

## **Artificial intelligence and international psychological security: academic discussion in Khanty-Mansiysk and Moscow**

The development of artificial intelligence and machine learning, embedded systems and devices, the Internet of things, augmented and virtual reality, big data analysis (data science) and cloud computing, block chain, etc. stimulate the transition to a new technological order. At the same time, positive expectations associated with scientific and technological progress are combined with clearly perceived threats of the approaching future, which are described and analyzed by representatives of various discourses – from mass media to academic and political circles. The development of artificial intelligence (AI) is a subject of discussion in the broad context of economic, political and related social transformations. The importance of the introduction of AI is now recognized by almost all states and international organizations. However, new threats in the field of psychological security, caused primarily by global international tensions and the growing influence of non-state actors, including criminal, terrorist organizations, are closely linked with the use of new tools provided by AI, in the field of communication in particular.

The attention of the academic community to this range of problems is evidenced by the active discussion that took place at the panel discussion “Malicious use of artificial intelligence and international psychological security” of the UNESCO Conference in Khanty-Mansiysk and continued at the eponymous research seminar at the Diplomatic Academy of the Ministry of Foreign Affairs of the Russian Federation.



**Conference participants**

The II International Conference “Tangible and Intangible Impact of Information and Communication in the Digital Age” was held within the framework of the [UNESCO Intergovernmental Information for All Programme \(IFAP\)](#) and [XI International IT Forum with the participation of BRICS and SCO](#) countries in Khanty-Mansiysk on June 9-12, 2019. The conference was organized by the [Government of the Khanty-Mansiysk Autonomous Okrug – Ugra](#), [Commission of the Russian Federation for UNESCO](#), [UNESCO Intergovernmental Information for All Programme \(IFAP\)](#), [UNESCO Institute for Information Technologies in Education \(IITE\)](#), [Russian Committee of the UNESCO IFAP](#), [Diplomatic Academy of the Ministry of Foreign Affairs of the Russian Federation](#) and the [Interregional Library Cooperation Centre](#).

A number of academic institutions provided academic support for the event. These are the [International Center for Social and Political Studies and Consulting \(ICSPSC\)](#), the [European-Russian Communication Management Network \(EU-RU-CM Network\)](#) and [the Russian – Latin American Strategic Studies Association \(RLASSA\)](#).

The conference was supported by the [publishing house “Ugra News”](#), the [Institute for Political, Social and Economic Studies – EURISPES](#) (Rome), the [Association of Studies, Research and Internationalization in Eurasia and Africa – ASRIE](#) (Rome), the [Geopolitics of the East Association](#) (Bucharest), the [International Association “Eurocontinent”](#) (Brussels) and the [International Institute for Scientific Research – IIRS](#) (Marrakech).



**The Governor of Ugra Natalia Komarova at the opening the panel discussion “Malicious use of artificial intelligence and international information and psychological security”**

The Government of Ugra attaches great importance to the development of cooperation with UNESCO, which has been counting down [since 2001](#). The Governor of Ugra **Natalia KOMAROVA** [took part in the conference](#). Opening the panel discussion “Malicious use of artificial intelligence and international information and psychological security”, the head of the region recalled that a year ago

the Ugra Declaration adopted at the first UNESCO conference included proposals for the preparation of a world report on socio-cultural transformations in the digital age, the formation of educational programs relating to ethical, legal, cultural, social aspects of life. “In the modern world, artificial intelligence is rapidly gaining popularity, the issues of its implementation for good purposes are discussed, concerns about the negative impact are expressed, – Natalia Komarova continued. – Russian President Vladimir Putin, speaking at the recent Economic Forum in St. Petersburg, said that by 2024 the world market of products using artificial intelligence will grow by almost 17 times”. She stressed that in the era of global transformations, competition for resources, and especially for human ones (intelligence), is growing.

According to the Governor, in connection with the mass involvement of people in the global communication space, the subtlety of the ideological setting of society is of particular importance: “Social identity is forced to comprehend changes, accepting, mastering changing values. Experts note that in a situation of instability of the value landscape identity can not be stable – it becomes fluid. In this case, the English sociologist Anthony Giddens notes, identity is constructed, modernity is distinguished by ‘appetite for new’, and this new encourages us to act, including the development of artificial intelligence, which in turn requires the creation, use of information security, prevention of cyber attacks, protection against data leakage. At the same time, according to Natalia Komarova, there is a need to ensure psychological security.

[Vice-Chair of the Intergovernmental Council for the UNESCO Information for All Programme \(IFAP\)](#); Chair of UNESCO IFAP Working Group for Multilingualism in Cyberspace; Chair of the Russian IFAP Committee, and President of the [Interregional Library Cooperation Centre](#) **Evgeny KUZMIN** recalled that since 2001, Ugra has hosted major information for all conferences. “This is a serious contribution of the Khanty-Mansiysk Autonomous Okrug and the whole of Russia to the implementation of this flagship intergovernmental program of UNESCO. The events were devoted to the discussion of such important issues as the preservation of the languages of the peoples of the world and the development of linguistic diversity in cyberspace, the formation of open government, transparency of governance, the interaction of governments and the population by improving information literacy among officials and citizens,” he stressed.



**Evgeny Kuzmin**

The [Interregional Library Cooperation Centre](#) under the leadership of Evgeny Kuzmin made significant efforts to ensure that the conference was held at a high academic level. In total, the conference was attended by representatives of 35 countries from all over the world.



**From left to right: Marco Ricceri, Evgeny Pashentsev, and Dorothy Gordon**

With academic support of EU-RU-CM Network the conference was attended by coordinators and network members: Darya Bazarkina (Russia), Evgeny Pashentsev (Russia), Olga Polunina (Russia), Marco Ricceri (Italy), Gregory Simons (Latvia/New Zealand/Sweden), Pierre-Emmanuel Thomann (Belgium) and Marius Vacarelu (Romania).

*At the opening of the conference **Evgeny PASHENTSEV**, Leading Researcher, Diplomatic Academy of the Ministry of Foreign Affairs of the Russian Federation; Director, International Centre for Social and Political Studies and Consulting (Moscow, Russia), Coordinator of the European-Russian Communication Management Network (EU-RU-CM Network), senior researcher, St. Petersburg State University presented a paper on "Artificial Intelligence: Current and Promising Threats to International Psychological Security". We further provide the full text of his paper.*

International security today is under threat due to destructive processes in the economic, social, military and other spheres of public life. Negative processes are developing at the national, regional and global levels. It is essential to have an adequate understanding of the existing problems by all sectors of society. But this goal is far from being realized. Perhaps because of the vested interests and/ or irresponsibility of the ruling elites who seek to hide the truth about the real state of things. Or, much worse, because of their intellectual and moral degradation, the elites are unable to respond to the increasingly obvious threats through a system of effective action, not even for the sake of social progress, but for the sake of their physical self-preservation.

