica in 2016 elections. We must underline too that internal competition is thematically and temporally limitless.

7. Are any measures (political, legal, technical or other) being taken in your country to overcome threats to psychological security caused by the malicious use of artificial intelligence? What are these measures?

Vian Bakir and Andrew McStay

Globally, the strategic communications industry (of which Cambridge Analytica was a part) remains unregulated and opaque, with self-regulation failing to stymie its activities. For instance, the Final Report from the UK Inquiry into Fake News and Disinformation observes that the strategic communications industry is largely self-regulated in the UK and requires regulation to curb bad behavior in the industry. The Inquiry recommended that the UK government consider new regulations to ensure transparency in strategic communications companies, with all campaigns that they work on at home and abroad on public record; that the government revisit the UK Bribery Act to gauge whether it prevents bad behavior abroad; and that the government explore the feasibility of adopting a UK version of the US Foreign Agents and Registration Act, which requires persons acting as political agents of foreign principals to disclose their relationships with the foreign principal, as well as the activities, receipts, and disbursements in support of those activities (DCMS 2019: 83–84). However, at the time of writing, most of the UK (namely, England, Northern Ireland, and Wales) still does not implement even basic recommendations that would help people understand who has sponsored political content online during elections. (New legislation has come into force in Scotland covering digital election campaign material about parties and campaigners. It legislates that both "paid for" and "unpaid" digital election campaign material must be clearly labeled with information about who is promoting it. This legislation applied to all parties and campaigners in the May 2021 Scottish Parliament election).

Raynel Batista Tellez

Cuba is working to develop a national cybersecurity strategy that covers issues in several fields:

- *Legal:* The new Constitution Act elected for Cuban people that includes principles for international relations and a Data Protection Law is soon to be implemented.

- *Political*: The last Congress of the PCC (Cuban Communist Party) identified MUAI threats as national security priorities.

- *Technical*: Several platforms and actions have been implemented to protect national sovereignty, for example the NOVA operating system for PC and mobile devices, the declaration of telecommunications as a public and exclusive State property, and the creation of a national university network to cooperate with and support government strategies (devices, infrastructure, operating systems, educational programs, etc.).

- *Social*: Cuban educational programs are focused on promoting best practices in the development and use of digital technologies with social responsibility from their early stages. Educational programs should make the public aware of these threats.

Robert Borkowski

So far, most of society does not perceive the threat of MUAI. The recent development of 5G technology has aroused much more concern, but only in some social environments. Most of society accepts all technical innovations uncritically. People are happy with all the novelties on the digital market. The state authorities also do not yet have any concerns about MUAI. Only some press publications, think tanks, and human rights foundations mention the new dangers of the dissemination and malicious use of artificial intelligence.

Anna Bychkova

Regulatory legal acts of a strategic nature have been adopted: the Decree of the President of the Russian Federation No. 490 of 10 October 2019, “On the development of artificial intelligence in the Russian Federation” (alongside the “National strategy for the development of artificial intelligence for the period up to 2030”); the Decree of the President of the Russian Federation No. 213 of 12 April 2021, “On approval of the foundations of the state policy of the Russian Federation in the field of international information security”; and the Decree of the President of the Russian Federation No. 400 of 2 July 2021, “On the National Security Strategy of the Russian Federation.”

Matthew Crosston

No legitimate measures beyond public international media shaming.

Svetlana S. Gorokhova

I believe that special measures aimed at overcoming threats to national and international psychological security caused by the malicious use of artificial intelligence do not yet exist in any country. Countering such threats is carried out within the framework of the application of methods and means of ensuring information security in general. Such measures, of course, exist in Russia and are expressed through competent authorities’ control over the information content of open resources, and principally for protecting children from information that can cause them psychological damage. So, as an example, we can cite the federal law “On the protection of children from information that harms their health and development” of 29 December 2010, N 436-FZ, which has been in force in Russia for more than 10 years.

Alexander Raikov

I think these questions are for the special organizations in the field of national security to answer. All that I know of is the creation of the Ethical Code, which will be approved in the AI Forum this year, and the growing scientific studies in advanced AI.

Marina Reshetnikova

In Russia, the state is the main fighter in overcoming national and IPS threats caused by the malicious use of AI. At the government level, legal and technical measures are being taken to overcome these threats. Experimental legal regimes are the most promising and effective tools for creating a special testing procedure and subsequent implementation of AI solutions. They ensure the required level of security, protection of citizens' rights and control by government agencies. An example is the introduction of an experimental legal regime in Moscow, where the maximum number of large IT companies is concentrated.

Vitali Romanovski

National legislation updates, the enhancement of multi-layered cybersecurity measures for critical infrastructure objects, special training programs for the personnel of national security entities, and information exchange at the regional level.

Sergey A. Sebekin

Russia is a state with a population that has a high level of Internet access and good quality higher technical education (with Russian programmers being among the best in the world). At the same time, it is also characterized by serious social problems that have been exacerbated during the COVID-19 pandemic and the activities of criminal and terrorist groups, which gives reason for serious concern about the threats of the malicious use of AI (MUAI) in the information and psychological sphere. The deteriorating international situation and various pressures on the country, in addition to the active use of AI-based high-tech ICTs, seriously increase the risks of MUAI.

Pierre-Emmanuel Thomann

A legal framework has been promoted at the EU level but there has been no real breakthrough so far because of various national opinions. More measures have been taken at the national level in France in some strategic areas like defense.

8. Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for Northeast Asia?

Vian Bakir and Andrew McStay

The use of emotional AI that has unproven accuracy in coercive situations will damage the global reputation of this emerging industry. For instance, exploiting the absence of laws restricting authorities’ access to biometric data on the grounds of national security or public safety, major Chinese technology companies, including Baidu and Alibaba, are experimenting with emotion

recognition. Smaller start-ups have begun to collaborate with academics and local governments to roll out emotion recognition technology without safeguards or public deliberation in surveillance systems in prisons, detention centers, remand facilities, and police checkpoints (to detect violence and suicide among prisoners and suspect populations). A case in point is VibraImage technology (video and emotion analytics developed by the Russian company Elsys Corporation to measure facial and head micro-movements and muscle vibrations to describe and categorize subjects' mental–emotional states). Branded as "Alphaeye" in China, it has been trialed for proof of concept to identify potentially dangerous people in crowds in major Chinese airports; was officially certified for use by Chinese police in 2017; and has been deployed in the Inner Mongolia Border Immigration Office and the train station in Yiwu City, Zhejiang Province. Such developments in China are globally significant in that, over the past few years, Chinese technology companies have fueled an international boom in foreign governments' acquisition of surveillance technology: China's One Belt, One Road initiative enables wide-scale implementation of Huawei's SafeCities policing platforms and Hikvision facial recognition cameras in democracies and autocracies alike.

Raynel Batista Tellez

Based on the developing inter-region race to develop AI, economies' growth perspectives, and power distribution patterns, the principal threats to international psychological security caused by the malicious use of artificial intelligence are:

- Fake news reports using deepfake technology
- Automated influence campaigns
- Automated social engineering advertisement practices
- Automated hacking

Robert Borkowski

The most serious threat, in my opinion, is the manipulation of the mood of the masses using disseminated fake news, arousing moral panic, hostility, and fear, and possibly fueling hostility towards the people of the region's states.

Anna Bychkova

Such threats include: the generation of fake information, especially deepfakes; MUAI in shaping the information agenda (including on social networks); the uncontrolled spread of chatbots; the interception of control over technological systems involving AI; and the introduction of programs that generate text and comments on a given topic.

Matthew Crosston

A real and explicit change in the regional political order.

Svetlana S. Gorokhova

I believe that threats to information and psychological security, by and large, are universal in nature, and can cause equal harm to any person, regardless of their region of residence.

Nguyen Quoc Hung

AI technology can help create a vast volume of virtual information and distribute it rapidly and on a large-scale on social media, which can provoke the effect of social unrest, agitate people, and raise discomfort among them. MUAI can be the reason for the deepening division concerning social issues, as well as the distortion of the information work of the state by foreign propaganda. Thus, there is an increased risk of AI being used to interfere in state affairs.

Alexander Raikov

Series of conflicts developed within Hong Kong and Xinjiang is a source of great threat to psychological security. The conflicts between Northeast Asia's countries and around the region also have only grown lately, and not without the West's participation, spearheaded by the United States. This is detrimental to the political and economic stability not only in the region but also worldwide. It is worth noting the lack of a science-based systematic understanding of the threat of the malicious use of AI in the context of psychological security in all countries of Northeast Asia. China has advanced the most in understanding such threats, but they are principally discussed in the context of military threats.

The countries of Northeast Asia have not fully taken into account the synergistic effect of the systemic use of AI in the psychological sphere. Now, there are many AI tools through which damage can be done in the field of psychological security; for example, deepfake technology, distortion of the information agenda (agenda-setting, the use of chatbots, predictive analytics, cognitive modeling, etc.), and so on. It is necessary to consider the non-formalized cognitive semantics of AI models, their chaotic behavior and non-local effects, and the fluctuations of atomic components of neural network structures of a human brain. This is already the subject of research in the field of strong AI, which may be a new force in the formation of malicious psychological weapons.

Marina Reshetnikova

The main threat to IPS caused by the malicious use of AI is the violation of constitutional rights, namely personal inviolability. This violation can give rise to global psychosomatic disorder, leading to the destabilization of world order. The formation of permanent physiological disturbance provokes feelings of personal insecurity. Furthermore, this can cause not only internal political and ethnic unrest, but also large-scale artificial disasters. In this situation, the salvation of humanity depends on the development of protective technologies. It is through them that people will technologically confront "Big Brother." These technologies will have to allow people to constantly simulate and model their "alternative personality" with parameters like geolocation and appearance. This is the only way to resist actors like NSO and their offspring, such as Pegasus.

Experimental legal regimes are the most promising and effective tools for creating a special testing procedure for and subsequent implementation of AI solutions. They ensure the required level of security, protection of citizens' rights, and control through government agencies. An example of this is the introduction of an experimental legal regime in Moscow, where many large IT companies carry out their activities.

Vitali Romanovski

The critical threats to Northeast Asia are AI employment for the data-poisoning of strategic ecosystems, such as public transport, power grids, and nuclear infrastructure, and targeted cognitive operations against the population, when AI-supported computational propaganda tools—such as

content manipulation, illegal data collection, microtargeting, and deepfakes—can assist malicious actors and their supporters in formulating political agendas and manipulating the population in a state or entire region.

Sergey A. Sebekin

The high level of interstate disagreements in the region and adjacent territories facilitates the malicious use of AI technologies, which may become a dangerous large-scale reality in the not-too-distant future.

Pierre-Emmanuel Thomann

The use of MUAI for geopolitical competition and state-sponsored terrorist threats is the most relevant threat for Northeast Asia.

Marius Vacarelu

The Global competition and wish to reach a better position in the world's hierarchy affects Northeast Asia more than other parts of the world. Such wishes attract more unfriendly behavior from other competitors and a strong pressure on the citizens from their own governments. In such a paradigm, it is much more likely to see a strong internal dissatisfaction because of a national government' actions rather than external actions.

9. How much does the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia today?

	Bakir V. and McStay A.	Batista Tellez R.	Borkowski R.	Bychkova A.	Crosston M.	Gorokhova S.	Hung N. Q.	Raikov A.	Reshetnikova M.	Romanovski V.	Sebekin S.	Thomann P.-E.	Vacarelu M.	Expert from Russia	Expert from Vietnam #2	Expert from Vietnam #3
Strongly						V			V							
Noticeably	V	V		V	V		V	V			V		V	V	V	V
Only slightly			V							V		V				
Not at all																

Vian Bakir and Andrew McStay

On the basis of our answer to Question 8, we would say "noticeably." These technologies are still in their trial stages, but they have certainly been noticed with a great deal of concern by human rights organizations and journalists.

