# Cyber Factors of Strategic Stability

## How the Advance of AI Can Change the Global Balance of Power

*Pavel A. Karasev*

**Pavel A. Karasev**
Primakov National Research Institute of World Economy and International Relations, Russian Academy of Sciences (IMEMO RAS), Moscow, Russia
Senior Researcher

ORCID: 0000-0002-7316-8491
IstinaResearcherID (IRID): 2712148

E-mail: karpaul@mail.ru
Tel.: +7 (916) 211-5260
Address: 23 Profsoyuznaya Str., Moscow 117997, Russia

The article has been supported by a grant of the Russian Science Foundation, Project 18-18-00463.

DOI: 10.31278/1810-6374-2020-18-3-24-52

**Abstract**
It is hard to disagree with most international relations experts that nuclear arms control is currently undergoing a systemic crisis. Opinions differ on its depth and possible ways out of it. At the same time, many experts consider it necessary to conceptualize the factor of the latest information and communications technologies (ICTs), including artificial intelligence (AI) systems, in the domain of strategic stability. The matter concerns not only the digitalization of nuclear communications, command, control, and intelligence systems (C3I), but also the development and use of lethal autonomous weapons systems (LAWS).

***Keywords:*** strategic stability, ICT environment, new technologies, artificial intelligence, cybersecurity, cyber warfare.

## STRATEGIC STABILITY

The starting point, when the two nuclear superpowers agreed on the same understanding of "strategic stability," was the adoption of the U.S.-Soviet "Joint Statement on Future Negotiations on Nuclear and Space Arms and Further Enhancing Strategic Stability" (Washington, June 1, 1990) (Joint Statement, 1990). According to this document, the essence of strategic stability is a state of strategic relations between the nations when incentives for a nuclear first strike are being removed. Although this pillar of strategic stability was formulated thirty years ago, it is still relevant today.

In recent years, the international situation has deteriorated drastically, and the development of strategic offensive and defensive weapons has accelerated. Besides, new non-nuclear factors of strategic stability have emerged, including those associated with the development of new ICTs and AI systems. All this has led, on the one hand, to discrepancies in assessments of strategic issues, and, on the other hand, to the emergence of ideas and initiatives aimed at "adapting" the concept of strategic stability to new realities. These ideas range from reviewing the notion of strategic stability to discarding it altogether and elaborating a new concept to replace it. For example, the Joint Statement by Russia and China of June 25, 2016, (President of Russia, 2016) refers to the concept of "Global Strategic Stability." It implies a departure from the interpretation of strategic stability as a military category derived from the state of nuclear weapons and its expansion to reflect the full breadth and diversity of the present state of international security.

Despite the fact that today the leading states, without doubt, take into account some of the new factors of strategic stability, their interconnection and impact remain insufficiently conceptualized, which is evidenced by their much generalized assessments at the national level. For example, the U.S. Nuclear Posture Review (NPR-2018) states that the current Nuclear Command, Control and Communications (NC3) system "is now subject to challenges from both aging system components and new, growing 21st century threats. Of particular concern are expanding threats in space and cyber space..."

(Department of Defense, 2018b, p. XIII). Russia also sees that the use of ICTs as information weapons for military-political purposes "poses a threat to international peace, security and strategic stability" (President of Russia, 2013b). It is not clear whether potential threats associated with the implementation of AI and the use of LAWS are implied by these statements.

Current research of interplay between new ICTs and various aspects of strategic stability is, on the one hand, diverse but, on the other hand, quite specific. Some experts focus on cybersecurity issues in general (Unal and Lewis, 2018; Futter, 2016), and others on cyber vulnerabilities of the nuclear triad (Abaimov and Ingram, 2017). At the same time, a number of papers explore the possibilities and threats pertaining to the use of AI in C3I (Fitzpatrick, 2019; Stefanovich, 2019) and implications of lethal autonomous weapons for strategic stability (Kozyulin, 2018; Altmann and Sauer, 2017). Yet another, more general, research venue, which is closely tied to these issues, is the military use of AI (Vilovatykh, 2019). And naturally, research in the field of AI is a broader topic. It seems that all experts have arrived at a general conclusion that the development of cyber and AI capabilities for military use will impact strategic stability. While the assessments of this impact and its extent vary, the results obtained in the abovementioned works by different experts and institutions highlighted the need for the present research. The results are substantial and clearly demonstrate the need to examine the implications of new ICTs for strategic stability from a number of intertwined angles. When analyzing and assessing possible consequences of the introduction of new technologies and AI for maintaining/degrading strategic stability, it is necessary to start by determining the facets of the problem under consideration. The analysis should take into account the concepts providing for the use of strategic offensive arms incorporated in the relevant doctrines, be it the first strike, launch-on-warning, or delayed second strike.

The first facet of the analysis is an assessment of the expected direct effect the use of new ICTs and AI may have on the survivability of strategic offensive arms and incentives for a nuclear first strike.

The second facet is the impact on strategic stability of the "classical" cyberattacks and cyber-cognitive impact, as well as unknown glitches and derivations in the operation of digital systems and AI (the likelihood of their manifestation increases with the complexity and scale of the systems used).

The third facet reflects the threats arising from the properties of the ICT environment and the essence of cyber weapons. This environment is transboundary and anonymous, which makes it very difficult to attribute an attack; and the "components" of cyber weapons—like zero-day vulnerabilities—are readily available.