The low level of civil self-organization of society, the lack of established progressive counter-elites testifies to the crisis not only of the "top", but also of the "bottom" strata of the society, i.e. of the civilizational crisis. The only way to solve the problems facing humanity is to have access to information, to the modern possibilities of its processing, analysis and dissemination. On this basis it is possible to offer scientifically-proved models of progressive development of mankind, their discussion, both in the professional environment, and at public discussion forums. Artificial intelligence can be of great help in the processing, analysis, verification of research results and implementation of relevant social development programs. Human beings' intellectual creative abilities, will, aspirations and feelings "Weak AI" will not replace, and "Strong AI" equal or superior to the modern human mind – is a matter of the future. Unfortunately, the rapidly growing practice of using AI to manipulate public consciousness at the international level, once again testifies to the large and dangerous potential of the negative consequences of the use of new technologies.

International psychological security (IPS) means protecting the system of international relations from negative information and psychological influences associated with various factors of international development. The latter include targeted efforts by various state, non-state and supranational actors to achieve partial/complete, local/global, short-term/long-term, and latent/open destabilization of the international situation in order to gain competitive advantages, even through the physical elimination of the enemy.

International actors engaging in hybrid warfare exert negative *direct* and *indirect* impacts on the enemy's public consciousness and, often, on themselves, their allies and neutral actors. For example, economic sanctions are intended not only to financially weaken/destroy the enemy, but also to reduce the readiness of target groups for further resistance by increasing the enemy's economic problems. Military-political confrontation with the enemy, based on aggressive interests and mass genocide against other nations, causes irrecoverable damage to the mentality and psyche of the aggressor country's population. At the same time, psychological warfare (PW) is always aimed at delivering direct (although often latent) blows to the enemy's public consciousness and achieving (through victory in this sphere) a bloodless total victory over the enemy. In fact, the modern global



world is witnessing hybrid warfare in the system of international relations, which has never completely stopped throughout history; rather it has had natural periods of exacerbation. We have clearly entered a long-term transition period in the development of humanity and the system of international relations in particular, which is accompanied by irregularly growing PW.

In our opinion MUIAI can allow hostile actors to be more successful than so far in:

- *provoking a public reaction to a non-existent factor of social development* in the interests of the customer of psychological impact. The target audience sees something that doesn't really exist.
- *presenting a false interpretation of the existing factor of social development* and thus provoking the desired target reaction. The audience sees what exists, but in a false light.
- significantly and dangerously *strengthening (weakening) public reaction to the real factor of social development*. The audience sees what exists but reacts inadequately.

**We can** suggest the *following MUIAI classification* according to the degree of implementability:

- current MUIAI practices;
- existing MUIAI capabilities that have not been used in practice yet (this probability is associated with a wide range of new rapidly developing AI capabilities — not all of them are immediately included in the range of implemented MUIAI capabilities);
- future MUIAI capabilities based on current developments and future research (assessments should be given for the short, medium and long term);
- unidentified risks, also known as “the unknown in the unknown.” Not all AI developments can be accurately assessed. Readiness to meet unexpected hidden risks is crucial.

It is important and necessary to use independent teams of different specialists and AI systems to assess MUIAI capabilities.

We can also propose the following MUIAI classifications:

- by territorial coverage: local, regional, global;
- by the degree of damage: insignificant, significant, major, catastrophic;
- by the speed of propagation: slow, fast, rapid;
- by the form of propagation: open, hidden.

Among the possible threats of MUIAI (see more detailed: Bazarkina and Pashentsev, 2019), which can cause a serious destabilizing impact on the socio-political development of a country and the system of international relations, including the sphere of IIB, are the following:

• ***The growth of integrated, all-encompassing systems with active or leading AI use*** increases the risk of malicious takeover of such systems. Numerous infrastructure facilities, for example, robotic self-learning transport systems with AI-based centralized management, can be convenient targets for high-tech terrorist attacks. If terrorists seize control over the transport management system of a large city, this may lead to numerous casualties, cause panic and create a psychological climate that will facilitate further hostile actions.

• ***The reorientation of commercial AI systems***. Commercial systems can be used in harmful and unintended ways, such as deploying drones or autonomous vehicles to deliver explosives and cause crashes (Brundage, et al., 2018, p. 27). A series of serious disasters, especially those involving celebrities, may cause international media hype and damage IPS.

• ***Attacks further removed in time and space***. Physical attacks are further removed from the actor initiating the attack as a result of autonomous operation using AI (Brundage, et al., 2018, p. 28).

The surprise effect of such attacks may destabilize the system of international relations. For example, nuclear devices can be simultaneously set off from afar in different countries of the world without direct human participation. Officials of all countries that possess modern technologies speak of the need to retain control over the combat uses of AI systems. This is understandable, since no government, reactionary or progressive, wants to lose control over their weapons. But this does not apply to non-state actors: for example, a group of techno-religious maniacs who want to eliminate humanity will have an increasing chance of success due to the continuous improvement of AI, the creation of complex cross-border AI systems, the propagation of new technologies, and other factors.

• **The creation of 'deepfakes'.** 'Deepfake' (a portmanteau of "deep learning" and "fake") is an AI-based [human image/voice synthesis](#) technique. Many celebrities, including Scarlett Johansson, Maisie Williams, Taylor Swift and Mila Kunis, have fallen victim to deepfake pornography. Deepfake hobbyists have begun using this technology to create digitally-altered videos of world leaders, including U.S. President [Donald Trump](#), Russian President [Vladimir Putin](#), former U.S. president [Barack Obama](#) and former presidential candidate [Hillary Clinton](#). Experts warn that the videos could be realistic enough to manipulate future elections and global politics as early as 2020 (Palmer, 2018). However, it could take years before researchers invent a system that can reliably detect deepfakes, which makes them a potentially dangerous lever for influencing the behavior of individual persons and large target groups. Deepfakes can be used in psychological warfare to provoke financial panic and trade or hot wars. Fake videos of Israeli Prime Minister Benjamin Netanyahu or other government officials – for instance, talking about impending plans to take over Jerusalem's Temple Mount and Al-Aqsa Mosque – could spread like wildfire (The Times of Israel, 2018). Just as dangerous is the possibility that deepfake technology spreads to the point that people are unwilling to trust video or audio evidence (Waddel, 2018).

• **'Fake People' technology.** After the sale of the first AI-generated painting in early 2018, deep learning algorithms now generate portraits of non-existent people. The NVIDIA company has recently published the results of the work of a generative adversarial network (GAN) trained to generate images of people (Karras, Laine and Aila, 2018). The technique is based on an infinite collection of images of real faces; this is why a neural network recognizes and applies many fine details in its work. It can generate hundreds of faces with glasses, but with different hairstyles, skin textures, wrinkles and scars, and add age signs, cultural and ethnic features, emotions, moods or effects of external factors, such as wind in the hair or an uneven tan. Back in 2017, NVIDIA experts held a similar experiment, but the images of faces they got then were blurry and were easily recognized as fakes. Today, neural networks are incomparably better and generate faces in high resolution. They can easily produce, for example, an image of a non-existent illegitimate child of a celebrity, with a perfect family resemblance, as a provocation.

• **Agenda setting and amplification.** Studies indicate that bots made up over 50 percent of all online traffic in 2016. Entities that artificially promote content can manipulate the "agenda setting" principle, which dictates that the more often people see certain content, the more they think it is important (Horowitz, et al., 2018, pp. 5-6). Reputational damage done by bots during political campaigns, for example, can be used by terrorist groups to attract new supporters or organize assassinations of politicians.