10. How much will the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia by 2030?

	Bakir V. and McStay A.	Batista Tellez R.	Borkowski R.	Bychkova A.	Crosston M.	Gorokhova S.	Hung N. Q.	Raikov A.	Reshetnikova M.	Romanovski V.	Sebekin S.	Thomann P.-E.	Vacarelu M.	Expert from Vietnam #2	Expert from Vietnam #3	Expert from Russia
Strongly		V		V		V	V	V	V							
Noticeably			V		V					V	V	V	V	V	V	V
Only slightly																
Not at all																

Vian Bakir and Andrew McStay

We do not know. The answer to this will depend on how governments react to the trials. If they decide to go ahead with the widespread adoption of this technology, the answer would be "significantly".

11. In which countries of Northeast Asia (no more than three) have the threats to international psychological security caused by the malicious use of artificial intelligence reached the highest level? Why?

Vian Bakir and Andrew McStay

China seems to be at the forefront of developing emotional AI for security reasons, and is a leading exporter of its surveillance technologies.

Raynel Batista Tellez

China, Japan, and South Korea are considered the countries with the fastest AI-driven development in the region. Their economic systems (representing different political systems) are also rising to the worldwide top.

Robert Borkowski

It seems that, based on the most developed economies in terms of digital technologies, and the societies with the highest rates of digital communication device usage, the highest level of MUAI risk is in Japan, South Korea, and China.

Anna Bychkova

1. For the first country, I think either South Korea or Japan.

2. China.

South Korea and Japan are countries with a high level of technological development, but, unlike Russia, they are not vast territories with a multinational population and they do not run the risk of state collapse due to the implementation of extremist threats. The psychological security of the population of these countries is rather threatened by information dependence.

As for China, this country has a type of protection, being the power with the highest level of sovereign Internet (the Golden Shield Project or Great Firewall of China). Spreading a kind of "mental plague" across the world in the form of the TikTok platform, China itself is taking a number of measures aimed at banning the use of gadgets in educational institutions. It is also conducting experiments on its own population by using artificial intelligence to create a "social rating" system.

Matthew Crosston

China (domestically), North Korea (domestically), Japan/South Korea (internationally).

Alexander Raikov

China, Japan, and South Korea. These countries are the most dangerous economic competitors for the United States.

Marina Reshetnikova

South Korea, Japan, and, of course, China are leaders in developing and applying AI technologies in the NEA region and globally.

Japan currently applies these technologies for industry optimization by improving overall cost-effectiveness and maintaining public safety through an extensive CCTV network. However, the Japanese government is trying to participate in the private lives of citizens. With the help of AI algorithms, the matchmaking industry is now actively developing. Whatever the purpose, this is an interference in the lives of citizens and poses a threat to IPS.

China is the world's leading country in the field of AI. In terms of the number of patents and AI talents, the country has bypassed its main competitor, the United States. Nevertheless, government actions such as tracking social activity, location, and health status raise some concerns.

The Chinese authorities deny that such technologies can be used to "control minds," stating that active surveillance is part of the smart city system, allowing for tracking and controlling traffic, as well as responding faster to emergencies.

On the one hand, the people of China agree that a system in which they can be quickly informed of being in the same space as a COVID-19-infected person is sufficiently effective and is beneficial to fighting the epidemic, helping to reduce incidences of the disease. On the other hand, it is an encroachment upon privacy and IPS.

The social credit system continues to lack transparency due to the sheer volume and decentralized storage of information. However, Chinese society cannot get away from it due to cultural and historical peculiarities.

Thus, it can be said that China is the leader in MUAI in the region. The malicious nature of the anti-Chinese activity of US companies in the region, conducted through the extensive use of AI technologies, should also be noted.

Vitali Romanovski

The countries with the highest levels of digital interconnectivity and societal digitalization are the most vulnerable to sophisticated cyber-attacks and cognitive operations. South Korea and Japan are among those countries.

Sergey A. Sebekin

I think China, Japan, and South Korea.

Marius Vacarelu

China (because of internal political regime pressure); South Korea (because of North Korea's international behavior), and Japan (because it is the key-country in Western Pacific security).

12. How well is the public in Northeast Asia aware of the threats to international psychological security caused by the malicious use of artificial intelligence?

Raynel Batista Tellez

The public must be aware of these threats. However, the citizens in Northeast Asia are not aware enough of MUAI threats. Cultures may limit the access to education, information, and economic opportunities. The use of AI to destabilize international relations through targeted, high-tech, informational–psychological tactics carried out on people is a reality.

Robert Borkowski

It is difficult to gauge how well the public is aware of these threats, but given the nature of political systems, the most developed level of awareness of these threats should be found in the more democratic countries of Asia, where there is greater freedom of communication and free media, compared to in autocratic states and closed societies.

Matthew Crosston

Quite aware.

Nguyen Quoc Hung

Not very well aware.

Alexander Raikov

The public in Northeast Asia is not well aware of the threats to international psychological security caused by the malicious use of artificial intelligence because the public does not have a deep awareness of AI itself. The public is getting used people at work being replaced with robots. They may have heard something about the possible dangers of robots, about creating different ethical codes, and so on. But they do not think about the real dangers of AI.

Vitali Romanovski

The World Economic Forum's Global Risks Report 2021 underlines that digital power concentration, digital inequality, and cybersecurity failure will be among the most likely risks in the next ten years. Moreover, the applications of today's emerging technologies, including AI, can exacerbate regional tensions and bring new considerations to longstanding security challenges. The policymakers in Northeast Asia seem to be aware of the possible risks that the digital divide and disruption of supply chains could bring to economies. However, there is a lack of understanding of the risks that AI-supported cognitive operations could entail.

Sergey A. Sebekin

The societies of Northeast Asia – such as South Korea, Japan, and China – which, in terms of technological development, live in the tomorrow, are currently watching how high technologies are being introduced into daily life. At the same time, they find themselves in the trenches of the new threats produced by these technologies. Therefore, it seems that these societies have a certain, although insufficient, understanding of the psychological and destructive possibilities that MUAI can bring.

Marius Vacarelu

In China, the press mainly presents the positive aspects of AI. In other countries, the public has a more balanced approach to its dangers and benefits.

13. How do you assess the degree of readiness of state bodies of the countries of Northeast Asia to counter threats to international psychological security caused by the malicious use of artificial intelligence?

Vian Bakir and Andrew McStay

A good first step would be the transparent publication of the trials determining how accurately the technologies achieve what they claim to achieve.

Raynel Batista Tellez

Their degree of readiness is only slight. Advancement in digital technologies produces changes in how information is disseminated and how diplomatic communication is conducted. Cyber-diplomacy has been recently recognized as an instrument of international cooperation to neutralize the proliferation of cyber-attacks and sustain the peaceful use of digital technology in the digital age.

However, the countries in Northeast Asia have different interests and applications of Internet norms and cyber-governance. Perceptions and understandings of threats to international psychological security caused by the malicious use of artificial intelligence limit nations' efforts to support multilateral models or common cyber-sovereignty strategies.

United Nations initiatives like the Governmental Group of Experts and the Internet Governance Forum have been established to address issues related to Internet governance and cybersecurity. However, regional efforts are close to succeeding due to cultural ties. Safe artificial intelligence requires cultural intelligence and changes in cultural codes, behaviors, and fields of knowledge to promote a regional response.

Robert Borkowski

Each country's governments in this region have different concerns and enemies. Thus, their actions in domestic politics are very strongly conditioned by the incentives flowing from the immediate international environment. China's policy is aimed at maintaining full control over its society. The policy of Japan and South Korea is more focused on the welfare of societies and maintaining the status quo in the region.

Anna Bychkova

I do not have information about other countries. In Russia, such readiness is expressed in the Decree of the President of the Russian Federation No. 213 dated 12 April 2021, "On approval of the fundamentals of the state policy of the Russian Federation in the field of international information security." An analysis of this decree allows us to assess the degree of readiness of Russia, as one of the countries of Northeast Asia, to counter threats to international psychological security as high.

Matthew Crosston

Degree of readiness = not very ready given the nature of how the threat impacts societal actors.

Svetlana S. Gorokhova

In my opinion, the degree of readiness of the state bodies of the countries of Northeast Asia to counter threats to international psychological security caused by the malicious use of artificial intelligence—like the state bodies of any other region of the world—cannot be too high for a number of reasons, the most important of which is that AI technologies themselves are still at the stage of formation and development, and, so far, there is not enough empirical data to objectively assess possible future harm, which these technologies can cause even if they are used in good faith, not to mention the possible abuse of these technologies for criminal purposes.

Alexander Raikov

Their degree of readiness is low. The threat of a split in the world is growing. The development of international dialogue under these conditions is slowing down and is excluded from most countries' foreign policy agenda and from the activities of international organizations and civil society. AI systems can reveal positive and negative tendencies and various nuances in the international dialogue, which is not always a prerequisite for leaders' control of various states. At present, along with the analysis of possible risks, an active policy is needed to design a future that is

moving in the direction of positive development. For this, an image of the future should be formed as a development goal. In this regard, the role of Russia among the Northeast Asian countries is increasing as this region of the world is capable of proposing a project to create a secure world. And AI systems can help implement this.

Vitali Romanovski

The pandemic has accelerated fragmentation in Northeast Asia. In the digitalization sphere, the bifurcation between the China- and US-dominated blocs is cause for serious concern. A multilateral mechanism that would promote dialogue among states is now greatly needed. Exchange in itself will not eradicate mistrust in the region, but lack of discussion can lead to serious miscalculations. This also applies to the understanding of the risks that can arise from the use of AI technologies by malicious actors.

Sergey A. Sebekin

To date, the degree of readiness is an extremely relative concept. It is enough to look at the dynamics with which high technologies are evolving today, and, with them, threats – both technological and psychological in nature. AI and the possibilities of its use in the near future will be in the trenches of this process of evolution. As for the technological societies of Northeast Asia – such as South Korea, Japan, and China – with their existing systems of government, they are perfectly capable of responding to incoming threats in a timely manner.

Marius Vacarelu

They are better prepared than the European countries and are probably most prepared at the global level.

The Questionnaire for Experts “Malicious Use of Artificial Intelligence and Challenges to Psychological Security”

This questionnaire is a part of the research project “Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia” funded by the Russian Foundation for Basic Research and the Vietnam Academy of Social Sciences, project number 21-514-92001.

Please answer the maximum number of questions at your discretion. Answers to open-ended questions (assuming a detailed answer) should not exceed 700–800 words in total.

The answers to the open-ended questions are expected to be published in English in the collection: "Malicious Use of Artificial Intelligence and Challenges to Psychological Security”. See a similar publication: "Experts on the Malicious Use of Artificial Intelligence: Challenges for Political Stability and International Psychological Security". Report by the International Center for Social and Political Studies and Consulting June 2020. Moscow.

By answering the questions, your consent is automatically granted. You can contact the researchers in charge at the following email addresses:

Prof. Evgeny Pashentsev: icspsc@mail.ru

Prof. Darya Bazarkina: bazarkina-icspsc@yandex.ru

1. What threats to psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for the modern world? Why?

2. How much does the malicious use of artificial intelligence increase the level of threat to international psychological security today?

- Strongly
- Noticeably
- Only slightly
- Not at all

3. How much will the malicious use of artificial intelligence increase the level of threat to international psychological security by 2030?

- Strongly
- Noticeably
- Only slightly
- Not at all

4. What measures (political, legal, technical or other) do you consider to be important to neutralize the threat to international psychological security caused by the malicious use of artificial intelligence?