Finally, it is necessary to review and assess efforts currently being taken by the international expert community (most importantly, UN groups of governmental experts) and their results in view of the reduction of threats to strategic stability posed by new ICTs and AI.

## FIRST FACET: EXPECTED OUTCOMES AND UNEXPECTED EFFECTS

Despite the fact that initiatives to introduce elements of AI into nuclear command, control and communications have not yet been reflected in nuclear doctrines, the wording in other documents suggests that the latest technologies are viewed as a means to increase the effectiveness of strategic potentials as the range of threats keeps expanding.

As far back as 2016, the U.S. National Science and Technology Council released a report titled "Preparing for the Future of Artificial Intelligence" which detailed the U.S.'s views on a number of challenges connected with the development and adoption of AI. Among possible uses of AI it named cybersecurity, where it is "expected to play an increasing role for both defensive (reactive) measures and offensive (proactive) measures" (NSTC, 2016, p.36) and military application, where AI may "play an important role in new systems for protecting people and high-value fixed assets and deterring attacks through non-lethal means" (NSTC 2016, p.39). It is important to note that one of the recommendations laid down by the report was the development of a single, government-wide policy, consistent with international humanitarian law, on autonomous and semi-autonomous weapons, although, as demonstrated by the results of the work of the UN Group

of Governmental Experts (Davis and Verbruggen, 2018, pp.384-385), there is still no consensus on how to apply international humanitarian law (IHL) in this area.

The proposals which would require military application of AI can be found in the 2018 U.S. Nuclear Policy Review: "strengthening protection against cyber threats, strengthening protection against space-based threats, enhancing integrated tactical warning and attack assessment, improving command post and communication links, advancing decision support technology, integrating planning and operations, and reforming governance of the overall NC3 system" (Department of Defense, 2018b, p. XIII). This echoes the tasks for AI outlined in the 2018 U.S. Department of Defense Artificial Intelligence Strategy: "improving situational awareness and decision-making" (Department of Defense, 2018a, p.6). If successful, this would be accomplished through AI analysis of images and extraction of useful information from raw data. The document also noted that the present moment is pivotal, and in order to ensure security and increase U.S. competitiveness it is necessary to seize the initiative in leading the world in the development and adoption of AI (Department of Defense, 2018a, p.17).

The Third Offset Strategy presented in 2017 also contains positive assessments and high expectations from the implementation of AI. The ideology of this Strategy was presented in 2014 by then-U.S. Secretary of Defense Chuck Hagel (Hagel, 2020). Former U.S. Deputy Secretary of Defense Robert Work, in his comments on the Strategy at an event organized by the Center for Strategic and International Studies, said that it was technologically based on five key areas: "autonomous learning systems, human-machine collaborative decision-making, assisted human operations, advanced manned-unmanned systems operations, and network-enabled autonomous weapons and high-speed projectiles" (Ellman, Samp and Coll, 2017, p.3).

As for the Russian plans, various pieces of information are available about the planned use of AI for ensuring the functioning of strategic nuclear forces (see Stefanovich, 2019; Isaev, Filatov, Fyodorov and Grevkov, 2015, p.59). Widely cited is an assessment of the prospects for

AI given by Russian President Vladimir Putin: "Artificial intelligence is the future not only of Russia, but of all humankind. There are tremendous opportunities and threats that are difficult to predict today. Anyone who becomes a leader in this field will be the ruler of the world" (RIA Novosti, 2017). "The National Strategy for the Development of Artificial Intelligence for the Period until 2030" (adopted in October 2019), in addition to the task of developing the industries related to AI, is aimed at overcoming the economic and socio-humanitarian challenges associated with this process, and does not cover the military use of AI. At the same time, one of the principles recognized in the Strategy is security—in particular, "prevention and minimization of the risks of negative consequences of the use of artificial intelligence technologies" (President of Russia, 2019).

Not only Russia and the U.S., but also other leading states have fully realized the expected benefits of AI and are actively developing this sphere. In particular, China is eager to claim leadership in the field of AI development. To this end, in 2017 it adopted a long-term strategy for the period until 2030 (China. State Council, 2017). One of the specifics of the Chinese approach is the concentration of efforts of civilian and military specialists.

It is important to note that at this stage, in spite of all assessments, there is no serious evidence supporting the expected effectiveness of the proposed solutions—in any case, while the general directions of the military use of AI are already outlined, so far there have been almost no specific examples. According to (Kozyulin, 2018), "machine learning and autonomy technologies make it possible to use nuclear weapons (for example, B61-12 low-yield high-precision nuclear bombs) to perform tactical missions and vice versa—to accomplish strategic missions using non-strategic weapons." It is likely that a survivable high-precision hypersonic weapon outfitted with AI would tip the balance in the system of strategic stability. In particular, this may create incentives for launching a first nuclear strike in response to an attack using such weapons. The impact of AI technologies on strategic stability is highly dependent on the specific scenario. A number of studies note (see National Science and Technology Council Committee