• **Sentiment analysis** is a class of content-analysis methods used in computational linguistics to identify emotionally loaded words in texts that reveal the author's opinion of the topic. Sentiment analysis is done on the basis of a wide range of sources, such as blogs, articles, forums, polls, etc. This can be a very effective tool in PW.

**The development of Artificial Emotional Intelligence (AEI).** Here research is conducted not only in the field of artificial, but also natural, human intelligence. The research here develops in several areas: firstly, the recognition of emotions of humans and animals, secondly, is the analysis, interpretation of these emotions and the necessary for that techniques. For this purpose, machine learning and big data analysis are used. The third direction is the reproduction of emotions in robotic systems. If we talk about the reproducibility of emotions in robotic systems, the Japan has achieved great success in this area. This direction is well applicable for example to the field of care for the elderly. The full creation of emotional AI is possible only within the framework of the creation of "Strong AI". Unfortunately, the development of emotional AI within the framework of "Weak AI" also poses many threats to the IPS as it opens up new forms of control over the human mind through AI, including the provocation of mass riots.

•AI, machine learning and sentiment analysis make it possible to **predict the future by analyzing the past** — quite a holy grail for the financial sector or government planning agencies. But various state and non-state actors can potentially use this possibility for MUIAI. Particularly important are **prognostic weapons**: predictive analytics methods based on big data and AI, which make it possible to correct the future from the present in one's own interests and contrary to the objective interests of the target. For example, the Intelligence Advanced Research Project Activity (IARPA) launched the Early Model Based Event Recognition Using Surrogates (EMBERS) program in 2012 to forecast socially significant population-level events, such as incidents of civil unrest, disease outbreaks, and election outcomes. For civil unrest, EMBERS produces detailed forecasts about future events, including the date, location, type of event, and protesting population, along with any uncertainties. The system processes a range of data, from open-source media, such as Twitter, to higher-quality sources, such as economic indicators, processing about five million messages a day. The system delivers more than 50 predictions about civil unrest alone for 30 days ahead (see: Doyle, et al., 2014).

•**Growing threats from phishing.** AI allows to increase dramatically data processing speed and respond faster to people's expectations, which makes phishing more dangerous. Progress in automated spear phishing has demonstrated that automatically generated text can be effective at fooling humans, and indeed very simple approaches can be convincing to humans, especially when the text pertains to certain topics such as entertainment (Brundage, et al., 2018, p. 3, 46). Main methods of using artificial intelligence hackers are phishing, spear phishing and whaling, i.e. phishing focused on senior managers responsible for financial decision-making.

•**Computer games using AI can also increase the effectiveness of psychological impact,** especially on children and adolescents. AI is already actively used in the creation of computer games. That computer games can have a certain manipulative affect, it is known well for a long time, however the analysis of use of AI for these purposes – one of perspective tasks of researchers. From the point of view of IPS, special attention should be paid to computer games, which are widely distributed in many countries of the world.

•It can be imagined that due to a combination of psychological influence techniques, sophisticated AI systems and big data, **synthetic information products** could emerge in the near future that would be similar in nature to modular malicious software. However, they will have an effect not on inanimate objects, social media, etc., but on humans (individuals and masses) as psychological and biophysical beings. These synthetic information products will contain software modules that will drive large numbers of people into depression. After that, suggestive programs will latently come into action. Appealing to habits, stereotypes, and even psychophysiology, they will encourage people to perform strictly defined actions (Larina and Ovchinskiy, 2018, pp. 126-127).



However, any of the above mentioned threats can also be more effectively neutralized with a help of AI. For example, *Swisscom Innovations* developed and trained an artificial intelligence based phishing detection system. It predicts reliably whether a formerly unknown website contains phishing or not (Bürgi, 2016). Another programme, *Lookout Phishing AI* continuously scans the Internet looking for malicious websites. *Lookout Phishing AI* detects the early signals of phishing, protects end users from visiting such sites as they come up, and alerts the targeted organizations (Richards, J., 2019).

The task today is to repel threats from the real and constantly developing “weak” artificial intelligence, which is a threat not in itself, but because of the actions of antisocial external and internal actors that turn it into a threat to the international security. In the not so distant future, there may be problems associated with “strong intelligence”, the possibility of which in the coming decades, forecast more and more researchers.

### References

Blinnikova, N., 2018. Emocional'nyj II podskazhet, kogda na rabote luchshe pojti popit' čaj, i pomozhet borot'sja so stressom (Emotional AI will tell you when to go to work to drink tea, and will help fight stress). ITMO News. <[http://news.ifmo.ru/ru/startups\\_and\\_business/initiative/news/7703/](http://news.ifmo.ru/ru/startups_and_business/initiative/news/7703/)> [Accessed 23 June 2019].

Brundage, et al., 2018. *The malicious use of artificial intelligence: forecasting, prevention, and mitigation*. Oxford, AZ: Future of Humanity Institute, University of Oxford.

Bürgi, U., 2016. Using Artificial Intelligence to Fight Phishing. *Swisscom [online]*. Available at: <<https://ict.swisscom.ch/2016/11/using-artificial-intelligence-to-fight-phishing/>> [Accessed 22 June 2019].

Crowder, J. A., and Friess, S., 2012. Artificial Psychology: The Psychology of AI. Conference Paper. <<https://www.comparethecloud.net/articles/artificial-intelligence-human-behaviour/>> [Accessed 23 June 2019].

Doyle, A., et al., 2014. *Forecasting significant societal events using the EMBERS streaming predicative analytics system*. *Big Data*, Vol. 4, pp. 185–195.

Horowitz, M. C., et al., 2018. *Artificial intelligence and international security*. Washington: Center for a New American Security (CNAS).

Karras, T., Laine, S., and Aila, T., 2018. A style-based generator architecture for generative adversarial networks. *arXiv of Cornell University [online]*. Available at: <<https://arxiv.org/pdf/1812.04948.pdf>> [Accessed 31 January 2019].

Larina, E., and Ovchinskiy, V., 2018. *Iskusstvenny? intellekt. Bol'shie dannye. Prestupnost' [Artificial intelligence. Big Data. Crime]*. Moscow: Knizhnyj mir.

Richards, J., 2019. What is Lookout Phishing AI? *Lookout Blog*. <<https://blog.lookout.com/lookout-phishing-ai>> [Accessed 22 June 2019].

*The Times of Israel*, 2018. 'I Never Said That!' The High-Tech Deception of 'Deepfake' Videos. *The Times of Israel [online]*. Available at: <<https://www.timesofisrael.com/i-never-said-that-the-high-tech-deception-of-deepfake-videos/>> [Accessed 31 January 2019].

Waddel, K., 2018. *The impending war over deepfakes*. *Axios [online]*. Available at: <<https://www.axios.com/the-impending-war-over-deepfakes-b3427757-2ed7-4fbc-9edb-45e461eb87ba.html>> [Accessed 31 January 2019].

Next, we give a summary of a number of papers presented at the panel “Malicious use of artificial intelligence and international information and psychological security”.