5. How important is international cooperation in successfully countering the malicious use of artificial intelligence? On what international platforms (and why) is this cooperation the most effective? What are the existing obstacles to this cooperation?

6. Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for your country?

7. Are any measures (political, legal, technical or other) being taken in your country to overcome threats to psychological security caused by the malicious use of artificial intelligence? What are these measures?

8. Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for Northeast Asia?

9. How much does the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia today?

- Strongly
- Noticeably
- Only Slightly
- Not at all

10. How much will the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia by 2030?

- Strongly
- Noticeably
- Only Slightly
- Not at all

11. In which countries of Northeast Asia (no more than three) have the threats to international psychological security caused by the malicious use of artificial intelligence reached the highest level? Why?

12. How well is the public in Northeast Asia aware of the threats to international psychological security caused by the malicious use of artificial intelligence?

13. How do you assess the degree of readiness of state bodies of the countries of Northeast Asia to counter threats to international psychological security caused by the malicious use of artificial intelligence?

Do you agree to the publication of your answers?

- Yes
- No

If you are willing, please indicate your first name, last name, affiliation, email address, your website (if any, email address and website will not be specified in the publication).

Expert Review

Evgeny Pashentsev

The analysis of the survey results, which does not claim comprehensiveness and takes into account the diversity of experts' points of view, focuses on some of the experts' most important conclusions. For the reader's consideration, the analysis presents both the general and prevailing conclusions and the assessments of a minority of experts on a particular issue, with brief comments and a conclusion by the author.

Preparation of a questionnaire for the survey of experts. A questionnaire with thirteen questions was prepared, the first six of which were designed to find out the opinion of experts on the global issues of MUAI and IPS. The seventh question was designed to find out whether any measures (political, legal, technical, or other) are being taken in the expert's homeland to overcome the threats to PS caused by MUAI, and what these measures are. The remaining questions were devoted to clarifying the threats of MUAI against IPS in NEA.

Four questions in the questionnaire are closed-ended and nine are open-ended[2]. Open-ended questions seem to be especially necessary when researching a new problem, where research approaches and terminology have not been established and it is natural to expect a significant difference in initial responses. In the case of the present survey, the open-ended questions were designed to find out which threats to IPS by means of MUAI are the most relevant for the modern world, which counteraction measures exist, and the relevance of these threats to the country represented by the expert, as well as to the NEA countries. The closed-ended questions were designed to find out how the threats of MUAI in the field of IPS are relevant today and how they may be relevant in 2030. These questions required a choice from several options.

Experts were asked to answer the maximum number of questions at their discretion. They were told that answers to open-ended questions (assuming a detailed answer) should not exceed 800 words in total. This word limit for open-ended questions allows the experts to highlight the main aspects of the problem under study, but may reduce the justification for the answer and the description of the factors, causes and consequences of a phenomenon that are less significant for the expert but objectively important for the formation of a holistic view of the various aspects of the survey topic. When analyzing the results, it is necessary to take into account whether each expert had enough time to answer the questions and their different levels of competence in the themes of particular questions. Thus, each expert was given the opportunity to determine how much to write for each question, and, if they were unsure of their opinion, not to give an answer.

Formation of a knowledge base of experts on the issues of the survey and related fields of science and practice

A. Preliminary selection. In organizing the preliminary selection of experts, the author of the study commenced based on the need to attract specialists from countries with different levels of socioeconomic and technological development in order to obtain a sufficiently broad cross-section of the assessment of the problem. Experts from both the leading countries in the field of AI technology and the countries that are mainly AI consumers were invited.

[2] Open-ended questions require a detailed answer and any explanations, whereas closed-ended questions require only "yes" or "no" answers or a choice between several options.

When choosing specific experts, the author considered it necessary to involve, first of all, researchers in the field of political sciences (including political psychology and political communication), as well as the fields of international security and law. This is due to the task of identifying the threats of MUAI to IPS and the resulting threats to political stability and the security of society as a whole. The author took into account the already-demonstrated research interest of experts in the role of AI in social development and, in particular, in the political aspects of MUAI and PS, as well as academic publications by experts in the subject area of the questionnaire and/or research in adjacent areas. As much as possible, the author took the experts' experiences of practical consulting in the field of AI into account. Because there are plans to invite researchers from NEA countries to participate in a separate survey on the issues of MUAI, no emphasis was placed on attracting experts from the region for this survey.

B. Experts who filled out the questionnaire. Of the nineteen experts who filled out the questionnaire, at least nine have peer reviewed publications in the field of MUAI and PS. Most of the publications of the other experts are in related fields (the historical and political aspects of cybersecurity, AI regulation, the spread of AI in the modern world, etc.).

All experts except one hold PhDs; three hold the title of professor. At least nine experts have performed the role of consultant for legislative and executive authorities, reputable international organizations including relevant UN bodies, and/or startups. Four experts are specialists in technical science in the field of AI, and are interested in the sociopolitical aspects and consequences of MUAI. This makes it possible to correlate the analytical assessments of specialists in the fields of technical and social sciences. At least a third of the experts have publications in or adjacent to the subject area of this survey based on their research on NEA as a whole or specific countries in that region. Prior to the invitation to participate in the survey, the author was personally acquainted with and had cooperated academically with seven experts from the nineteen who filled out the questionnaire (and with 21 of the 58 experts who were invited to participate in the survey). Taking into account the still-narrow circle of specialists on the problem under consideration, it was possible to obtain a high level of expert agreement to fill out the questionnaire (19/58≈33%).

Sending questions to experts and the response received. Extramural forms of work with experts, as is known, make it possible to disregard geographical boundaries during an expert survey and reduce the risks of the experts mutually influencing each other, but make the work of expert groups less operational. In this case, the survey was launched in June 2021 and completed in October 2021. The survey materials were published in December 2021. To obtain a certain representativeness of the assessments, the questionnaire was sent to 58 specialists from fourteen countries. Experts from Belarus (1), Belgium (1), Cuba (1), France (1), the United Kingdom (1 questionnaire —presented jointly by two researchers), Poland (1), Romania (1), Russia (7), the USA (1), and Vietnam (3)[3] responded positively and completed the questionnaire.

[3] Fifteen experts from Belarus, Cuba, France, Poland, Romania, Russia, the United Kingdom, the USA, and Vietnam have agreed to have their answers published. Four experts (one from Belgium, one from Russia, and two from Vietnam) did not give such consent. The answers of these four experts are used in the analytical part of this publication, with their answers to closed questions taken into account. These experts are referred to as the "expert from Belgium," "expert from Russia," "expert from Vietnam #2," and "expert from Vietnam #3". During the survey, two completed questionnaires were received from unknown sources, which are not taken into account in this analysis.

The main results obtained

The answers to the *first question* (made up of two interrelated questions)—"What threats to psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for the modern world? Why?"—revealed a wide panorama of capabilities possessed by MUAI to threaten the PS of individuals, groups, nations, and the whole of humankind. The experts also recognized (directly or indirectly) the presence of antisocial actors who are ready to have such an impact. None of the experts blamed MUAI on AI as such.

Unfortunately, in society, including in the leading countries in the development of AI, there are major yet completely unfounded concerns about AI. This may cause the consequent movement of new Luddites (for example, in connection with the further rapid growth of automation and the growth of the quantitative and qualitative capabilities of AI). Remedying this requires not just proactive explanatory work on the part of the scientific community, but also adequate strategic communication (the synchronization of the deeds, words, and images of the state in important and long-term areas of social development, as well as their perception by various target audiences).

There is no fundamental difference in the approaches of experts in the field of social and technical sciences to the definition of the MUAI threat. The experts' responses list both technologies of psychological influence utilized through MUAI and areas of antisocial activity where MUAI is or may be a tool directed against the PS of society (activities of criminal organizations, including terrorists; targeted malicious influence on the results of elections and referendums; etc.). When listing specific MUAI technologies, despite their answers not overlapping, most experts clearly indicated the general manipulative nature of such activities and their complex nature.

On this basis, it is possible to talk about the multi-variance of tasks, methods, fields of application, territorial coverage, and social conditions, as well as MUAI actors. The multi-variance of MUAI, however, is not limited to this list and requires an adequate systemic public response.

Answers to the closed-ended *second question*—"How much does the malicious use of artificial intelligence increase the level of threat to international psychological security today?"—showed that most experts (n = 10, or ≈53%) answered "noticeably," five (≈26%) answered "strongly," and four (≈21%) answered "only slightly." Thus, most of the experts note a significant or strong influence of MUAI on the growth of IPS threats today. Notably, none of the experts denied such influence; the question is on its extent. Here, experts differed in their assessments, which is probably partly due to the limited statistical base of MUAI in the field of PS.

For the closed-ended *third question* about the situation in 2030—"How much will the malicious use of artificial intelligence increase the level of threat to international psychological security by 2030?"—, the assessments change, compared to the answers to the second question, leaning toward the deterioration of the situation: ten (≈53%) experts answered that MUAI will "strongly" raise threats to international IPS and nine (≈47%) answered "noticeably." No one pointed to an insignificant level of such a threat ("only slightly"), let alone its absence. This implies the need to take preventive measures against a negative scenario. But, as indicated in the introduction to this review of expert responses, the worst-case scenario follows not from the prospect of the further development of AI technologies (which open up wide opportunities for social progress for humanity), but from the high probability of deepening the crises of modern society, strengthening the role of antisocial actors. The latter naturally leads to increased MUAI risks, including in the sphere of IPS.

The *fourth question* is "What measures (political, legal, technical or other) do you consider to be important to neutralize the threat to international psychological security caused by the malicious use of artificial intelligence?" The answers to this question generated a large number of specific (and often interrelated) proposals to prevent and neutralize MUAI and minimize its negative

consequences. This seems to be an extremely important result of the survey because a successful response to a complex threat implies complex solutions. In several responses, experts placed justifiable emphasis on the need for people to be educated in order to successfully resist MUAI. Nevertheless, some experts drew attention to the fact that there are objective social and political disagreements that make it difficult to make coordinated decisions at the level of the state authorities of individual countries or within the framework of interstate cooperation.

A clear minority of experts drew attention to the extreme importance of AI in neutralizing MUAI. The fact is that the technical methods of countering MUAI are best known to a narrow circle of specialists with relevant profiles (illustratively, only four technical specialists took part in this survey). Meanwhile, humanitarian experts naturally focus their attention on the political and legal aspects of countering MUAI, in which they are competent. This once again confirms the need for an interdisciplinary approach to assessing MUAI threats. Conceivably, because of the future growth in the practice of the effective use of AI in order to neutralize MUAI, it will be better known and understood, albeit in a simplified form, by humanitarians. What remained outside the scope of discussion when answering this question were approaches based on the possibility of the successful, socially oriented development of AI and the prevention of MUAI as a result of the progressive transformation of the system of social relations, the more complete integration of AI with humans, and other revolutionary solutions that can, under certain conditions, become breakthroughs for human progress.

The *fifth question* is “How important is international co-operation in successfully countering the malicious use of artificial intelligence? On what international platforms (and why) is this cooperation the most effective? What are the existing obstacles to this co-operation?” The majority of experts support the idea of such cooperation, although they noted serious difficulties in its implementation and insufficient effectiveness; in other words, their responses reflected the structure “International cooperation is very important, but…” The two points of view that differed from this come from more pessimistic assessments. The first is expressed by Matthew Crosston, who wrote, “International cooperation is almost irrelevant in countering MUAI, as it operates at a sub-level far below where international laws, sanctions, and countermeasures could successfully operate.” The second point of view is reflected in the very close positions of the two experts Marius Vacarelu and Pierre-Emmanuel Thomann: “…International cooperation will exist only between countries that do not compete for the same territories, resources, or geo-political positions” (Vacarelu) and “Ad hoc coalitions might be more successful than large international organizations” (Thomann).