on Technology, 2016, p.11; Stefanovich, 2019) the emerging possibility of hitting previously invulnerable targets, as one of the consequences of using AI to increase situational awareness (for an overview of some of the works in this area see (National Science and Technology Council Committee on Technology, 2016. p.71, footnote 15)). For example, it is theorized that nuclear submarines can be detected by autonomous sea drones. Such a decrease in survivability of one of the elements of the nuclear triad will create incentives for a first strike. On the contrary, under launch-on-warning posture (which is adopted in Russia) the improvement of situational awareness of a hypothetical attack (including the quality of sensor data analysis) can strengthen strategic stability, since it would to a certain extent reduce the risk of an erroneous launch-on-warning strike. AI-based expert systems could also help to alleviate the time deficiency by processing more information with greater speed. At the same time AI in itself is not infallible and can be a target for specific attacks—leading to disruptions and degradation of its functionality and to wrong decisions. In their research J. Altmann and F. Sauer arrived at an interesting conclusion that while "speed is undoubtedly a tactical advantage on the battlefield, and humans are slower than machines… when [strategic stability] comes under threat, some remainder of human slowness is a good thing" (2017, p.136).

## SECOND FACET: CYBERATTACKS AND OPERATIONAL ERRORS

At the moment, the international community has not yet developed a universally accepted definition of artificial intelligence. In discussions about possible types of AI, experts distinguish two large categories: "weak" and "strong." They fundamentally differ in their characteristics. Today, in all cases of practical application of AI technologies we can see "weak" or "specialized" AI aimed at performing strictly defined tasks, including image recognition, speech recognition and synthesis, and the work of expert systems. A weak AI does not have an independent goalsetting capability and does not always have safeguards that would allow it to recognize and diagnose mistakes. Strictly speaking, there is no "intelligence" in the weak AI, meaning self-awareness and self-knowledge, but we can talk about the imitation of certain thought

processes. Current technology allows the implementation of a weak AI through the construction of artificial neural networks. From this point of view, the main factors affecting the work of a weak AI (speed of obtaining a result and its quality) are: 1) the ability to develop an adequate (meaning sufficiently accurate) model of the simulated process (in practice it is determined by the number and quality of specialists); 2) the availability of an extensive database of accumulated data (in practice it is determined by introduction and implementation of policies in the field of big data, including the policy of collecting, storing, protecting, processing and transmitting information); 3) the availability of computing power (in practice it is determined by the availability and accessibility of data centers and supercomputers).

The issues connected with the development and functioning of a weak AI are: 1) the difficulty of developing adequate models—both in terms of completeness and accuracy; 2) the difficulty of obtaining quality data for neural network training—"bad" data can disrupt the learning process, and even plant a "cognitive bomb"—an AI backdoor, which would provoke a certain response upon receiving a specific data pattern; 3) vulnerability to malicious cyber actions, and the emergence of hybrid, cyber-cognitive threats.

An important characteristic of "smart" combat systems in which AI elements can be used is the degree of autonomy or, in other words, the degree of combat functionality delegated from a person to the "intellectual component" of the system. Autonomy, or the ability to make algorithm-based independent decisions, can vary from robotization and intellectualization of certain functions while maintaining the key role of an operator, up to complete independence in decision-making and the self-learning ability. The United States proposed (Department of Defense, 2017, pp.13-14) a three-level classification of combat systems by degree of autonomy:

1. Autonomous weapon system—a weapon system that, once activated, can select and engage targets without further intervention by a human operator;
2. Human-supervised autonomous weapon system—an autonomous weapon system that is designed to provide

human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur;

3. Semi-autonomous weapon system—a weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator.

It seems that a degree of autonomy of intelligent combat systems (nuclear communication, command and control systems can be placed among them) reduces the likelihood of some errors (primarily human), but increases the risk of errors of a different nature, in no small part related to errors in receiving, processing and perceiving data.

The threats to artificial intelligence can be technical in nature—they are cyber threats—and cognitive ones which are aimed at the malicious use of flaws in the data processing algorithms of artificial intelligence. It follows that in order to ensure the cyber security of artificial intelligence systems, those methods and means that were developed to protect against cyber threats will not be enough. As M. Fitzpatrick notes, "…while disinformation is a long-standing intelligence and strategic problem that pre-dates the cyber age, the integrity of AI systems is especially vulnerable to it" (2019, p.91). It seems necessary to define a new, narrower, concept for that category of specialized ICT tools that affect consciousness—both natural and artificial. They are "cognitive attack weapons"—a set of means and tools designed to influence natural and artificial (man-made) adaptive algorithms for processing and assimilation of information.

As already mentioned, one of the most important conditions for creating a weak AI is the elimination of "bad" data that can negatively affect the learning process—and introduce "cognitive bookmarks." The threat persists after the learning phase. If attackers, or potential adversaries, have an understanding of the operation of algorithms and knowledge of their limitations, specially prepared data supplied to the "input" of an intelligent system can also lead to distortions in its operation. For example, experiments conducted with AI image

recognition systems have demonstrated how the smallest changes in the processed image can radically affect the result (see Heaven, 2019; Matsakis, 2017; Nguyen, Yosinski and Clune, 2015). It seems that in the case of military use of such systems, errors could amount to human lives.