**Darya BAZARKINA**, Professor, Russian Presidential Academy of National Economy and Public Administration; Senior Researcher, Saint Petersburg State University (Moscow, Russia)

*Artificial Intelligence as a Terrorist Weapon: Information and Psychological Consequences of Future Terrorist Attacks and Ways to Minimize Them*

The threats posed by the use of artificial intelligence by terrorist organizations can be divided into two groups:

- 1) Use of AI for destruction of physical objects, killing and harming the health of citizens;
- 2) Use of AI in the propaganda activities of terrorist groups.

It can be assumed that the first group of threats has no less pronounced psychological impact on the target audience than the second, for which the psychological effect follows from the definition. A terrorist act is an act of communication in itself, and considering possible terrorist acts in which an AI becomes a murder weapon, it is worth considering whether such a murder of citizens is more shocking than similar in number of victims, but performed by "traditional" means.

It is no accident that the organization "Islamic State" (IS) is actively recruiting specialists in the field of high technologies. Even now, terrorists are experimenting with crypto currencies that allow you to transfer funds across borders, avoiding bank control. It is already clear that machine learning technologies are becoming increasingly available. Drones are already equipped with AI, and the use of military equipment that can operate without human help, has become the subject of lively discussions. Unfortunately, the documented use of social media, encryption and drones by terrorists suggests that once new technologies become widely available to the consumer, terrorists will also be able to use them.

In the field of working with information, AI capabilities are very wide. The analysis of big data based on the contents of social media has already been used by North African militants in the attack on the Tunisian city of Ben Gardane in March 2016. Available evidence, including effective ways of killing key members of the security service, showed that the terrorists had pre-studied the habits and schedules of the victims. This case shows that with the development of social media and their monitoring mechanisms (processing of "big data", which enhances AI), the possibilities of open-source intelligence are becoming more accessible to all sorts of non-state actors. It is only a matter of time before less technically advanced extremist groups connect these mechanisms. For example, the far right in Europe exchange information about possible targets for attacks on sites such as "Redwatch", created in Poland on the British model (the site contains photos of activists of the left movement, which is collected by the far right). The analysis of possibilities of AI already allows to draw conclusions on the facilitation of the collection of data on potential victims and for the selection of priority targets for cyber-attacks based on machine learning.



**Darya Bazarkina**

Terrorist propaganda adapts to the expectations of the target audience, of which potential and actual recruits are an important part. Finally yet importantly, to recruit young people into its ranks, IS publishes materials aimed at developing a more “high-tech” image of a terrorist, which combines the features of a fanatic and, for example, a skilled hacker. These phenomena could be a prologue to a new, much more dangerous phase of terrorist activity, in which terrorist acts could become much more destructive and their perpetrators, operating at a distance with the advanced technology, would become extremely difficult to detect. In this regard, it is worth mentioning the magazine “Kybernetiq”.

It is advisable to widely use predictive analytics mechanisms by state and supranational agencies to prevent social unrest (through timely social, economic and political measures to achieve social stability in the long term). Among measures not directly related to AI (but potentially optimized with its help), governments should develop long-term policies towards the social integration of people of different religions and sects into socially significant projects. In countries and regions experiencing social and economic instability, apart from taking measures to improve the well-being of citizens, governments should explain to the population the economic and political goals of terrorist organizations and the essence of the ideology of terrorism. The larger the scope of terrorist activities, the higher international agencies should predict and take such measures.

**Fatima ROUMATE**, Associate Professor, Mohamed V University; President, Institut International de la Recherche Scientifique (Marrakech, Morocco)

*Malicious Use of Artificial Intelligence: New Challenges for International Relations and International Psychological Security*

Nowadays, AI offers new opportunities for international and bilateral cooperation, and facilitates the inclusion of all actors within global governance. However, the malicious use of AI represents a threat to the international psychological security whether we are speaking about social, economic or military activities. In fact, this threat is an important feature of the new cold war characterized by the race toward AI. A new international order is in a progress considering the rise of new technological and economic forces which means emerging of new players and new rules of the international relations.

We'll have a closer look at the impact of the malicious use of AI on the international community and the challenges the international community faces today considering the AI race in economic and military areas.

International actors are using AI to achieve their specific beneficial goals. However, they are investing more efforts to limit their vulnerabilities. The consequence is that international society faces the psychological impact of non-trusted information which influences policy-maker decision and political changes in global affairs.



**Fatima Roumate**

Malicious use of AI leads us to think about one of the most important negative impacts, especially the attacks on democracy. In fact, AI is not only expanding existing threats, but it's creating new threats. Spear phishing attacks, for example, increased significantly since 2016 in several

countries such as Canada, France, Italy and the USA where attacks against specific targets accounts for more than 86% of all phishing attacks.

There is a new voice technology that can reproduce a believable fake voice. There is also the concept of machine learning software that creates fake videos. Moreover, "AI systems are expanding the phishing attacks space from email to other communication domains, such as phone calls and video conferencing". This shows how it's challenging for diplomats and countries to check the trustworthiness of the information and its sources before making a decision.

Governments invest in malicious use of AI, first, for surveillance and defense, second, by other states and non-states actors to create or support social movement aimed at specific political changes. In systems that combine data from satellite imagery, facial recognition-powered cameras, and cell phone location information, among other things, AI can provide a detailed picture of individuals' movements as well as predicting future movement and location. It could therefore easily be used by governments to facilitate more precise restriction of the freedom of movement, at both the individual and group level and by foreign actors who are targeting political changes. Voting behavior and election campaigns are also influenced through social media.

Malicious use of AI can influence a lot of domains as defense, diplomacy, cyber security, economic and financial sector. According to the 2017 Official Annual Cybercrime Report, cyber crime costed \$3 trillion in 2015 and it's estimated that cybercrime will cost \$6 trillion annually by 2021.

The malicious use of AI creates new challenges for states as an original actor in international relations. This invites researchers and policymakers to rethink many concepts linked to the state's notion as sovereignty, diplomacy and security, considering the appearance of new notions as artificial diplomacy, cyber security, cyber war...

AI is creating big changes in international relations. It facilitates the integration of new actors in global issues, especially in this age characterized by the diffusion of power in international society and the expansion of transnational relations. In fact, in the age of AI, it is necessary to rethink all institutions outside and inside countries. The massive interconnection between all actors imposes the need to update diplomatic tools.

As for the influence of AI systems in global affairs, Hillary Clinton argues that the use of information and communication technologies has an influence on global debates and is playing greater roles in international affairs for good and bad.

In this sense, the future of international psychological security is conditioned by the State's response to the challenges imposed by the cyber era. For that, the American security strategy focuses on the improvement of strategic planning and intelligence. At that juncture, real coordination between States and transnational corporations specialized in ICT (GAFA in USA and BAT in China) is a sine qua non-condition considering advances in artificial intelligence which are reshaping the practice of diplomacy.

The principal goal of the competition between China and USA is racing toward technological sovereignty which means, according to Nicholas Westcott, having a seat at the international table at the age of AI.

Malicious use of AI imposes new challenges related to international law and human rights, especially with the charter of principals and human rights in the internet which recognizes the access to the internet as a fundamental right. AI age is a new phase in the development of international law which becomes heavily traditional. In the same context, the appearance of the Lethal Autonomous Weapons (LAWs) creates a controversial discussion between States and it requires an urgent review

of the use of force as it was cited in the UN charter. States competition toward LAWs lead us to think that current trade crisis between China and USA can be escalated to an open military conflict with the use of AI weapons.