The difference between the approaches is sometimes conditional, which was quite clearly demonstrated by the approach of an expert from Belgium: “While political cooperation on the matter is unavoidable in order to take effective countermeasures, … unfortunately, such initiatives are blocked by geopolitical tensions and political interests.”

Because MUAI is mainly developing in the arena of the Internet today, unless it is stopped, it will be possible to suspend MUAI with further quantitative and qualitative growth in the absence of effective international cooperation only by leaving the single global virtual space. A country choosing to retreat from the Internet would broadly protect its population from external psychological attacks, but it would greatly limit its residents in terms of free communication and the use of the achievements of modern civilization. It would be even worse if, due to the increase in global conflicts, the Internet as a unified information space ceased to exist. This would not be the best basis for the development of mutual understanding and interaction between peoples. The losses for each nation would be enormous. Understanding this is what can stop states from taking extreme measures. However, if the threats from MUAI in the psychological field continue to grow and not only connect with traditional Internet propaganda (which is already happening), but, worse, synchronize with

larger and more dangerous cyber-attacks against critical infrastructure, the position of the public and nation-states regarding the Internet may undergo radical changes.

The *sixth question* is “Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for your country?” On this topic, experts are united in their view that psychological manipulation is the main threat. They highlighted different spheres and forms of this manipulation using AI. Some experts focused on specific MUAI technologies that allow (or will allow) for the successful manipulation of people. Another focus was on the qualitative characteristics of the public environment that makes (or can make) MUAI effective and successful. A third focus was on specific areas where the threat is greatest (for example, national security or the use of AI to disrupt administrative management) or on possible MUAI actors (for example, political parties). All of the above points—and many more—are important, indicating that the issue undoubtedly requires an interdisciplinary approach and international cooperation, even if it is limited.

The concentration of society on socially oriented tasks of sustainable development seems to be a means of limiting (an elusive but realistic goal), if not of excluding (an utopian task for a foreseeable future) the activity and very existence of antisocial actors in general and in the field of AI in particular. It is hardly possible to systematically and effectively restrict the activities of antisocial actors exclusively in the field of AI usage without limiting them in the basic economic, social, and political spheres. Antisocial actors are increasingly seeing the growing capabilities of AI in financial, political, and military terms. Hopes that they will voluntarily give up an important tool to strengthen their positions in society are largely unfounded. Similar illusions that the Internet and, later, social networks would become exceptionally positive phenomena of technical and social progress existed at the beginning of the 21st century, but this did not turn out to be the case. The positive role of AI in the transformation of society is an important aspect of socially oriented strategic communication. Conversely, negative scenarios of social and geopolitical development can lead antisocial actors to trigger potentially disastrous MUAI-led outcomes.

The *seventh question* is “Are any measures (political, legal, technical or other) being taken in your country to overcome threats to psychological security caused by the malicious use of artificial intelligence? What are these measures?” The approaches mentioned by the experts vary from a denial of the adoption of such measures to an enumeration of specific political, legal, social, and technical decisions made in a particular country. This indicates great differences both in the states’ approaches to the adoption of such measures and in the experts’ assessments of their breadth and effectiveness. It is crucial to consider the following three circumstances: the insufficient theoretical elaboration on the MUAI and IPS problem, the different degrees of application of various AI technologies for malicious purposes, and the objective limitations of publicly available data. The large variation in estimates may be due, among other factors, to the greater or lesser restrictions on the dissemination of such information in different countries.

The *eighth question* is “Which of the threats to international psychological security caused by the malicious use of artificial intelligence do you consider the most relevant for Northeast Asia?” There is also a large variation in estimates for this question. Some experts found their responses on the fact that threats to IPS, by and large, are universal in nature, whereas others pay attention to the severity of internal political conflicts, interstate disagreements in the region, and the clash of geopolitical interests in NEA, which is a factor that stimulates and facilitates MUAI. These two approaches do not contradict but complement each other: today, the region uses basically the same MUAI technologies as the modern world as a whole. However, the intensity of MUAI may increase, both due to interstate conflicts and due to the greater prevalence of certain technologies in the region. For example, the abuse of video games, where AI is actively used, leads to an increase in

painful attachment to them, and, according to some sources, half of porn deepfakes originate from South Korea. Moreover, if AI continues to develop rapidly, and if conflicts increase, China and, to a lesser extent, Japan and South Korea, aiming to become world leaders in AI by 2030, will be able to become fields for testing local "innovative developments" of MUAI, which will then penetrate into other countries.

Answers to the closed-ended *ninth question*—"How much does the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia today?"—showed that most experts (n = 12, or ≈63%) answered "noticeably," two (≈10%) answered "strongly," three (≈16%) answered "only slightly," and two (≈10%) did not give an answer. None of the experts denied such influence; the question is on its extent.

Regarding the closed-ended *tenth question* about the situation in 2030—"How much will the malicious use of artificial intelligence increase the level of threat to international psychological security in Northeast Asia by 2030?"—, the estimates change in comparison with assessments of the current situation in the answers to the ninth question, tending towards the deterioration of the situation: six (≈32%) experts believe that, already, MUAI will "strongly" increase threats to IPS, nine (≈47%) believe it "noticeably" will, and two (≈10%) noted that "the answer … will depend on how governments react to the trials. If they decide to go ahead with the widespread adoption of this technology, the answer would be "significantly."" None of the experts pointed to an insignificant level of such a threat, let alone its absence. Two experts (≈10%) did not answer.

The answers to the *eleventh question* (made up of two interrelated questions)—"In which countries of Northeast Asia (no more than three) have the threats to international psychological security caused by the malicious use of artificial intelligence reached the highest level? Why?"—produced the following results. China is the clear leader: six experts placed it first in terms of threat level, and two experts placed it second. Japan was ranked first by one expert, second by four experts, and third by three experts. South Korea was ranked first by one expert, second by four, and third by two. Matthew Crosston went beyond the proposed classification by giving a detail answer of interest: "China (domestically), North Korea (domestically), Japan/South Korea (internationally)." When ranking NEA countries by the level of threat to IPS, the experts took into account, among other factors, the level of economic development, the development and implementation of AI technologies, the severity of the MUAI problem under different political systems, and related internal and external conflicts.

The *twelfth question* is "How well is the public in Northeast Asia aware of the threats to international psychological security caused by the malicious use of artificial intelligence?" The most common answer (expressed by seven experts) was that the public in the region is not well aware ("not aware enough of MUAI threats," "not very well aware," "the public … is not well aware," etc.). Four responses provide a positive assessment of public awareness, although the perceived degree of such awareness sometimes differs by country.

The *thirteenth question* is "How do you assess the degree of readiness of state bodies of the countries of Northeast Asia to counter threats to international psychological security caused by the malicious use of artificial intelligence?" The experts' votes were divided approximately equally. A number of experts determined the readiness of government agencies to counter the threats of IPS through MUAI as being of very different degrees, depending on the level of technological development, the degree of openness/closeness of the political system, and other factors.

Limitations

It was only possible to attract experts from Cuba, EU countries, Russia, the USA, and Vietnam to the study, which does not allow for an accurate idea of MUAI and PS in other regions and, to a certain extent, reduces the accuracy of the assessment of the problem in the context of the modern world.

The circle of experts on the integrated assessment of the MUAI and IPS situation is still being formed, as this problem has become acute only in recent years. This happened both because of the rapid development of several AI technologies and MUAI practices, and because of the global crisis covering all major spheres of public life, in particular the sphere of international relations, where acute psychological warfare is taking place as AI usage grows.

The questions allow for some of the most important aspects of the threat of MUAI to IPS to be revealed, but are limited in number, which does not allow for a comprehensive analysis of the problem. This would require the formation of not only more detailed questionnaires and the expansion of the geographical location of the experts involved, but also, and above all, a combination of expert surveys that use other methods of academic research.

There are many opportunities for the development of the methodology of this study, including methods based on the use of AI in expert surveys.

Conclusion

This project demonstrates limited but real opportunities for international cooperation in a new, very important, and yet extremely problematic area of interdisciplinary research that is taking its first steps: MUAI and IPS. The high degree of readiness of specialists to take part in the survey and, in general, the comprehensiveness and professional competence of the answers received are highly encouraging. During the survey, experts expressed coinciding, significantly different, and even mutually exclusive points of view, which is understandable given the novelty and particularity of the issues being discussed.

The political science aspect of the problem is especially important in the context of the aggravation of the psychological warfare between state and non-state actors against the backdrop of acute economic and sociopolitical conflicts in the modern world.

The consideration of the regional aspect of the MUAI and IPS problem using the example of NEA provided a valuable cross-section of data on the nature and dynamics of the formation of a new type of threat. The leading NEA countries, due to their high level of development and application of AI, face serious problems in some areas of MUAI (malicious use of deepfakes, emerging negative aspects of computer gaming with the rising use of AI, etc.). Their analysis can be useful and applicable to developing countermeasures not only in the countries in the region, but also far beyond its borders.

Overcoming the global crisis requires a clear understanding of its causes, driving forces, and consequences, as well as a broad, public discussion about ways to overcome it. We need to act strategically, and, for this, we need clarity of strategic thought.

People can be disoriented by a combination of skillful propaganda in conditions of information hunger, the prohibition of alternative information channels, and open violence. Fascism and many other obvious forms of dictatorship were based on these components. However, in an open dictatorship, people feel encouraged to search for a social alternative and fight for it despite the threat of repression and death. The historical doom of such dictatorships has been proven in practice, but does not prevent dangerous relapses of the dark pages of history in a new environment. An implicit dictatorship, hidden from the public consciousness, is more dangerous under its conditions of

the skillful manipulation of a multitude of half-truths that place people in the kingdom of crooked mirrors, where, in fact, they are deprived of choice. Yet, although it is not easy, there is a chance to find a way out of the labyrinth of lies of traditional forms of propaganda.

The growing MUAI by antisocial actors poses a serious threat at the international level, further narrowing the ability of people to understand the current situation when it is extremely necessary for them and all future generations. Today, we are closer than ever to the end of human history. Further progress of AI technologies, MUAI and its large-scale use as a psychological weapon may lead, at the least, to an even greater delay in an adequate public response to the dangerous behaviors of antisocial actors. On a global scale, this will facilitate the formation of conditions for various kinds of fabricated and social disasters, including a Third World War. With a qualitatively perfect MUAI, the matrices of thinking and behaviors that are hostile to people in the near future may become practically insurmountable at the level of public consciousness and political practice. It may become an important element in the formation of techno-fascism, with the subsequent almost conflict-free liquidation of the population because of continued automation, robotization of production processes, and widespread incorporation of AI in the interests not of society but of a narrow oligarchic elite.

MUAI threats to IPS should be considered at three levels.

At the *first level*, MUAI threats are associated with a deliberately distorted interpretation of the circumstances and consequences of AI development for the benefit of antisocial groups, and the spread of the false negative image of AI can slow its incorporation and cause sociopolitical tensions. At the same time, the deliberately inflated expectations for the use of AI that are transmitted to society through various channels are no less dangerous: they, for example, can be effectively used to disorient the general public, interested commercial and non-profit structures, and public authorities, and, ultimately, can also turn into disappointments, wrong decisions, and social and political conflicts.

Where MUAI is aimed primarily not at managing target audiences in the psychological sphere but at committing other malicious actions (for example, destroying critical infrastructure), we can talk about the *second level* of the effect of MUAI on IPS. Such attacks can have a great psychological effect due to the damage caused.