Taking into account the fact that intelligent systems, including military ones, physically have the form of hardware and software systems, they are also susceptible to cyber threats that could be used by potential adversaries to accomplish their tasks. Such threats are increasing in proportion to the increasing pace of computerization and the introduction of new ICTs into military operations. Available studies (see Unal and Lewis, 2018; Abaimov and Ingram, 2017; Futter, 2016) indicate that nuclear command, control and communications systems are not free from cyber vulnerabilities either. At the same time, due to secrecy, the amount of available information on the systems used in this area is extremely limited, which makes it difficult to assess the real aggregate level of such a threat.

A number of factors are considered when assessing the general state of cybersecurity. These include the degree of use of commercial publicly available products, and/or imported components as well as software; connectivity to public networks—such as the Internet; level of computer literacy and "cyber hygiene" of personnel. A factor of a different nature is the possible existence of so-called "insiders," that is, informants and agents infiltrated into relevant organizations.

It is obvious that cyberattacks constitute a significantly greater threat if the command and control systems use ready-made commercially available and mass-produced components (off the shelf), since the vulnerabilities of these components are widely known and better studied. In addition, such hardware and software could contain specialized undocumented functions (so-called "backdoors"). Finding and identifying them is an extremely difficult task of cybersecurity. Backdoors can be introduced even without the consent of the manufacturer—by using loopholes in the supply chain of components. For example, in 2014, information about the activities of the Office of Tailored Access Operations of the U.S. National Security Agency, which carries out counter-terrorism operations, cyberattacks and espionage,

was leaked. The list of equipment that the Office can access includes servers, workstations, firewalls, routers, mobile phones, telephone lines, and SCADA systems (Spiegel Staff, 2013). This leak was followed by disclosure of the catalogue of the used hardware and software (Appelbaum, Horchert and Stöcker, 2013).

On the contrary, the conduct of cyberattacks against custom systems that are designed to solve specialized problems is a more complicated task. Firstly, they more often than not use closed hardware-software architecture. Secondly, these systems are usually isolated from public networks. Thirdly, such systems often use an outdated element base, where the presence of non-documented functions introduced from outside is minimal. According to a report by the U.S. Government Accountability Office (GAO, 2016, p. 3), in 2016 the average age of systems used by the U.S. Department of Defense, including intercontinental ballistic missile control systems, was 53 years.

However, today these advantages are rapidly disappearing, and the solution of a previously complex task is already theoretically possible. Firstly, new vulnerabilities may be introduced during an upgrade and/or repair of obsolete or defective components. The Russian Defense Ministry has announced that by 2020 the Strategic Missile Forces will have completely switched to modern digital information transfer technologies (Russian Defense Ministry, 2019). According to the latest information (Insinna, 2019), the U.S. is also in the process of upgrading outdated command, control and communications systems used by its Strategic Command.

Secondly, the isolation of the systems from public networks can no longer provide guaranteed protection against cyberattacks. One recent study (Abaimov and Ingram, 2017) pointed to cyber vulnerability of British Vanguard-class strategic nuclear-powered submarines (NPS) and of the launch control systems of Trident II missiles deployed on them. According to the provided information, the control systems of these submarines and some other warships use a modified version of the commercial operating system (OS) Windows XP (MacAskill, 2017). Although it has been officially stated that the potential vulnerabilities of this OS are mitigated due to the isolation of ships and submarines

that are at sea (Allison, 2018), the authors of the report note that there are vectors for cyberattacks that do not require external commands or connection to networks. As it was already mentioned, it is impossible to exclude the presence of agents among the service personnel who can reveal the internal workings, principles of operation and architecture of the installed networks and systems and/or participate in implanting malicious ICT tools—for example, during maintenance work (Abaimov and Ingram, 2017, p.354). That is exactly how the isolated and classified facilities of Iran's nuclear program were infected (Kaspersky daily, 2014). It is assumed that the removable storage medium could be planted by an employee of an enterprise, including during the construction phase (Cherry, 2010).

Likewise, satellite systems cannot be excluded from the list of possible targets for cyberattacks. In 2017, Symantec specialists identified a number of attacks, which were called Operation Thrip (Security Response Attack Investigation Team, 2018). It was targeted, inter alia, at systems that control satellites and monitor telemetry, and this suggests that the possible target of the attacks may have been not only espionage, but also disruption of the spacecraft. With regard to strategic stability, it is known that the missile attack warning systems in Russia and the United States consist of ground- and space-based echelons. As illustrated above, restricted access facilities can be a target for cyberattacks, and thus early warning systems are also at risk. Backdoors and malicious software could be introduced during assembly and testing of a spacecraft and ground stations can be accessed with the help of insiders.

So, on the one hand, a security threat arises due to insufficient knowledge of the impact that the process of intellectualization of weapons can have on strategic stability. On the other hand, the danger is associated with uncertainty about the level of protection of such weapons from cyber threats. For example, a 2018 report by the U.S. Government Accountability Office noted that critical vulnerabilities had been discovered in almost all major arms and military equipment procurement programs that underwent operational tests during the period of 2012-2017 (GAO, 2018. p. 21). Despite the

fact that the report itself does not indicate specific weapon systems, according to *The New York Times*, these include submarines, missiles, cargo rockets, radars, fighter jets, refueling tankers, aircraft carriers, destroyers, satellites, helicopters, and electronic jammers. In an interview (Sanger and Broad, 2018) officials noted that among the affected systems were two elements of the nuclear triad—the Columbia-class submarines and Ground Based Strategic Deterrent missiles, which are being developed as a replacement for the Minuteman III ICBMs.