First, the future of humanity will be decided by no state actors when they will own LAWs. Second, all these new technologies are growing faster than international law and diplomacy. Thus, international law norms such as those concerning the use of force and defense need to be revised.

States need to invest more on the LAWs to prevent violation of international humanitarian law. According to the Report of the Special "Rapporteur" on extrajudicial, summary or arbitrary executions, States must be transparent about the development, acquisition and use of armed drones. The goal is to ensure international psychological security which is a sine qua non-condition of international security.

The growing investment in AI for commercial and military will expand the challenges and threats to international psychological security. These challenges are significant because AI is growing rapidly while the development and updating of international mechanisms is very slow. This leads us to another challenge which is the creation of right balance, first between commercial and military funding dedicated to AI and second between investment in AI and protection of human rights in peace and in war.

Malicious use of AI invites all actors (States, international institutions, NGOs, transnational corporations and individuals) to collaborate and give a written riposte in the political, juridical and institutional level. The goal is to ensure international psychological security.

The challenges imposed by malicious use of AI are pushing international society towards a new global order with several and fundamental changes of players and rules in the international game.

**Frederic LABARRE**, analyst and Education management consultant at the Royal Military College of Canada, co-chair of the Regional Stability in the South Caucasus Study Group (Partnership for Peace Consortium) (Canada)

***The Mechanics of Social Media and AI-aided Radicalization: Impact on Human Psychology*** (A digest from "Mapping Social-Media Enabled Radicalization: A Research Note" by P. Jolicoeur and F. Labarre (2017). The paper was sent to and presented at the research seminar at the Diplomatic Academy in Moscow, June 14<sup>th</sup>, 2019).

In 1957, at a moment when the discipline of political science was being developed, famed scholar David Easton came up with a systems analysis approach to examine and explain political action. At the time, his effort was ground-breaking, and for a while, hindsight looked on the systems approach of political analysis as hopelessly descriptive, and not analytical at all. Critics said that, for one, a systems approach could only explain the functioning of democratic systems. Sixty years later, with most regimes on the planet being democratic, it would seem that Easton's approach is more relevant than ever. All the more so since half of the individuals on the planet, thanks to technology, carries this democratic power in their pocket; through their cell phones and devices.

As we have argued above, technology expands the individuals' horizons and seemingly provides direct and instantaneous access to the political system. The neat compartmentalization that existed before the advent of the internet is no longer possible. Everything is instantaneous, and everyone has a voice. Clearly, this puts added pressure on elected officials. However, it also provides them with tools to channel demands. Contrary to what the Internet promised some quarter of a century ago, individuals are not exposed to many facets of a story or event, enabling them to make more



“optimal” choices. Today, technology produces a perversion of democracy at the individual level; electronic systems deliver what the individual wants as opposed to what they need.

This is because individuals are their own systems. They are biological systems. Biological systems are subject to demands and pressures, and, like any other system, produce outputs and decisions which inform future demands from their immediate environment.

Groups that are adept at seizing upon this verity also leverage the power of technology to further pressure individuals, so that the individual is, wittingly or not, “enrolled” into a project or a political vision which is not their own initially.

Individuals, themselves bombarded by demands of all sorts, instinctively seek to make sense of the world by seeking reassuring biases. They will be less likely to challenge their own views. They will tend to keep company with like-minded people in communities of thought. These communities, however, are not always physical; they are often remote, brought perceptually closer thanks to the ubiquitous Internet and social media.

The problem comes when algorithms begins “feeding” individuals with “expected” support, thereby reinforcing pre-existing biases within individuals. Social-media provides a wealth of information on individual habits, allowing virtual communities to provide individuals with messages and images that are soothing and apparently give meaning and structure to what is seemingly a raw and chaotic world. The recent scandal involving Cambridge Analytica’s role in pushing negative messaging on targeted audiences is a case in point. The current malaise with fake news has not lifted the veil from peoples’ eyes; they have merely reinforced perceptual biases; it is the others’ news outlet which are fake, never our own.

Under pressure, social media like Facebook have resorted to artificial intelligence (AI) to reduce the incidence of fake, hateful, or radical content on its site. But AI, currently in the form of Bots and other automatic trolling devices, can also be leveraged to produce and reproduce radical content exponentially. This seems to be a losing battle, and the human mind never feels overwhelmed, as the content is always so emotionally fitting. Very soon, it will be difficult to determine what is genuine content from what is human-generated content.

Be that as it may, the output (in Eastonian terms) – the radical decisions – will always be human, and so radicalization is generated from the following path of exposure.

1) The person interacts “normally” in life and online. Choices made there are reflected in search histories, websites and other visits.

2) Data generated finds itself in data aggregators (i.e. contribute to “Big Data” pools).

3) Data aggregators inform radical groups of communities and individuals’ vulnerabilities and psychological predispositions. These “markers” may be social, economic, political and ethnic.

4) Physical and electronic messaging (by community and local political leaders, but by the media as well) begins by reflecting existing beliefs, reinforcing them.

5) Trolls and bots intervene in social media to activate (some old hands in Russia might say “agitate”) individuals. Some trolling might be favourable, other might be unfavourable to an idea. What matters is generating belief and feeling among the target population

6) AI takes over by self-generating messaging on social media, and also by directing and redirecting advertisement, searches, etc. towards “expected” outcomes. Individuals become progressively isolated within the pool of opinion. The outcome of this is that the people start to believe that there is only one stream of dominant opinion, or that a minority problem is ever present.

7) Psychological radicalization is achieved. Within a certain period of time, the individual takes steps that bridge the gap between the consumption of radical messaging and acting out a program propagated by ideologues.

The steps outlined above are the same for any program of “sale” of any idea. Whether it is to buy a car or medication, or changing the world through voting for a candidate, going down in the street, or planting a bomb in a café. Anyone and everyone is vulnerable in the same way. Only the socio-political outcome is different. And our definitions of the act. But that is another story.

Without time to reflect, without reasoned contact with competing or contrary opinion, and yet, even with assurances of perfectly clean data and statistics on a problem, individuals will always side with their preferred biases. The aim of the state is to avoid unnecessary bloodshed or upheavals. But technology provides other states and groups with the power to cause mayhem elsewhere. Technological advances in communications are not merely a double-edged sword; it’s only a blade with no handle, sure to slip from the bloody hand that wields it.

**Aleksandr RAIKOV**, Leading Researcher, Institute of Control Sciences, Russian Academy of Sciences (Moscow, Russia)

***Strong Artificial Intelligence, Its Features and Ethical Principles of Safe Development***

Artificial Intelligence (AI) is currently developing in a digital economy. It increasingly penetrates the socio-humanitarian and industrial sphere, helps to resolve a state and municipal government’s issues.

AI is a technology that enhances a person’s creative possibilities and helps him in his work. AI makes it possible to understand and use the power of the human mind, to get closer to the mystery of the human spirit. However, before the AI was a harmless human’s helper in routine, now it has already become a dangerous competitor for any employee.