MUAI designed primarily to cause psychological damage belongs to the third and highest level of threat to IPS. The use of AI in psychological warfare already makes covert perception management campaigns more dangerous. Examples include AI phishing and the use of deepfakes and smart bots in information campaigns for various purposes, such as marring the reputation of an opponent, be it a person, an organization, or even a country. At some point, this can allow aggressive actors to control the public consciousness and eventually lead to the destabilization of the international situation.

PS threats posed by MUAI can exist in both pure forms (for example, the misinformation of citizens about the nature of AI without its malicious use) and combined forms. For example, overestimating the effectiveness of current AI technologies, forming expectations of certain highly favorable results of their implementation or any products based on them would be a first-level attack with a communicative effect (for example, a speculative boom in the stock market). However, if the perpetrators were to accompany their actions with physical attacks on critical infrastructure or people and a widespread, malicious psychological campaign using different AI tools, the threat would become a combined attack.

Manipulation of broad segments of the population in targeted perception management campaigns is particularly dangerous, as many experts pointed out in their answers with different nuances in their wording.

Given the extreme tensions of today's world, it seems that attention should be paid to the first level of threat to IPS through MUAI in combination with subsequent levels because this is where very disturbing phenomena and trends are observed. This may become the subject of a future survey and the subject of a broad discussion. In the present publication, the author pays attention to only some of these phenomena and trends.

The largest companies in the field of high technology actively use AI according to their narrow corporate interests, which rather often go against the interests of society. It is clear that companies with access to large amounts of data to power AI models are leading AI development. Key groups within AI include GAFAM—Google (Alphabet), Apple, Facebook (Meta), Amazon, and Microsoft, also known as the Big Five, which is a name given to the five largest, most dominant, and most prestigious companies in the information technology industry of the United States, BAT (the BAT is the Chinese name given to the leading internet and software companies in China: Baidu, Alibaba, and Tencent), early-mover IBM, and hardware giants Intel and NVIDIA (Lee, 2021).

It is hardly accidental that among the individuals with the ten largest fortunes in the world, six represent Amazon (1), Microsoft (2), Google (2), and Facebook (1) (Forbes, 2021). $7.5 trillion: that was the combined market capitalization of GAFAM at the end of 2020, according to an analysis by the Wall Street Journal. At the end of 2019, these firms' combined market capitalization was $4.9 trillion, which means they increased in value by 52% in a single year. As of November 12, 2021, the capitalization of these companies has grown by another $2.5 trillion and reached approximately $10 trillion (Statista, 2021a). That is nearly a quarter of the combined $41.8 trillion market capitalization of all companies in the S&P 500 (La Monica, 2021). It is appropriate to recall that the United States' nominal GDP in 2020 was around $21 trillion. Japan, the world's third-largest economy, had a GDP of about $5 trillion, and Russia had one of only about $1.5 trillion.

However, the positioning and consolidation of these individuals in the top ranks in terms of assets coincided with the degradation of the reputational capital of most of their associated companies.

The 2021 Edelman Trust Barometer, an annual survey conducted by the global public relations firm Edelman for more than two decades, shows this clearly. Although technology has long been the most trusted industry sector, trust has plummeted more than in any other sector over the past ten years. In 2012, 77% of survey respondents expressed trust in tech companies to "do what is right." The 2021 research shows that that percentage has dropped to 68%. This percentage decline is three times that of any other industry in the study (Shabir, 2021). Three of the Big Five companies—Google, Amazon, and Microsoft—have dropped in rank year after year in the Global RepTrak 100 rankings. A fourth, Facebook, did not appear in the rankings in 2020-2021. The fifth—Apple—managed a decent improvement. Apple's gain was overshadowed, however, by Amazon plummeting by 50 places, from 42nd in 2020 to 92nd in 2021 (Abdulla, 2021).

Four Big Five CEOs testified before the U.S. House Antitrust, Commercial, and Administrative Law Subcommittee in an antitrust hearing on July 29, 2020. Amazon founder and CEO Jeff Bezos, Facebook founder and CEO Mark Zuckerberg, Apple CEO Tim Cook, and Alphabet and Google CEO Sundar Pichai defended their companies against accusations of anticompetitive practices (Rev, 2021). Former Facebook product manager Frances Haugen testified before the US Senate on October 5, 2021, that the company's social media platforms "harm children, stoke division and weaken our democracy" (Menczer, 2021), and that Facebook did not use AI technologies ethically. "Right now, Facebook is closing the door on us being able to act. We have a slight window of time to regain people control over AI" (Browne & Shead, 2021). In November 2021, a new bipartisan Senate bill aimed at restricting tech companies' "anticompetitive" acquisition was introduced by Senators Amy

Klobuchar (D-MN) and Tom Cotton (R-AR) and would greatly limit the ability of Big Five companies to acquire other tech companies.

US tech giants GAFAM have been accused in the EU of not paying enough taxes, stifling competition, stealing media content, and threatening democracy by spreading fake news. An EU court in November 2021 rejected a Google appeal against a 2.4-billion euro (2.8-billion dollar) anti-trust fine. Amazon was fined 746 million euros in July 2021 by Luxembourg authorities for flouting the EU's data protection rules. France has also fined Google and Amazon a total of 135 million euros for breaking rules on computer cookies. The European Parliament and member states agreed to force platforms to remove terrorist content, and to do so within one hour. EU rules now also forbid the use of algorithms to spread false information and hate speech, which some major platforms are suspected of doing to, among other efforts, increase advertising revenue (AFP, 2021).

The Chinese government strengthened control measures over the country's technology companies in 2021. More than one trillion dollars were wiped off the collective market capitalization of some of the Chinese largest Internet groups, such as Tencent, a gaming and social-media giant, and Alibaba, China's e-commerce powerhouse (He, 2021). The Cyberspace Administration of China (CAC) reported on March 18, 2021, that representatives of CAC and the Ministry of State Security met with employees of Alibaba Group, Tencent, ByteDance, and other companies specializing in information technology to discuss potential problems with deepfakes. China's state control authorities have instructed local CAC offices and state security agencies to strengthen the security assessment of voice software and deepfakes. To do this, it is proposed to take into account the "Law on Network Security," "regulations on the Assessment of the Security of Information Services on the Internet" (**中共中央网**络安全和信息化委员会办公室 (Cyberspace administration of China), 2021), and other laws and regulations. In August 2021, the Cyberspace Administration of China announced draft regulations for Internet recommendation algorithms. It wants to halt algorithms that encourage users to spend large amounts of money or spend money in ways that "may disrupt public order" (Frater, 2021). In September 2021, regulators told gaming companies, including Tencent, that they should stop focusing on profits and instead concentrate on reducing adolescents' addiction to playing. The short-video industry, dominated by companies such as ByteDance, Kuaishou, and Bilibili, may receive similar treatment (**中共中央网**络安全和信息化委员会办公室 (Cyberspace administration of China), 2021).

In Russia, the state communications regulator Roskomnadzor demanded in November 2021 that thirteen foreign (mostly U.S.-based) technology companies be officially represented on Russian soil by the end of 2021 or face possible restrictions or outright bans. This year, Russia fined foreign social media giants Google, Facebook, Twitter, and TikTok and messaging app Telegram for failing to delete content it deems illegal. Apple, which Russia has targeted for the alleged abuse of its dominant position in the mobile applications market, was also on the list. Roskomnadzor said firms that violate the legislation could face advertising, data collection, and money transfer restrictions or outright bans (Marrow & Stolyarov, 2021).

In Russia, the authorities want to force major foreign streaming services (like YouTube) to pay local operators for using their traffic. This proposal was put forward in 2021 by the Ministry of Digital Development, Communications and Mass Media of the Russian Federation. If the proposal is accepted, it should serve as a new barrier against foreign Internet giants. The Ministry explained its position by stating that operators are put in a difficult position due to the overloading of Russian networks. Thus, the authorities can force foreign companies to finance, to some extent, the development of communications infrastructure in Russia (Tsargrad, 2021). The Russian authorities have been trying to regulate foreign Internet services as one of their instruments of influence for at least the past six years, but they took decisive action only at the end of 2020. Then, a number of laws

were passed that toughened the responsibility of companies. These are, for example, amendments to the law “On measures to influence persons involved in violations of fundamental human rights and freedoms, the rights and freedoms of citizens of the Russian Federation,” according to which Roskomnadzor received the right (by decision of the Prosecutor General’s Office) to slow down the traffic of services not only because of security threats, but also if they restrict access to “socially important information”. In addition, on February 1, 2021, the law came into force, according to which social networks are obliged to identify and block illegal content. So far, it does not involve sanctions. The authorities first used the technology of slowing down traffic within the framework of the law “on the sovereign Internet” in March 2021, recognizing Twitter as a security threat for providing access to information prohibited in Russia (Shestopyorov & Lebedeva, 2021).

Thus, restrictive measures are being taken in many countries and, in the context of PS and MUAI, these measures are related to the use of AI technologies, which indicates an unintentional or intentional disregard for the interests of the public on behalf of the biggest high-tech companies. Are these measures in different countries sufficient and balanced enough to prevent negative antisocial phenomena associated with the new technical (primarily based on AI technologies) and financial capabilities of high-tech companies? Any answer to this question would be premature.

Elon Musk’s electric car giant Tesla can rightfully be included in the arena of Big Tech. It recently passed the $1 trillion mark in market capitalization and has since surged to about $1.25 trillion. The fortune of Musk, the richest man on the planet, reached a record high of $305 billion on November 22 (Forbes, 2021). (At the end of 2019, it was less than $30 billion.) Musk’s total net worth is now greater than the market value of Exxon Mobil Corporation or Nike Inc. The basis of Tesla’s success is the widespread development and application of AI. As stated on Tesla’s official site, *“*We develop and deploy autonomy at scale in vehicles, robots and more. We believe that an approach based on advanced AI for vision and planning, supported by efficient use of inference hardware, is the only way to achieve a general solution for full self-driving and beyond*”* (Tesla, 2021). However, to what extent will AI technologies in Musk’s projects (Tesla, Neuralink, etc.), as well as in numerous Big Tech projects in general, primarily serve society rather than financial elites? To what extent will AI technologies (for example, in medicine) be publicly available? Perhaps the boom of Big Tech based on AI and other technologies will be successful and sustainable, but will it unite rather than destroy humanity if decisions continue to be made within the framework of a modern socioeconomic model without a clear vision of the goals and means of our movement toward a more progressive dynamic and socially oriented model? So far, the examination of the effect of the coronavirus pandemic on mankind suggests that there is no alternative to the transition to new technologies. However, it is extremely socially unbalanced and gives a completely disproportionate amount of preferential material dividends to an insignificant minority, and, to the overwhelming majority, mostly limited to promises of a better life in the future.

Furthermore, to what extent is the rapid growth of Big Tech as a whole not inflating a huge financial bubble on inflated expectations from highly promising and extremely important technologies for humanity? In the near future, a crushing financial and economic crisis can follow, which will further enrich the few and ruin hundreds if not billions of people around the world. Big Tech Companies Amass Property Holdings During Covid-19 Pandemic, they “…are sitting on record piles of cash. They are getting paid next to nothing for holding it, and they are running out of ways to spend it” (Bangkok Post, 2021). For example, Alphabet Inc., Google’s parent company, held $135.9 billion in cash, cash equivalents and short-term investments as of the second quarter of 2021—more than any other publicly traded company, not counting financial and real-estate firms, according to S&P Global. Alphabet is now one of the biggest real-estate owners in New York City and the U.S. It held $49.7 billion worth of land and buildings as of 2020, up from $5.2 billion in 2011. Amazon, which owns many warehouses, held $57.3 billion worth of land and buildings—more than any other U.S.

public company except Walmart (Bangkok Post, 2021). The European Central Bank warned in November 2021 of bubbles in property and financial markets (The Liberty Beacon, 2021). According to Willem H. Buiter, an adjunct professor of international and public affairs at Columbia University, “the next financial crisis is fast approaching” (Buiter, 2021).