The above allows us to make a general conclusion that the tendency of developed states to fill their armed forces with autonomous systems and AI elements should be accompanied not only by a significant rethinking of approaches to ensuring their cybersecurity, but also by raising awareness of new threats. Autonomous intelligent systems can theoretically make more informed decisions due to their ability to process a larger amount of information than a human operator. However, they cannot guarantee the correctness of the decisions made—the received data can be misinterpreted, and the systems can be hacked. A "hybrid" cyberattack can change or even replace data transmitted to an operator, who in this case will incorrectly assess the situation and take erroneous action. It means that manipulation of decision-making processes may be the result of successive cyber actions (to penetrate systems) and/or information attacks (to affect data processing).

## THIRD FACET: PROPERTIES OF THE ENVIRONMENT

ICT tools are increasingly attractive for solving military-political tasks due to a number of their features related to the properties of the ICT environment. First of all, they are transboundary and anonymous, which complicates the attribution of an attack. This means that an attack can come from any geographic location with a connection to the global network, and even if it is detected, the existing technical means cannot quickly and accurately determine its source—and in the field of strategic stability, time can be a critical resource (especially under a launch-on-warning strategy). In addition, acquisition of cyber weapons is not a prerogative of great powers, or even states, as actors in politics.

Attacks can be carried out using ordinary personal computers with Internet access. Particularly dangerous are "false flag" cyberattacks, which under certain circumstances (especially during a crisis, or if the targets are part of critical infrastructure or works or installations containing dangerous elements such as dams, dykes and nuclear electrical generating stations) can cause an unintended escalation of the conflict. Carrying out a powerful cyberattack, which is potentially comparable in its effect to weapons of mass destruction—and create prerequisites for the retaliatory use of nuclear weapons—requires more resources and time-consuming preparations by a team of specialists using models of specific targeted systems. According to Kaspersky Lab experts' estimates, the cost of Stuxnet development amounted to $100 million (see Fisher, 2020). At the same time, it is obvious that this would not require the creation of an advanced industrial and scientific base. Thus, cyber weapons have relatively low barriers for entry, and this can lead to a significant rise in the number of actors capable of deploying and using them. Theoretically, they may include not only states, but also terrorist organizations, and organized criminal groups. It seems that even a small country can become a great cyber power.

According to some reports (Valentino-DeVries and Yadron, 2015), a significant number of states are already developing various ICT tools for military and political purposes, and there is no doubt that their number will increase over time. This means that the militarization of the ICT environment, as a new area of interstate confrontation, is already a fait accompli. The so-called cyber weapons, or cyber capabilities, are, in essence, specialized software—and this directly affects the possibility of its proliferation. Malicious software and/or its components (like the so-called zero-day vulnerabilities) can be purchased, and specialists can be hired (on cyber vulnerabilities and hackers' tools market see Ablon, Libicki and Abler, 2014).

States are interested in creating a database of vulnerabilities that can be used to "produce" cyber weapons. For example, the so-called Vulnerabilities Equities Process has been functioning in the United States since 2008 as part of the relevant policy. It is designed to facilitate decisions regarding the disclosure of new and unknown cyber

vulnerabilities to the general public. The assessment commission may decide to conceal information if there is a direct interest in using the discovered vulnerability for surveillance, law enforcement or national security purposes carried out under the law (White House, 2017, p.1). The fact that one of the fundamental documents used as a basis for the creation of the Process was the Joint Plan for the Coordination and Application of Offensive Capabilities to Defend U.S. Information Systems can provide some insight into the nature of intended use for national security purposes (White House, 2017, p.1).

At the same time, it can be stated that the proliferation of cyber weapons and their components is a practically uncontrolled process, which is developing outside the existing system of international security. "The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies" (Wassenaar Arrangement, 1996) could have played a significant role in this. In 2013, the participating states decided to expand the control lists and introduced a number of restrictions regarding "network communications surveillance systems" and "intrusion software" (Wassenaar Arrangement, 2019b, p.47), that is, in fact, a component of cyber weapons. However, in 2017, two exceptions were introduced regarding the process of distribution and transfer of intrusion software technologies (Wassenaar Arrangement, 2017, p.2). The first concerns the "vulnerability disclosure"—"the process of identifying, reporting or communicating a vulnerability to, or analyzing a vulnerability with, individuals or organizations responsible for conducting or coordinating remediation for the purpose of resolving the vulnerability" (Wassenaar Arrangement, 2019a, p.236). The second is "cyber incident response," that is, "the process of exchanging necessary information on a cybersecurity incident with individuals or organizations responsible for conducting or coordinating remediation to address the cybersecurity incident" (Wassenaar Arrangement, 2019a, p.219). Such exceptions create prerequisites for quasi-legal proliferation of cyber weapons technologies.

Another factor that has a detrimental effect on strategic stability is the widespread use of the so-called "public attribution." States that are

perceived as authority in the international arena make statements about the responsibility of a specific country for an incident in cyberspace, and those statements are taken by some members of the international community for granted without any evidence. At that, no international legal mechanisms have yet been created for a legitimate investigation and legal review of incidents in cyberspace, including those that may be considered an armed attack. An escalation of tensions can become an extremely negative background that can be used by third parties in their interests—including undermining strategic stability. In this regard, the following thesis adopted in the U.S. National Cyber Strategy is especially dangerous. To deter opponents from malicious acts in cyberspace that threaten U.S. national interests, allies or partners, all the instruments of national power are available, including military force (White House, 2018, p.21). It is noteworthy that, in accordance with the same Strategy, Russia, China, Iran, and North Korea were identified as the main opponents. One cannot be dismissive of the nuclear status of the first two countries—and, accordingly, the possibility of nuclear escalation.