However, the AI capabilities are expanding and deepening. It infiltrates deeper into the secrets of the sensual and emotional human sense levels, human’s meditative abilities, as well as the collective unconscious. And with this, features of the next generation of AI - Artificial Super-Intellect (ASI) begin to appear: “Intellect, which is much smarter than the best human mind in almost all areas, including scientific creativity, wisdom and social skills”. With the advent of ASI, its danger to society is not excluded. And on the way of its creation there can meet traps, hit in which is capable of causing irreparable damage to society.



**Aleksandr Raikov**

The first trap is the digitalization trap, when a continuous (analog) image of reality in a computer is replaced by a digit (bits and bytes). And there is no matter how accurately a computer restores a continuous signal to its individual points, the error of computer models accumulates. The digital signal has a limited frequency spectrum. Because of this in particular a number of tasks on supercomputers are solved in months, instead of fractions of a second. And more importantly, such a signal is unable to reflect the full depth of human emotions and feelings, which ultimately can lead to a decrease in the level of culture and spirituality of society.

The second trap is the rationality trap. A human tries to find a rational grain everywhere: at home, at work. He analyzes. Analysis is the division of the whole into the parts. Synthesis is a reverse operation. This is a much more complicated operation that requires the connection of a human's creative abilities. With a purely rational (algorithmic) approach to the synthesis, the possibilities of obtaining good solutions using AI are very limited. As a result, the risk of growing errors in solving vital problems, especially strategic ones, increases.

The third trap is the causality trap. Modern AI uses logic and statistics in its conclusions, which reflects causal connections and parameters correlations. The classical science canons require this. Forecasts are often made on the basis of experience accumulated over a certain period of time and established trends in the development of events. But life often offers to solve unexpected problems and in completely new circumstances, it behaves in a different illogical way, which modern AI cannot realize.

The fourth trap is phenomenological. Natural human intelligence is enhanced by emotions and feelings. There are also deeper levels of consciousness – meditative and transcendental. An even more complex phenomenon is the collective unconscious. These phenomena are sources of mysterious insight - instant comprehension of the whole, afflatus and insight of the human mind.

They are characterized by complete nonformalizability, illogicality, intensity, duration, objectivity, tonality, etc. Traditional AI cannot yet embrace these levels of consciousness.

The listed traps (the list is incomplete) are due to the stereotypes of conducting scientific research, insufficient coverage of disciplines and the lack of relevant international collaborations. But sooner or later, these limitations in the development of AI will be removed and the ASI will enter the arena. This requires the development of interdisciplinary basic research, a more critical attitude to digitalization, immersion of information models in infinite-dimensional spaces, removal of contradictions between quantum mechanics and the theory of relativity, appeal to the potential of Space and much more. It is also necessary to start teaching people the future!

The possible dangers associated with the future appearance of ASI make you think about their warning. Scientists, government and public figures, professionals and experts develop ethical principles that should be followed in order for the development of AI to move in a safe and moral manner. For this, in particular, the well-known Asilomar AI Principles have already been formed.

To support these principles, as well as taking into account the public and state significance of the issue of possible risks of ASI development, which may increase in unpredictable and abrupt manner, we consider it reational to offer the government authorities, the scientific and expert-analytical community the following minimum set of principles in the development of ASI:

1. The most important goal of ASI development should not only be to increase the efficiency of labor, but also to master the deep levels of human consciousness (emotions, feelings, meditative layers of consciousness), improve civic participation, strengthen the consideration of the socio-humanitarian factor.

2. ICI should be absolutely, 100% safe for human, including environmental and moral cleanliness, regardless of where it is used: government, robotics, advertising, manufacturing, intelligent assistants, etc.

3. Creating a comfortable and transparent network environment for the constant growth of the effectiveness of virtual (network) cooperation in various fields of activity on the basis of ASI authorities, scientists, teachers, engineers, students, schoolchildren and other sectors of society.

4. During initiating projects, concluding contracts for creating ASI systems, the personal responsibility of managers, scientists, developers and engineers for possible harm and the procedure for compensation for possible damage to production, the human hostel, should be determined.

5. The functioning of autonomous collective and individual systems of ASI should not contradict ethical and universal norms and values, established canons of freedom of conscience and religion.

6. The development of ASI should not threaten the employment of any person. Any introduction of ASI should be accompanied by an increase in people's satisfaction, an increase in the number of modified jobs.

7. A human should always have the right of responsible choice: to make a decision independently or to entrust it to the ASI system, and any such system should be designed taking into account that the human has the opportunity to intervene in the decision process implemented by this system.

8. Absolutely all the risks of the ASI development should be controlled and preempted by appropriate organizational, scientific and engineering techniques.

9. In the development of systems based on AI methods, special attention should be paid to the construction of the ASI (including, strong, general, collective, cognitive, etc.). It is exactly one

will be able to manifest the superhuman possibilities of the sensory, emotional and transcendental levels of consciousness.

10. Systems of Artificial Super-Intellect, including collective ones capable of autonomous self-controlled behavior, self-development and self-reproduction should be under especially strict control of the person.

**Pierre-Emmanuel THOMANN**, President/Founder, Eurocontinent (Brussels, Belgium)

***Artificial Intelligence and Geopolitics. Who controls AI controls the world! What Role for Europe?***

AI will contribute to change the power hierarchy and the international order in the 21st century, accelerating the dynamics in which new technology and power mutually reinforce each other. AI has the potential to transform the paradigms of geopolitics through new relationships between territories, spatio-temporal dimensions and immateriality. Geopolitics is characterized today by the rivalry between states, alliance of states or private actors for the control of different spaces: ground, sea, air and cyberspace. The emergence of AI is adding a new dimension, that is space-time dominance. Alliances of states able to exert full spectrum dominance in the different spaces like ground, sea, air, cyberspace and space-time (AI) will be able to have a decisive geopolitical advantage, because the mastery of territory and time in the service of a political objective is a decisive advantage and a central element of sovereignty.

AI will be influencing geopolitics at a tactical level, but also on a more strategic and long-term level.

The malicious use of AI at tactical level can have direct effects to change the balance of power in a conflict for geopolitical influence between rival states.

On a more long-term perspective and contrary to idea that digital revolution and the use of AI programmes necessarily triggers political and economic decentralization, it is actually possible that AI provokes a global movement of centralization of power for the benefit of a handful of states and private actors. AI could as a result reinforce monopolies by one or groups of states and create new geopolitical hierarchies, and new digital empires. Malicious use of AI can therefore have the potential to destabilize the system of international relations.

On a more strategic level, the introduction of AI might lead to reinforced competition between actors for full spectrum dominance, a combination of ground, sea, air, cyberspace and space-time (AI) dominance, and result in the transformation of the global geopolitical configuration.

AI research programmes need accumulation of data to be developed and this is why big data is therefore the fuel to AI. The geopolitical balance will probably change between actors and states possessing AI capacity and big data sovereignty and those who do not possess technological sovereignty and are dependant on other states or private actors.

The issue of AI and big data mastery is related to the question of historical memory, identity, education and ultimately control on populations (mind and behaviour) and states that will be under the pressure of extraterritorial influence and malicious geopolitical and transnational strategies.

The analysis of big data based on the accumulation of information on citizens from contents of social media (Twitter, Facebook, LinkedIn...), facial recognition systems but also digital libraries (digitalization of books and film archives), historical and diplomatic archives, and satellite imagery

and Geographical information systems (GIS) will be enhanced by the use of AI. It will add the space-time dimension in the geopolitical arena.