Unfortunately, if this negative crisis scenario turns out to be true, is not the rapid deterioration of the international situation the natural evolution, fraught with, if not a world war, a very large military provocation? Will it not become a necessary trigger for the collapse of the markets? The “culprit/guilty party” will, of course, be found where it is necessary, since the world’s main information resources are, dangerously, controlled by and subordinated to the interests of global corporate structures. Further, MUAI already exists on a global scale as a game based on inflated expectations of benefits from the incorporation of AI. This game is played through a versatile psychological impact on target audiences who are particularly susceptible and vulnerable to perception management in a crisis situation. In whose hands are the most advanced tools of global psychological influence and whose financial interests are at stake? Not only is there enough objective data to answer this question; there is an abundance. Therefore, the possible and specific scenarios of combined, targeted impact—not only with the help of specific AI technologies, but also of the very perception of AI—on the public consciousness for the purpose of speculative enrichment and the destabilization of public order requires the most serious attention and comprehensive study by specialists from different countries and with different scientific specializations.

Of course, one cannot blame only high-technology information companies for the antisocial use of AI: one of the main reasons for increased MUAI is the increasingly uncontrolled behavior of large businesses, which has only become more obvious during the crisis. Just 1,275 wealthy families paid $9.3 billion in estate tax to the U.S. Treasury in 2020. As recently as 2018, the IRS collected more than $20 billion from nearly 5,500 families. The dramatic decline—to the point where the tax is paid by 0.04% of dying Americans—is largely the result of the tax overhaul enacted by Republicans in 2017, which doubled the amount the wealthy can pass to heirs without triggering the levy (Bloomberg News, 2021). Between 2010 and 2020, the U.S. and its allies accounted for only 5% of worldwide increases in democracy. But a staggering 36% of all backsliding occurred in U.S.-aligned countries. On average, allied countries saw the quality of their democracies decline by nearly double the rate of non-allies, according to V-Dem’s figures (Fisher, 2021). According to a new national survey organized in the U.S. by the nonprofit Public Religion Research Institute, nearly one in five (18%) of overall respondents said they agreed with the statement: “Because things have gotten so far off track, true American patriots may have to resort to violence in order to save our country” (Dickson, 2021).

In a wide-ranging interview with UN News in September 2021, UN Secretary-General António Guterres called on world leaders to “wake up,” make an immediate course correction at home and abroad, and unite. “The institutions we have, have no teeth. And sometimes, even when they have teeth, like in the case of the Security Council, they have not much appetite to bite,” the UN chief said (UN Affairs, 2021). The same month, the UN secretary general warned that the world is “on the edge of an abyss and moving in the wrong direction” in an urgent and sometimes angry address to the world’s leaders at the UN general assembly. “We are seeing an explosion in seizures of powers by force. Military coups are back,” he said. When democracies fail to deliver on the basic needs of their people, Guterres added, “it provides oxygen for easy fixes, silver solutions and conspiracy theories” (Borger, 2021).

Total global military expenditure rose to $1,981 billion last year, an increase of 2.6% in real terms from 2019, according to new data published today by the Stockholm International Peace Research Institute (SIPRI). The five biggest spenders in 2020, which together accounted for 62% of global military expenditure, were the United States, China, India, Russia, and the United Kingdom.

The 2.6% increase in world military spending came in a year when global GDP shrank by 3.3% (Statista, 2021b).

Under the conditions of continuing acute economic problems, the impoverishment of hundreds of millions of people and the rapid concentration of world wealth in the hands of very few, an arms race, and acute geopolitical contradictions, MUAI against IPS, conducted by the forces of a variety of antisocial actors, can play an extremely negative and dangerous role. This is why countries, especially those with leading scientific and technical potential, can and should cooperate in order to prevent the antisocial actors' use of information technologies based increasingly on new AI capabilities. On November 3, 2021, without calling for a vote, the United Nations General Assembly First Committee adopted a drafted Russian–U.S. resolution on the rules of behavior in cyberspace. The document will be considered by the General Assembly in December (Suciu, 2021). This is a good example of the possibility of such cooperation.

Meanwhile, it is clearly insufficient and ineffective to resist increasingly successful MUAI in a society where the influence of antisocial actors is increasing via separate, unrelated decisions of a political, legal, and technical kind. Under these conditions, the countermeasures are nothing more than a palliative, at best allowing one to gain time, at worst a cover for the systemic deterioration of the situation. It is important to take into account the following factors:

First, MUAI is qualitatively more dangerous for a sick social organism than for a healthy one. We need a socially oriented transformation of society, part of which will be a complex of systemic and effective political, legal, technical, and organizational solutions to prevent MUAI and minimize its negative impact on the public consciousness. This is not about copying models of the past, but forming a progressive model that meets the realities, risks, and opportunities of the 21st century.

Second, increasing investments in science and education in order to develop the capabilities of the main productive force of modern society—people—is an important response to the threats of MUAI in the broad context of the formation of a comprehensively developed responsible citizen of a democratic society rather than a one-dimensional consumer who is convenient to manipulate for selfish purposes. A multidimensional, harmonious, and socially responsible person can protect themselves, their loved ones, and their society more successfully than a one-dimensional consumer. This rule holds for a developed, civil society; for a different society. In an unhealthy society, a clearly expressed civic position often means increased risks for its bearer, including the threat of physical destruction. We know this not only from history, but also from the modern reality of many countries.

Third, there are well-known estimates that indicate that society is becoming more complicated, and that the volume of incoming information is many times greater than the ability of the existing personal, group, and public consciousness to assimilate it and use it adequately in decision-making. This situation increasingly does not meet the needs of further dynamic sustainable development. One of the new, specific mechanisms for solving the problem may be augmented intelligence (also referred to as intelligence amplification, cognitive augmentation, or enhanced intelligence) is a design pattern for a human-centered partnership model of people and AI working together to enhance cognitive performance, including learning, decision-making, and new experiences (Gartner, 2021). Taking into account the growing possibilities of cyborgization (Pro Robots, 2020), a closer (and, in the future, symbiotic) connection between human and machine is also being facilitated, which will increase our capabilities to obtain, process, and verify data, and therefore to resist MUAI.

The author would like to believe that this survey is just a prologue for future joint international research in the field of MUAI and IPS. Such research will not only be designed to solve important scientific problems; its principal practical task will be to help ensure the PS of society. People must have a clear systemic understanding of the surrounding reality to make conscious choices in their lives. AI is a means to take away this choice in the interests of antisocial actors, but, to a greater

extent, it is also a tool for the protection and self-development of the individual and society as a whole. We still have a choice.

References

Abdulla, N. (2021). Only One of Big Tech's Big Five Comes Out Unscathed in RepTrak's 2021 Global Reputation Rankings. Retrieved 28 November 2021, from https://www.trustsignals.com/blog/big-tech-plummets-in-reptrak-100

AFP. (2021). Europe's battle to curb Big Tech. Retrieved 28 November 2021, from https://sg.finance.yahoo.com/news/europes-battle-curb-big-tech-040606211.html

Bangkok Post. (2021). Big Tech Companies Amass Property Holdings During Covid-19 Pandemic. Retrieved 28 November 2021, from https://www.bangkokpost.com/business/2189935/big-tech-companies-amass-property-holdings-during-covid-19-pandemic

Bloomberg News. (2021). Ultra-rich skip estate tax, sparking 50% drop in IRS revenue. Retrieved 28 November 2021, from https://www.investmentnews.com/ultra-rich-skip-estate-tax-sparking-50-drop-in-irs-revenue-214350

Borger, J. (2021). António Guterres 'sounds the alarm' over global inequalities in UN speech. Retrieved 28 November 2021, from https://www.theguardian.com/world/2021/sep/21/antonio-guterres-united-nations-unga-speech

Browne, R., & Shead, S. (2021). 'Facebook is closing the door on us being able to act,' whistleblower says in UK hearing. Retrieved 28 November 2021, from https://www.cnbc.com/2021/10/25/facebook-whistleblower-frances-haugen-testifies-in-uk-parliament.html

Buiter, W. (2021). The next financial crisis is fast approaching. Retrieved 28 November 2021, from https://www.marketwatch.com/story/the-next-financial-crisis-is-fast-approaching-11633447555

Dickson, C. (2021). 'Alarming finding': 30 percent of Republicans say violence may be needed to save U.S., poll shows. Retrieved 28 November 2021, from https://news.yahoo.com/prri-poll-republicans-violence-040144322.html?fr=sycsrp_catchall

Fisher, M. (2021). U.S. Allies Drive Much of World's Democratic Decline, Data Shows. Retrieved 28 November 2021, from https://www.yahoo.com/news/u-allies-drive-much-worlds-194121292.html

Forbes. (2021). The World's Real-Time Billionaires. Retrieved 28 November 2021, from https://www.forbes.com/real-time-billionaires/#1d7a52b83d78

Frater, P. (2021). Celebrities Disappear From Internet As China Moves Against Fan Culture. Retrieved 28 November 2021, from https://variety.com/2021/digital/asia/china-celebrities-disappear-internet-fan-culture-crackdown-1235050381/

Gartner. (2021). Definition of Augmented Intelligence - Gartner Information Technology Glossary. Retrieved 28 November 2021, from https://www.gartner.com/en/information-technology/glossary/augmented-intelligence

He, L. (2021). China's 'unprecedented' crackdown stunned private enterprise. One year on, it may have to cut business some slack. Retrieved 28 November 2021, from https://edition.cnn.com/2021/11/02/tech/china-economy-crackdown-private-companies-intl-hnk/index.html

La Monica, P. (2021). The race to $3 trillion: Big Tech keeps getting bigger. Retrieved 28 November 2021, from https://edition.cnn.com/2021/11/07/investing/stocks-week-ahead/index.html

Lee, G. (2021). Big Tech leads the AI race – but watch out for these six challengers. Retrieved 28 November 2021, from https://www.airport-technology.com/features/big-tech-leads-the-ai-race-but-watch-out-for-these-six-challenger-companies/

Marrow, A., & Stolyarov, G. (2021). Moscow tells 13 mostly U.S. tech firms they must set up in Russia by 2022. Retrieved 28 November 2021, from https://finance.yahoo.com/news/moscow-says-13-foreign-tech-122138251.html

Menczer, F. (2021). Facebook whistleblower Frances Haugen testified that the company's algorithms are dangerous – here's how they can manipulate you. Retrieved 28 November 2021, from https://news.yahoo.com/facebook-whistleblower-frances-haugen-testified-122343232.html?fr=sycsrp_catchall

Pro Robots. (2020). Cyborg Revolution: Latest Technologies and TOP of Real Cyborgs. Retrieved 28 November 2021, from https://www.youtube.com/watch?v=TyWohWpozp0

Rev. (2021). Big Tech Antitrust Hearing Full Transcript July 29. Retrieved 28 November 2021, from https://www.rev.com/blog/transcripts/big-tech-antitrust-hearing-full-transcript-july-29

Shabir, S. (2021). Four Steps To Winning Over An Increasingly Skeptical Public. Retrieved 28 November 2021, from https://www.technologytimes.pk/2021/02/02/four-steps-to-winning-over-an-increasingly-skeptical-public/

Shestopyorov, D., & Lebedeva, V. (2021). Mimo zamedlennogo dejstviya. Vlasti ishchut novye rychagi davleniya na zarubezhnyj IT-biznes (The authorities are looking for new levers of pressure on foreign IT business). Retrieved 28 November 2021, from https://www.kommersant.ru/doc/4783593