If we agree that a small country, or even a non-state actor, can become a great cyber power, and at the same time recognize the influence of the ICT environment on nuclear command, control and communications systems, it would turn out that the entities that are not bound by the spirit and letter of relevant arrangements are indirectly included in the strategic stability equation. Uncontrolled proliferation of military cyber capabilities is thus detrimental to strategic stability. The rising level of uncertainty coupled with a lack of agreed technical framework that would provide for fast and precise attribution of cyberattacks, as well as political "appointment" of the perpetrator through the use of public attribution has a potential to escalate tensions in international relations and further degrade strategic stability.

## EFFORTS TO DE-ESCALATE AND CONTROL

While previously the path to strategic de-escalation and control mainly ran along the lines of bilateral relations, it would seem that today, to overcome the challenges and threats identified in the previous sections,

it has to go through wide international discussion. Not all challenges posed by the use of AI and cyber technologies in the military sphere can be fully resolved at the level of individual countries. Unfortunately, despite the general concern about potential conflicts in the ICT environment and the use of lethal autonomous weapons systems (LAWS), there are significant political disagreements between leading states over how to prevent such conflicts.

The discussion of issues related to LAWS and international information security has been going on at the UN for years. The UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (GGE) held five sessions in 2004-2005, 2009-2010, 2012-2013, 2014-2015, and 2016-2017 (for a summary of GGE efforts and achievements see Boiko, 2016). Among significant achievements of the GGE is the recognition of the applicability of international law in the ICT environment, as well as the report of the fourth GGE (the Group included representatives from Belarus, Brazil, Ghana, Germany, Egypt, Israel, Spain, Kenya, China, Colombia, Malaysia, Mexico, Pakistan, the Republic of Korea, the Russian Federation, the United Kingdom, the USA, France, Estonia, and Japan), which presented norms, rules and principles for the responsible behavior of states in the ICT environment. The following norms take into consideration the issue of critical infrastructure protection and therefore could be applied to the sphere of strategic stability:

f) A state should not conduct or knowingly support ICT activity contrary to its obligations under international law that intentionally damages critical infrastructure or otherwise impairs the use and operation of critical infrastructure to provide services to the public;

g) States should take appropriate measures to protect their critical infrastructure from ICT threats, taking into account General Assembly Resolution 58/199 on the creation of a global culture of cybersecurity and the protection of critical information infrastructures, and other relevant resolutions. (United Nations, 2015, p. 8)

But these norms, rules and principles are not binding, and fall into the category of "soft law." Moreover, the experts did not produce a final

understanding of how to apply these norms and rules, and, therefore, the search for an answer to this question became one of the tasks of the GGE formed to work in 2016-2017. The existing contradictions between the states (for a view on contradictions see Karasev, 2018) did not allow the adoption of a coordinated report. It can even be stated that in 2017 the international community was divided, and a year later two competing associations were created to continue discussing security issues of the ICT environment—the Group of Government Experts and the Open-Ended Group. It remains to be seen how effective these groups will be in moving the discourse on international information security forward.

Issues related to LAWS (they are the closest to the topic of military use of AI) have been discussed at the UN since 2014 under the "Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects (CCW)." The main object of these discussions is the military trend to "delegate" to an AI part of the operator's authority regarding decision-making on the lethal use of weapons. Experts share concerns regarding the threat of violating the principles of warfare laid down in international humanitarian law (principles of proportionality, distinction and humanity), as well as regarding legal complexity of attribution of responsibility for actions committed by AI.

At the Fifth CCW Review Conference held in 2016, the High Contracting Parties decided to establish a Group of Governmental Experts (GGE) on LAWS. In addition to determining the key characteristics of LAWS, the GGE was to look into the applicability of the principles of international humanitarian law, analyze the impact of LAWS on regional and global stability, and identify possible cyber risks to LAWS (United Nations, 2016, p.18).

It seems that the most significant findings of the 2017 GGE report on LAWS are as follows. Firstly, it was noted that the CCW offers an appropriate framework for dealing with the issue of emerging technologies in the field of lethal autonomous weapons systems. Secondly, it was affirmed that international humanitarian law

continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems. Finally, states must ensure accountability for lethal action by any weapon system used by their forces in an armed conflict in accordance with applicable international law, in particular, international humanitarian law (United Nations, 2017, p.4).

In addition to confirming the principles outlined in the 2017 Report, an important result of the GGE meetings in 2018 was ten "Possible Guiding Principles" that could form the basis for legal regulation of LAWS. The principle contained in paragraph "e)" is essential for ensuring the security of LAWS, as it states: "When developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, physical security, appropriate non-physical safeguards (including cyber-security against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered" (United Nations, 2018. pp.4-5). At the same time, we can regard as a significant step backward the refusal of a number of leading states (among them the U.S., Russia, Israel, South Korea, and Australia) to develop and conclude a legally binding document that would prohibit the development and production of LAWS (See: Bergstrom, 2019). Russia's position on LAWS was outlined by the head of the Russian Delegation, V. Yermakov, who stated, in particular, that there are "…doubts about further prospects of the GGE in the absence of functioning samples of such systems, established basic specifications and definitions of the LAWS as well as significant discrepancies in the approaches of the participants in discussions to this matter" (Russian Ministry of Foreign Affairs, 2018).