The states or private actors who are able to combine the storage of big data (In clouds and databases) and dominance for the mastery of AI programmes to analyse these data will be more powerful than other actors. A loss of sovereignty is what is awaiting those states who do not possess digital sovereignty (control of their own big data) and artificial intelligence knowledge. When big data (cloud infrastructures) of a state is outsourced in another country, they will risk losing control in the long term of their own historical, scientific, cultural and civilisational memory. Introduction of AI in Geographical Information Systems (GIS) and mapping will also facilitate the control of state territory for spatial planning. This geographical information can also be used by external actors (states or terrorists) for geopolitical objectives through the disturbance of infrastructures (water, transport, pipelines...) or influence and reorientation of economic, energetic and demographic flows. If states and nations will no longer have the full control of their own historical memory and full control on their own geographical space, they will be losing control of their destiny in space and time (this is why AI introduces the space-time dimension of geopolitics).

The populations as a whole or minorities of these countries could be manipulated by more powerful actors possessing AI and big data dominance through the construction of "fake" or "self-fulfilling" predictive analytics methods and scenarios or reinterpretation of historical narratives based on these big data and AI. The use of communication war techniques associated with dissemination through AI programmes of new historical and geopolitical representations aimed at manipulating populations would lead to an effective mobilization of peoples and achieve new geopolitical objectives.

There is therefore is a risk of colonization of minds of citizens in these states unable to master big data and IR programmes to perform data mining and new analysis and research in an autonomous capacity. This will reinforce the inability to think independently for these citizens and they will more subject to manipulation. As result, they could easily change political loyalty for external geopolitical and ideological visions of the world imposed by poles of power possessing full spectrum dominance and to which their state or nation will be subordinated.

Facing the risk of strengthening geopolitical imbalances due to unequal access to AI, it is necessary to seek through international cooperation for a more balanced distribution of AI research results with common international platforms.

Increased international cooperation is necessary to promote open access to digitized Big Data international platforms and centers since they constitute the main fuel for research projects using AI and AI international algorithms libraries. This could also foster the development of the poorest countries. The use of artificial intelligence could be an opportunity for countries whose development is lagging behind to make a technological leapfrogging and avoiding former obsolete industrial development phases.

The principle of sovereignty of states should be reaffirmed in order to contain the political interference facilitated by the malicious use of artificial intelligence programs by some states looking for supremacy through dominance of the AI technology. The GAFA (Google, Apple, Facebook, Amazon) should not monopolize AI research and States could negotiate international legislation to secure the most possible open access of AI databases.





**Pierre-Emmanuel Thomann**

It also necessary to warn of the dangers of the extraterritorial principles with regard to access to private data, in order to strengthen the protection of citizens, private companies and the sovereignty of states against malicious use of Big Data and AI programmes.

The European Union and its member states are promoting an ethical approach to AI. If this posture is useful, it is not enough. EU and international partners will be able to promote the ethical dimension of AI only if it reaches a position of strength and sovereignty and not a position of dependence and weakness. It would be useful for UNESCO member states to defend this point through the largest possible alliances but also with the least developed countries in order to promote a global balanced access to AI especially Africa and South America.

In order to promote an equitable distribution of the benefits and benefits of the use of AI and to avoid the emergence of dominant positions in the world in terms of mastery of artificial intelligence,

1) Multilateral cooperation between international organizations on IA research projects, under the coordination of UNESCO, could be promoted, particularly between the EU and the Eurasian Economic Union, the European Union and the Shanghai Cooperation Organization, the EU and Organization of African Unity (OAU);

2) Bilateral cooperation between the EU and Russia, especially within the principle of "selective cooperation".

In order to avoid the emergence of any AI domination from a group of countries, the EU and Russia could work together to maintain the widest possible international cooperation in artificial intelligence

The issue of cultural diversity, and the non-discrimination of languages should be a central theme in the development of educational programs using AI in the primary, secondary and higher education sector

The development of Cloud storage according to the principle of the sovereignty of each State for its own Data could be promoted. Each state should be in a position to build databases on spatial data and information on the population on its own territory, in order to maintain control over one's destiny and sovereignty.

**Erik VLAEMINCK**, Researcher, University of Edinburgh; Research Associate, International Cultural Relations Ltd (London, UK)

*Culture in the New Technological Paradigm: From Weaponization to Valorization*

Throughout history, changes in technology have impacted societies and their peoples all over the world in the most thorough ways, often for the better, but also for the worse. The latest development in the sphere of technology and communication are not an exception to this paradigm. From the digitalisation of our economies, the rise of social media to the advances in the field of machine learning and AI, the impact on people's daily life is tremendous and most probably still in its initial phase.

Besides the many benefits, among which interconnectivity and the partial erasure of (geographical) boundaries between people, societies and economies, this new technological paradigm has also brought various challenges and threats to our societies and democratic institutions, well-exemplified by the dissemination of propaganda and fake news to the hacking of elections and manipulation of political identities on a global scale. The future advancements in the field of AI might worsen these threats considerably as state and non-state actors with bad intentions might turn against society in the pursuit of political interests. In order to counter these potential threats, it will be of importance to conduct more research and advocate for international cooperation.



**Olga Polunina, Erik Vlaeminck**

This paper builds upon a constructivist view on culture and explores the role of culture, arts and various cultural phenomena in the new technological paradigm. After introducing the concept of culture and pointing at the interrelation between culture and global (political) processes, this paper enquires into the interrelation between culture and the malicious of new technologies.

This paper contends that more research should be conducted on the interrelation between culture and the malicious use of new technologies. A range of examples will be provided of how cultural identities, values and beliefs have become the primary target of actors who pursue 'cognitive hacking' on our cultural and social identities. The use of AI-driven technologies and their relation to identity-based conflicts will be scrutinised. As the future developments in the area might worsen these processes, it is important to take the risks of following potential threats into account: the incitement of a global culture war, the manipulation and rewriting of cultural memory, cognitive framing through cultural products, and cultural (social) engineering on a mass scale.

At the same time, it will be important to think about how culture (and arts) can protect us against these same threats as well as toward the building of a sustainable future. Culture and the concept of cultural relations could take a more central role in the strategic efforts to counter propaganda and attempts of psychological operations. An ethical approach should stand central in these efforts as they run the risk of politicisation themselves.

Culture and particularly the arts have also a role to play in the fostering of intercultural dialogue on the future digital threats and opportunities as well as in relation to prevention and strategic advocacy. We should think about how culture, arts and the critical humanities could play a more

central role in education and media literacy. In this way, culture could become a shield against threats and part of a human-centered approach to the prospects of a hyper-rationalised future.

Continuous efforts should be made to advocate against the risk of an increasing cultural and creative divide between those who are able follow the progress and those who will be left behind (and will become subsequently more vulnerable). Similar attention should go out to the promotion of a more inclusive AI which avoids social and cultural biases in relation to race, class and gender.

Overall, this paper points at the utter necessity to start talking about culture in the debate centering on the new technological paradigm. In order to be able to counter the malicious use of AI-driven cultural warfare, it will be important to work towards an integrated and human-centered approach to education. We should similarly take measures to protect the foreseeable victims of what might become a cultural and creative divide in our societies. In addition, it will be important to invest in research dealing with the politicisation of cultural products on a mass scale and their implications on the human brain.