Statista. (2021a). S&P 500: largest companies by market cap 2021. Retrieved 28 November 2021, from https://www.statista.com/statistics/1181188/sandp500-largest-companies-market-cap/

Statista. (2021b). Growth of the global gross domestic product (GDP) from 2016 to 2026. Retrieved 28 November 2021, from https://www.statista.com/statistics/273951/growth-of-the-global-gross-domestic-product-gdp/#:~:text=In%202020%2C%20the%20global%20economy%20fell%20by%20about,by%20a%20country%20in%20a%20certain%20time%20period

Suciu, P. (2021). Is a U.S-Russian Cyber Alliance in the Works?. Retrieved 28 November 2021, from https://nationalinterest.org/blog/buzz/us-russian-cyber-alliance-works-196194

Tesla. (2021). Artificial Intelligence & Autopilot. Retrieved 28 November 2021, from https://www.tesla.com/AI

The Liberty Beacon. (2021). Whistleblowers Torpedo Big Tech And Big Pharma: Who's Next?. Retrieved 28 November 2021, from https://www.thelibertybeacon.com/whistleblowers-torpedo-big-tech-and-big-pharma-whos-next/

Tsargrad. (2021). V Rossii obsuzhdayut platu za YouTube: Novyj zaslon dlya internet-gigantov (YouTube Payments Discussed in Russia: New Barrier for Internet Giants). Retrieved 28 November 2021, from https://tsargrad.tv/news/v-rossii-obsuzhdajut-platu-za-youtube-novyj-zaslon-dlja-internet-gigantov_451654

UN Affairs. (2021). UN chief's message to world leaders: 'Wake up, change course, unite'. Retrieved 28 November 2021, from https://news.un.org/en/story/2021/09/1100152

中共中央网络安全和信息化委员会办公室 (Cyberspace administration of China). (2021). **国家互**联网信息办公室、公安部加强对语音社交软件和涉深度伪造技术的互联网新技术新应用安全评估. Retrieved 28 November 2021, from http://www.cac.gov.cn/2021-03/18/c_1617648089558637.htm

About the Experts

Vian BAKIR

Prof. Vian Bakir is Professor in Journalism and Political Communication at Bangor University, UK. She is an expert on the impact of the digital age on strategic political communication, dataveillance, and disinformation. She has advised UK national research councils (Engineering & Physical Sciences Research Council (EPSRC), Arts & Humanities Research Council (AHRC) on their major investments into digital citizenship, AI, ethics, and governance; and the European Commission on its Horizon 2020 work program on digital disinformation. She has reviewed >150 grant applications on technology and society for such councils. She has been awarded multiple UK research council grants on data governance and transparency (from the Economic & Social Research Council, AHRC, EPSRC, Innovate UK, and Arts Councils). Her books include: *Intelligence Elites and Public Accountability: Relationships of Influence with Civil Society* (2018); *Torture, Intelligence and Sousveillance in the War on Terror (2016); Sousveillance, Media and Strategic Political Communication: Iraq, USA, UK (2010) and Communication in the Age of Suspicion (2007).* She has recently advised the UK All Party Parliamentary Group on AI (2020), the UK All Party Parliamentary Group on Electoral Transparency (2019-20), the UK Parliament's Fake News Inquiry (2017-19), and the Parliament of Victoria (Australia) Electoral Matters Committee on social media and political campaigning. She has recently worked with the UK's National Union of Journalists to prepare guidance for journalists seeking to avoid surveillance by the security state; and she has advised business on public attitudes towards emergent technology.

Raynel BATISTA TELLEZ

Principal Professor and Chair of Lectures at National Cybersecurity Engineering Program, Universidad de las Ciencias Informáticas (UCI), Cuba. He is a member of UCI Artificial Intelligence Lab and the National Association of Pattern Recognition. He was Chief Editor of Futuro Publishing House, editorial board member of academic journals and collaborator of *International Sociology journal*. He holds an AMBA Certificate at International Management Institute (IMI), India. He participated in several international projects *Chilenische-Deutsch Jugend Kulturtreffen* (Germany), *Philosophischen Fakultät der Universität Zürich* (Switzerland), *Caracas Social Work* (Venezuela), *Social Network Analysis* (Cuba). His main research areas are Big Data Analytics, Cognitive Automation, Data driven innovation, and social

network and social media analysis. His current PhD research is based on sociocybernetics and anthropology, assuming cross-cultural competency to understand the influence of artificial intelligence on the global distribution of power and regional balance.

Robert BORKOWSKI

Holds a PhD. Associate professor at Andrzej Frycz Modrzewski Krakow University, Faculty of Security Sciences, Krakow, Poland. Scholarship holder of the US Department of State. Member of the Scientific Council of the Center for Research on Terrorism in Warsaw, Poland. He deals with the issues of terrorism and counter-terrorism prevention, as well as religious fundamentalism and psychological aspects of terrorism. Rescue is another area of his research and activity. Author of the books *Postmodern terrorism – a study of political anthropology*, *Foundations of Europe*, *Civilization – Technology – Ecology*, editor of twelve collaborative books. Decorated by the President of Poland with the Medal of Sacrifice and Courage and the Silver Cross of Merit for his contribution to rescue.

Anna BYCHKOVA

Anna Bychkova holds a PhD in Law. Associate Professor. Psychologist. She graduated with a Master's degree in Journalism. Head of the Scientific Research Department of the Irkutsk Institute (branch) All-Russian State University of Justice (Russian Law Academy of the Ministry of Justice of Russia). Expert of the Federal Service of the Russian Federation for Supervision in the Field of Communications, Information Technologies, and Mass Communications. Deputy editor of the *Prologue: Law Journal*. Author and co-author of publications on the problems of using artificial intelligence in countering, predicting, preventing and investigation crime, and also how crime is adopting technologies with elements of artificial intelligence. She deals with issues of legal evaluation of acts committed using such technologies. Conducts research on the development of artificial intelligence algorithms in the media. Deals with the problems of digital addiction; content and communication risks in social networks. Her current research area includes artificial intelligence in legal practices (Legal Tech, LawTech, Regtech); digitalization of the educational process; implementation of blockchain technology; the use of big data in criminology.

Matthew CROSSTON

Dr. Matthew Crosston is an acclaimed author and educator who consults with governments and academic institutions on a range of issues covering education innovation, human rights conflicts, resource dilemmas, intelligence, and cyber. He serves under the Provost at Bowie State University just outside of Washington DC as its first-ever Director of Academic Transformation. Additionally, he serves as Senior Research Fellow at the Institute for National Security Studies in Tel Aviv, Israel, Senior Advisor for the Research Institute for European and American Studies in Athens, Greece, Senior Fellow at the China Eurasia Council for Political and Strategic Research in Nanjing, China, Executive Vice Chairman of ModernDiplomacy.eu, was the first American invited to conduct a political analysis column for the Russian International Affairs Council in Moscow, Russia, and is on the Senior Advisory Board of the International Journal of Intelligence and Counterintelligence. He has published top-tier research that has made required readings lists at US STRATCOM, CYBERCOM, the Army and Naval War Colleges, Mossad, the US Department of State, the Ministry of State Security in China, and the Russian Ministry of Foreign Affairs. Overall, his work has been translated into Russian, Arabic, Mandarin, French, Indonesian, Hebrew, Spanish, Turkish, Farsi, Greek, and Uzbek. He has a BA from Colgate University, MA from the University of London, PhD from Brown University, and completed a Post-Doctoral Fellowship at the University of Toronto.

Svetlana S. GOROKHOVA

Associate Professor S. S. Gorokhova holds a PhD in Law, is a lecturer of legal disciplines at the Financial University under the Government of the Russian Federation (Moscow), and at the same time is a leading researcher at the Center for Legal Research and Expertise of the Faculty of Law of the Financial University. She is a member of the Expert Council on control issues in the field of state (municipal) finance and other resources under the State Duma Committee on Control and Order of the Russian Federation. Svetlana Gorokhova is a member of the editorial board of the Russian peer review journal *Scientific Notes of Young Researchers*. As a part of the research team, she has been participating in a fundamental scientific research on the state task of the Financial University on the topic "Theory of legal regulation of artificial intelligence, robots, and robotics in the Russian Federation" for three years (2019-2021). She is the author/co-author of more than 100 academic publications published in Russian, English, and Spanish, including 10 monographs, 15 textbooks and manuals for higher educational institutions, and more

than 70 scientific articles. A significant part of the published research is devoted to issues of national (including psychological) security, as well as issues of legal regulation of relations complicated by the element of artificial intelligence.

Nguyen Quoc HUNG

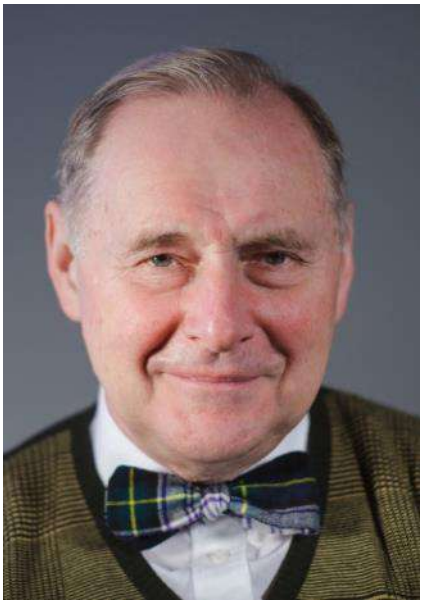
He holds a PhD in economic sciences and is a Senior Researcher at the Center for Russian Strategy in Asia at the Institute of Economics of the Russian Academy of Sciences. He has published more than 40 academic publications, including publications on the topic of the digital economy and the development of scientific and technological cooperation between Russia and Vietnam. He is one of the executors of the Russian Foundation for Basic Research (RFBR) grant "Problems of implementation and expected effects of the Free Trade Agreement between the Eurasian Economic Union and the Socialist Republic of Vietnam." His research interests are: Russian–Vietnamese economic relations, trade and economic relations of Russia with the countries of Indochina and ASEAN, and integration processes.

Pavel KARASEV

He holds a PhD in political sciences and is a Senior Researcher at the Center for Information Security Issues at the Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University. His research interests encompass the political dimension of ICT use in international relations, varying from issues of cyber warfare to problems of terminology to the filtration of objectionable content and application of international law to cyberspace. He has authored and co-authored reports and research articles on issues of information security, cybersecurity, and international information security. He has been a participant and speaker at major conferences and summits on cybersecurity and information security in Russia, Germany, India, and the USA.

Andrew MCSTAY

Prof. Andrew McStay is Professor of Digital Life at Bangor University, UK. He is a world-leading scholar in how emotion-sensing technologies are transforming society, with his work regularly featured in international media, and his major books translated into Chinese and Indonesian. His 8 monographs include, most recently, *Emotional AI: The Rise of Empathic Media* (2018) which examines the impact of technologies that make use of data about affective and emotional life. Director of The Emotional AI Lab, current projects include cross-cultural social analysis of emotional AI in UK and Japan. Non-academic work includes IEEE membership (P7000/7014) and ongoing advisory roles for start-ups, NGOs, and policy bodies. He has also appeared and made submissions to the United Nations Office of the High Commissioner on the right to privacy in the digital age, the UK House of Lords AI Inquiry and the UK Department for Culture, Media and Sport Inquiry on emotion, news, and reality media.