In general, while some states often express the need to regulate important aspects related to LAWS and the functioning of the ICT environment at the highest level possible, the corresponding processes in the UN are yet to achieve concrete results. In this regard, many see the feasibility of progress only within regional formats. While the divide on the issues of international information security has already taken shape, the process is still far from being complete with regard

to LAWS and AI. This means that bilateral agreements, or new "rules of the road" that could adapt the formula of strategic stability to new factors, may become possible only after the rules of the road are defined within like-minded groups. However, it is worth mentioning that the lack of multilateral agreements and lack of common terminology did not prevent Russia and the U.S. from issuing a "Joint Statement by the Presidents of the United States of America and the Russian Federation on a New Field of Cooperation in Confidence Building" (President of Russia, 2013b) in 2013. Issues concerning the implementation of this document are a subject of separate consideration, but its adoption illustrates another possible path of action for strengthening relations, increasing transparency, and building confidence between the two nations

* * *

Due to the fact that at this stage there is no accumulated experience of applying new technologies in nuclear command, control and communications, an assessment of their impact can only be hypothetical. At the same time, the potential for unintentional escalation is obvious and steps are needed to counter it, at least by recognizing the new threats and moving towards common understanding of their potential, which implies their inclusion in the concept of strategic stability.

AI has an equally great potential to either strengthen or degrade strategic stability. Using it to raise awareness may lower the risk of unintentional launch-on-warning strikes, but at the same time new capabilities that would make previously invulnerable elements of the nuclear triad open to a first strike (nuclear or conventional) will negatively affect strategic stability.

Given the fact that AI and digitalized nuclear command, control and communications systems are prone to adversarial cyber influence, this may affect strategic stability in new ways. Cyber-cognitive attacks against AI are a possibility and could disrupt the decision-making process, which becomes especially dangerous in time-critical operations, as it leaves little time for verification.

Other factors which have a negative effect on strategic stability include the proliferation of offensive military cyber capabilities, which goes largely outside of any arms control frameworks. The growing number of cyber actors coupled with the deficiencies of cyber attribution, and the use of public attribution, introduces great uncertainty into international relations, and, by extension, into the sphere of strategic stability.

Taking into account the difficult political situation and relations between Russia and the United States, the development and adoption of relevant norms and rules of responsible behavior of states at a multilateral level could have become an effective mechanism to reduce the risk of conflicts and resolve interstate conflicts related to the use of AI in the military sphere and the hostile use of ICTs by states. However, in the process of developing such rules for the ICT environment the international community recognized the existing contradictions among different groups of nations; the process of developing similar rules for LAWS has yielded similar results. The lack of enforceable regulation today may result in greater problems in the near future. "In general, the haste of military development of AI can lead to a new arms race and eventual disregard for the norms and principles of international law" (Vilovatykh, 2019, p.189).

Bilateral cooperation on the most pressing issues of common concern in the sphere of strategic stability may be a viable way to overcome its crisis. In particular, priority should be given to the development of recommendations and frameworks for reducing cyber- and cyber-cognitive threats from third parties to nuclear command, control and communications systems, and missile defense.

There is a well-known saying that war is an engine of progress, and indeed, military acquisitions often contributed to economic growth, scientific discoveries and inventions. Peaceful use of atomic energy would hardly have received such development if it had not been backed up by considerable military needs. However it should be taken into account that the situation with AI technologies is opposite—although the military use of automated systems laid the foundation for further research, they have long gone beyond closed laboratories, and today

the global market has become the driver of progress in this area. The McKinsey Global Institute estimates that by 2030 AI will contribute about $13 trillion to the global economy (Bughin et al., 2018, p.3). This means that AI technologies are becoming more accessible to an ever wider range of actors, and it remains unknown how these technologies would be used for malicious purposes. In fact, even today the ICT environment is a place for an invisible war of mathematical models and algorithms, and in the near future the battle of two AIs, albeit "weak" ones, can turn from science fiction to reality.

## References

Abaimov, S. and Ingram, P., 2017. Hacking UK Trident: A Growing Threat. *British American Security Information Council*, June [online]. Available at: <http://www.basicint.org/publications/stanislav-abaimov-paul-ingram-executive-director/2017/hacking-uk-trident-growing-threat> [Accessed 1 June 2020].

Ablon, L., Libicki, M. and Abler, A., 2014. Markets for Cybercrime Tools and Stolen Data. *Santa Monica: RAND Corporation* [online]. Available at: <https://www.rand.org/pubs/research_reports/RR610.html> [Accessed 1 June 2020].

Allison, G., 18 June 2018. Despite CND Claims, Trident Doesn't Run on Windows XP. *UK Defence Journal* [blog]. Available at: <https://ukdefencejournal.org.uk/despite-cnd-claims-trident-doesnt-run-on-windows-xp/> [Accessed 1 June 2020].