Advocacy for autonomous art and artistic freedom should stand central in these efforts. Given the fact that culture, as a social construct, has its own rules (it doesn't follow the rules of money and power), the implementation of cultural factors will require long-term thinking. Therefore, the involvement of large international organisations is crucial. More particularly, culture could become a separate and primary vector in our thinking about sustainability and democratisation.

\*\*\*

Discussion of the problems of malicious use of AI continued on June 14 at the research seminar "Artificial Intelligence and Challenges to International Psychological Security". The seminar was organized by the [Centre for Euro Atlantic Studies and International Security](#) at the [Diplomatic Academy](#) of the MFA of Russia and [International Centre for Social and Political Studies and Consulting with the academic support of the European-Russian Communication Management Network](#) and the [Department of International Security and Foreign Policy of Russia, Russian Presidential Academy of National Economy and Public Administration](#).



Participants of the seminar in Diplomatic Academy

The participants of the seminar adopted a final document aimed at explaining to the authorities and civil society institutions the threats associated with the fall of AI tools into the hands of criminal actors.

### ***Literature***

#### Academic journal articles, monograph chapters, research papers

Alexander Raikov. Accelerating technology for self-organising networked democracy // *Futures*. Vol. 103, 2018. – P. 17 – 26.

Darya Bazarkina. L'intelligenza artificiale nella propaganda terroristica. L'analisi. *L'Eurispes.it*. 10 Maggio (May) 2019. <<https://www.leurispes.it/intelligenza-artificiale-nella-propaganda-terroristica-oggi-e-domani/>> [Accessed 17 May 2019].

Darya Yu. Bazarkina, Evgeny N. Pashentsev. *Artificial Intelligence and New Threats to International Psychological Security. Russia in Global Affairs. N. 1. 2019. Pp.147-170.*

Evgeny Pashentsev. Big data, political communication and terrorist threats: Russian experience in ethical dimension. *Russian Journal of Communication*, 2017, 9:3, 298-299. <<http://www.tandfonline.com/doi/full/10.1080/19409419.2017.1376564>>

Evgeny Pashentsev. Destabilization of Unstable Dynamic Social Equilibriums Through High-Tech Strategic Psychological Warfare // 14th International Conference on Cyber Warfare and Security ICCWS 2019 Hosted By

Stellenbosch University and the CSIR, South Africa, 28 February – 1 March 2019. Ed. Noëlle van der Waag-Cowling and Dr. Louise Leenen. Reading, UK: Academic Conferences and Publishing International Limited. – P. 322 – 328.

Evgeny Pashentsev. Kachestvennyye sdvigi v sovremennyh tehnologijah i ih vlijanie na otnoshenija Rossii i ES (Qualitative Changes in Modern Technologies and Their Impact on Russia-EU Relations). Public Administration. Electronic Bulletin. Issue 69. August 2018.

*Evgeny Pashentsev*. Sophisticated Technologies in Counteraction to Terrorism in Datafied Society: From Big Data to Artificial Intelligence. In: Understanding the War on Terror: Perspectives, Challenges and Issues (ed. Riku Flanagan). 2019. P. 99-136.

Evgeny Pashentsev. Strategic Communication in Russia – EU Relations under Global Shifts // Strategic Communication in EU-Russia Relations: Tensions, Challenges and Opportunities (edited by Evgeny Pashentsev and Erik Vlaeminck). Moscow: International Center for Socio-Political Studies and Consulting (ICSPSC), with the academic support of the Institute of Contemporary International Studies and the Department of International Security at the Diplomatic Academy of the Ministry of Foreign Affairs of the Russian Federation and European-Russian Communication Management Network (EURUCM Network), 2018. – P. 23 – 96.

Evgeny Pashentsev. Strategic Communication under Rising International Tension: Challenges and Opportunities for the EU and Russia Security. In: What a 'New European Security Deal' Could Mean for the South Caucasus. Ed. Frederic Labarre und George Niculescu. Vienna: Study Group Information, 2018 – P. 55 – 74.

#### Interviews, conference reviews

Antonio Occhiuto. Conference Report "Advanced Technologies and Terrorism. Future Threats: Defence and Prevention" (04 April 2019. Euro-Gulf Information Center HQ, Rome, Italy). Keynote Speech "Advanced Technologies and Terrorism: Future Threats, Defence and Prevention". EGIC. <<https://www.egic.info/report-technologies-terrorism>> [Accessed 17 May 2019].

Darya Bazarkina, Alexander Vorobyev. Artificial Intelligence and the Challenges to International Psychological Security on the Internet Professor Evgeny Pashentsev held a talk at the round table on "International Safer Internet Day. Global Trends and Common Sense" at MIA "Russia Today", February 5, 2019. Russian-Latin American Strategic Studies Organization. <<http://globalstratcom.ru/wp-content/uploads/2019/02/Artificial-Intelligence-and-Challenges-to-International-Psychological-Security-in-Internet-2.pdf>> [Accessed 22 June 2019].

Darya Bazarkina, Alexander Vorobyev. 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. Russian-Latin American Strategic Studies Organization. <<http://globalstratcom.ru/wp-content/uploads/2018/12/MFA-Round-Table-December-2018.pdf>> [Accessed 22 June 2019].

Darya Bazarkina, Diego Jimenez. Advanced Technologies and Psychological Warfare: Focusing on Latin America (Results of conferences, round tables and workshops of Russian researchers in South America) August 27 – September 10, 2018. Russian-Latin American Strategic Studies Organization. <<http://globalstratcom.ru/wp-content/uploads/2017/11/RLASSA-Advanced-Technologies-and-Psychological-Warfare-%D0%BA%D0%BE%D0%BF%D0%B8%D1%8F1.pdf>> [Accessed 22 June 2019].

Darya Bazarkina, Kaleria Kramar. Experts from Six Countries Discussed the Strategic Communication Issues in Russian Presidential Academy. Eurocontinent. <<https://www.eurocontinent.eu/2019/05/experts-from-six-countries-discussed-the-strategic-communication-issues-in-russian-presidential-academy/>> [Accessed 22 June 2019].

Darya Bazarkina, Olga Polunina, Jaivin Van Lingen. Russian Researchers on Strategic Communication in South Africa: Focusing on the Malicious Use of Artificial Intelligence. Russian-Latin American Strategic Studies



Organization. <<http://globalstratcom.ru/wp-content/uploads/2019/03/Summary-of-the-trip-to-South-Africa-Focusing-on-MUAI.pdf>> [Accessed 22 June 2019].

Edwin Audland. Artificial intelligence poses serious global terrorist threat. The Italian Insider. 12 April 2019. <<http://www.italianinsider.it/?q=node/7958>> [Accessed 22 June 2019].

**Darya Bazarkina DSc,**

Professor at the Chair of the International Security and Foreign Policy of Russia, RANEP;  
Research Coordinator on Communication Management and Strategic Communication, International Centre for Social and Political Studies and Consulting

**Mark Smirnov,**

Research Intern of the International Centre for Social and Political Studies and Consulting