Alexander RAIKOV

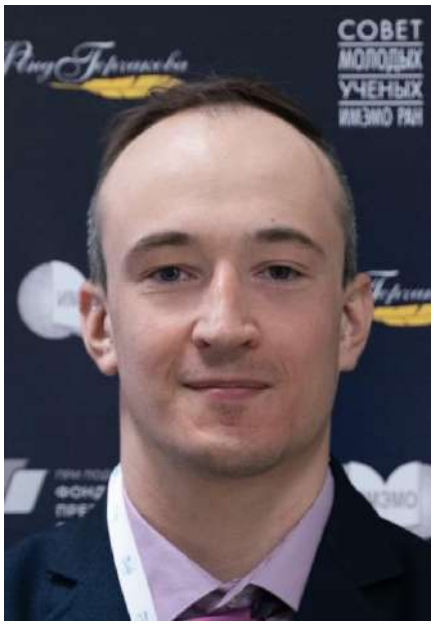
Professor, Doctor of Technical Sciences, State advisor of the Russian Federation of the 3rd class, Winner of the Russian government award in the field of Science and Technology, General Director of the New Strategies Agency (NSA) Ltd. Professor of the State Technological University (Russia), Leading researcher of the Institute of Control Sciences of the Russian Academy of Sciences, Senior Researcher of the National Center of Excellence in the field of Digital Economy of the Lomonosov Moscow State University, Chief researcher of the Institute of Philosophy, Russian Academy of Sciences. He is a member of the International Research Group on Threats to International Psychological Security through Malicious Use of Artificial Intelligence (Research MUAI). Prior to 1993, Prof. A. Raikov developed automated control systems for the Russian Government. In 1993 – 1999, he was the head of the Information-Analytical Department of the Presidential Administration and coordinator of the Situation Center of the Russian President. Since 1999, he and his team have completed about 60 projects for: the Administration of the President of the Russian Federation, Ministry of Education of Russia, Ministry of Economic Development of Russia, the Government of the city of Moscow, the Republic of Kazakhstan, etc. Prof. Raikov developed strategies for improving the quality of health care and education in the regions and scientific cities of Russia. Author of more than 350 research papers, 6 monographs, 9 patents in the field of strategic management, socio-economic development, information and analytical technologies, AI, situation centers, decision support systems, and network expertise.

Marina RESHETNIKOVA

Marina S. Reshetnikova is an Associate Professor at the Department of Economic and Mathematic Modelling at the Peoples' Friendship University of Russia (RUDN University), holds a PhD, and is a researcher in economic sciences and an innovation expert. She has been a member of the following research associations: Triple Helix Association (THA), The Free Economic Society of Russia**,** Centre d'Etudes en Macroéconomie et Finance Internationale (CEMAFI). She is also a deputy editor-in-chief of *RUDN Journal of Economics*. Marina is the author of over 30 research articles and a participant of more than 20 international academic conferences and seminars in Russia, Great Britain, Czech Republic, Italy, Greece. Her main research interests include the problems of ensuring sustainable development within the uncertainty of the global economy, innovation, and the problem of AI proliferation and especially its malicious use.

Vitali ROMANOVSKI

Vitali currently holds the position of analyst with the Belarusian Institute of Strategic Research. He has substantive field experience in the Middle East, including an advisory role with an intergovernmental program on military-technical cooperation in the UAE and an analytical post with the United Nations Assistance Mission in Iraq, Baghdad. Regular participant of the Chatham House events on information security and counter-terrorism in the EU, Middle East, and Russia. Member of the International Studies Association (ISA) and East European Studies Association (CEEISA). His research interest includes intelligence studies, psychological warfare, artificial intelligence, and information security.



Sergey A. SEBEKIN

He defended his PhD thesis "The Genesis and Development of Strategies to Deter Cyber Threats in the United States, China and Russia (1990s – 2014)" in 2020. He is a lecturer at the Department of Political Science, History and Regional Studies of Irkutsk State University. He is a Fellow of the Oxford Russian Foundation 2014–2015. In 2016, he completed an internship at Hokkaido University as part of the Russian–Japanese program Russia–Japan East 3 (RJE 3), Sapporo, Japan. In 2020–2021, he completed an internship at the Moscow-based the Russian Center for Policy Studies (PIR Center) under the program "New Technologies and International Security". His research interests are: issues of international cybersecurity, theories of cyber warfare, artificial intelligence and the future of international relations, and the impact of high technologies on international relations. He has authored 30 academic articles, analytical notes and papers on various aspects of international cybersecurity, published by such journals and organizations operating in the field of international relations as *Russia in Global Affairs*, *Russian International Affairs Council*, the *Valdai International Discussion Club*, *the PIR Center*, and *the Primakov Center for Foreign Policy Cooperation*.

Pierre-Emmanuel THOMANN

Pierre-Emmanuel Thomann is a doctor in geopolitics. His doctorate was obtained at the French Institute of Geopolitics (IFG, University Paris 8) in 2014. The thematics of his expertise cover Franco-German relations, pan-European and global geopolitical issues, geopolitical cartography, and the geopolitical dimension of artificial intelligence issues. He teaches at Lyon 3 Jean Moulin University, Lyon, France. He is president / founder of an international association Eurocontinent based in Brussels (website www.eurocontinent.eu) whose objective is to broaden the debates about the European project on geopolitical issues. Pierre-Emmanuel Thomann is a member of the International Research Group on Threats to International Psychological Security through Malicious Use of Artificial Intelligence (Research MUAI). Pierre-Emmanuel regularly publishes articles in academic journals and specialized magazines on diplomatic and geopolitical issues, and contributes regularly to international conferences and seminars on international affairs organized by UNO, OSCE, and UNESCO, at international and national conferences in France, Italy, Russia, Uzbekistan, and other countries.

Marius VACARELU

Marius Vacarelu holds a PhD and is a researcher in political sciences and a legal expert. He graduated from the Law Faculty in Bucharest. Marius Vacarelu teaches public law in the National School of Political Science and Public Administration since 2005, he is a member of the committee, which edits the Romanian magazine *GeoPolitica*, and is a head of "The Geopolitics of the East Association" which runs the website www.geopoliticaestului.ro. Marius Vacarelu is a member of the International Research Group on Threats to International Psychological Security through Malicious Use of Artificial Intelligence (Research MUAI). He is an author/co-author/coordinator of 22 books and more than 200 academic articles including the articles/book chapters on the role of AI in international relations and political processes. Marius Vacarelu is a frequent speaker on Romanian television on geopolitics issues. He is a blogger for Romania's most important journal *Adevarul*. Marius has presented papers and published articles in Russia, Czech Republic, France, Poland, UK and the US.

International Center for Social and Political Studies and Consulting (ICSPSC)

The International Center for Social and Political Studies and Consulting (ICSPSC) was founded in March 2002 as an association of researchers and consultants from different countries. Over the years, the ICSPSC has organized hundreds of international academic conferences, roundtable discussions, and workshops concerning the issues of national and international security and strategic communication, and published about 30 books and different reports. Monographs and collections of articles published by the ICSPSC in Russian and English include:

- Armies and Politics (in English);
- Russia and Latin America (in Russian);
- Russia and India – Strategic Partners(in English);
- Public Relations Training Courses (in Russian);
- Avenir Khanov – a Person, a Citizen, and a Diplomat (in Russian);
- India – Russia: A Dialogue between Civilizations(in English);
- India – Russia: Trade and Economic Relations (in English);
- Genesis of Russia's Market Reforms (in Russian);
- Mass Media and PR in Bulgaria (in Russian);
- Hugo Chavez and the Bolivarian Revolution (in Russian);
- Communication Management. Consulting in Public Relations (in Russian);
- Public Relations and Communication Management: The Foreign Experience (in Russian);
- The Foreign Policy of the USA: The Communication Aspect (in Russian);
- Communication Management in World Politics and Business (in Two Volumes, in Russian);
- The Rising Role of Communication Management in World Politics and Business (in English);
- Ultra-Left Terrorism in Germany: Major Trends in the Activity of the Red Army Fraction (RAF) and its Communication Maintenance (in Russian);
- Communication Management in the Foreign Policy of France in the Late 20th Century (in Russian);
- Communication Management and Strategic Communication(in Russian);
- Crisis, Army, Revolution (in Russian);
- The Presidents in Media Focus: The Practice of Psychological Warfare in Latin America;
- Hugo Chavez and Psychological Warfare in Venezuela (in Russian);
- Communication Management and Strategic Communication: The Modern Forms of Global Influence and Control (in Russian);
- "Ukraine" Strategic Provocation (in Russian);
- Communication and Terrorism (in Russian),
- Strategic Communication in EU-Russia Relations: Tensions, Challenges, and Opportunities (in Russian);
- Malicious Use of Artificial Intelligence and International Psychological Security in Latin America (in English).
- Malicious Use of Artificial Intelligence as a Threat to Psychological Security: Northeast Asia and the Rest of the World (in Russian and English).

Among the authors of these books are more than 102 researchers from 28 countries in Europe, Asia, and North and South America.

One of the most recent projects of the ICSPSC is the development of international associations that work in various fields of strategic studies and strategic communication. Leading scholars, CEOs, and employees of public and private structures and non-governmental organizations from Asia, Oceania, Africa, Europe, and South and North America are taking part in the activities of these associations (See more at GlobalStratCom: http://globalstratcom.ru/globalstratcom-eng/).

E-mail: icspsc_office@mail.ru, icspsc@mail.ru

GlobalStratCom

Russia is developing cooperation with different regions of the world. The GlobalStratCom platform aims to develop five associations in various fields of strategic studies and strategic communication. The following are currently in progress:

- European – Russian Communication Management Network (EU-RU-CM Network)
- Russian – Latin American Strategic Studies Association (RLASSA)

Leading scholars, heads, and responsible employees of public and private structures and non-governmental organizations from Asia, Oceania, Africa, Europe, and South and North America are taking part in the activities of these associations.

Research Areas

- Challenges and threats to national and international security: joint interests and possible areas of collaboration between Russia and other countries;
- Armed Forces and politics;
- Conflict resolution and crisis management;
- Participation in peace missions;
- Malicious use of artificial intelligence and psychological security
- Participation in wars and military conflicts;
- Prospective models of social and political development;
- New technologies and their influence on social development and security issues;
- Activities of law enforcement agencies;
- Terrorism and communication;
- Armed Forces, State, and Society;
- Strategic communication;
- Military history;
- Strategic studies as an area of cooperation between Russia and other countries;
- War and peace studies.

For more information, see the website of GlobalStratCom.

Evgeny N. PASHENTSEV

Prof. Evgeny Pashentsev is a Leading Researcher at the Diplomatic Academy at the Ministry of Foreign Affairs of the Russian Federation, and the Director of the International Center for Social and Political Studies and Consulting (Moscow). He is a coordinator of the European-Russian Communication Management Network (EU-RU-CM Network), the Russian-Latin American Strategic Studies Association, and the International Research Group on Threats to International Psychological Security through Malicious Use of Artificial Intelligence (Research MUAI). He is a partner of the European Association for Viewers Interests in Brussels, a member of the International Advisory Board of *Comunicar* (Spain), and the Editorial Board of the *Journal of Political Marketing* (USA). Prof. Pashentsev has authored, co-authored, and edited 37 books and more than 200 academic articles published in Russian, English, Spanish, Portuguese, Italian, Serbian, Vietnamese, and Bulgarian. He has presented his papers at more than 180 international conferences and seminars over the last 10 years in 24 countries. The areas of his current research include strategic communication and malicious use of AI.

Experts on the Malicious Use of Artificial Intelligence and Challenges to International Psychological Security

Report by Evgeny Pashentsev
Edition by the ICSPSC. December 2021, Moscow

ISBN 978-5-00166-528-1

Please send your letters and comments to the author at icspsc@mail.ru; icspsc_office@mail.ru.

Published in the Russian Federation in the printing house «OneBook.ru» LLC «SAM Polygraphist».