Altmann, J. and Sauer, F., 2017. Autonomous Weapon Systems and Strategic Stability. *Survival*, 59(5), pp.117-142.

Appelbaum, J., Horchert, J. and Stöcker, C., 2013. Catalog Advertises NSA Toolbox. *Spiegel Online,* 29 December [online]. Available at: <http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994.html> [Accessed 1 June 2020].

Bergstrom, L., 2019. The United States Should Drop its Opposition to a Killer Robot Treaty. *Bulletin of Atomic Scientists,* 7 November [blog]. Available at: <https://thebulletin.org/2019/11/the-united-states-should-drop-its-opposition-to-a-killer-robot-treaty/> [Accessed 1 June 2020].

Boiko, S., 2016. Gruppa pravitel'stvennyh ekspertov OON po dostizheniyam v sfere informatizatsii i telekommunikatsij v kontekste mezhdunarodnoī bezopasnosti: vzgl'ad iz proshlogo v budushchee [UN Group of Governmental

Experts on Developments in the Field of Information and Telecommunications in the Context of International Security: An Outlook from the Past into the Future]. *International Affairs*, (8), pp.54-71.

Bughin, J., Seong, J., Manyika, J., Chui, M. and Joshi, R., 2018. Notes arom the AI Frontier: Modeling the Impact of AI on the World Economy. *McKinsey Global Institute,* September [online]. Available at: <https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018.ashx> [Accessed 1 June 2020].

Cherry, S., 2010. How Stuxnet Is Rewriting the Cyberterrorism Playbook. *This Week in Technology,* 13 October [podcast]. Available at: <https://spectrum.ieee.org/podcast/telecom/security/how-stuxnet-is-rewriting-the-cyberterrorism-playbook> [Accessed 1 June 2020].

China. State Council, 2017. *State Council's Plan for the Development of New Generation Artificial Intelligence.* [online]. Available at: <http://chinainnovationfunding.eu/dt_testimonials/state-councils-plan-for-the-development-of-new-generation-artificial-intelligence/> [Accessed 1 June 2020].

Davis, I. and Verbruggen, M., 2018. The Convention on Certain Conventional Weapons. In: I. Davis (ed.) *SIPRI Yearbook 2018 Armaments, Disarmament and International Security.* Oxford: Oxford University Press, pp.381-392.

Department of Defense, 2017. Department of Defense Directive 3000.09, 21 Nov. 2012, Incorporating Change 1, *USA. Department of Defense,* 8 May. Available at: <https://fas.org/irp/doddir/dod/d3000_09.pdf> [Accessed 1 June 2020].

Department of Defense, 2018a. Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity. *USA. Department of Defense,* February. Available at: <https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF> [Accessed 1 June 2020].

Department of Defense, 2018b. U.S. Nuclear Posture Review. *USA. Department of Defense,* February. Available at: <https://media.defense.gov/2018/Feb/02/2001872886/-1/-1/1/2018-NUCLEAR-POSTURE-REVIEW-FINAL-REPORT.PDF> [Accessed 1 June 2020].

Ellman, J., Samp, L. and Coll, G., 2017. *Assessing the Third Offset Strategy.* [online]. Washington, DC: CSIS: Washington. Available at: <https://

csis-prod.s3.amazonaws.com/s3fs-public/publication/170302_Ellman_ThirdOffsetStrategySummary_Web.pdf> [Accessed 1 June 2020].

Fisher, D., 2014. Cost of Doing APT Business Dropping. *Threat Post,* 6 February [blog]. Available at: <https://threatpost.com/cost-of-doing-apt-business-dropping/104093#.UvO5yGjQ-jI.twitter> [Accessed 1 June 2020].

Fitzpatrick, M., 2019. Artificial Intelligence and Nuclear Command and Control. *Survival*, 61(3), pp.81-92.

Futter, A., July 2016. Cyber Threats and Nuclear Weapons: New Questions for Command and Control, Security and Strategy. *London: Royal United Services Institute for Defence and Security Studies* [online]. Available at: <https://rusi.org/sites/default/files/cyber_threats_and_nuclear_combined.1.pdf> [Accessed 1 June 2020].

GAO, 2016. Federal Agencies Need to Address Aging Legacy Systems. *Government Accountability Office (GAO)*, May. Available at: <https://www.gao.gov/assets/680/677436.pdf> [Accessed 1 June 2020].

GAO, 2018. Weapon Systems Cybersecurity—DOD Just Beginning to Grapple with Scale of Vulnerabilities. *Government Accountability Office (GAO),* October. Available at: <https://www.gao.gov/assets/700/694913.pdf> [Accessed 1 June 2020].

Hagel, C., 2014. A New Era for the Defense Department. *Defense One*, 18 November [online]. Available at: <https://www.defenseone.com/ideas/2014/11/new-era-defense-department/99392/> [Accessed 1 June 2020].

Heaven, D., 2019. Why Deep-Learning AIs Are So Easy to Fool. *Nature*, 574(7777), pp.163-166.

Insinna, V., 2019. The US Nuclear Forces' Dr. Strangelove-Era Messaging System Finally Got Rid of Its Floppy Disks. *C4ISRNET*, 17 October [online]. Available at: <https://www.c4isrnet.com/air/2019/10/17/the-us-nuclear-forces-dr-strangelove-era-messaging-system-finally-got-rid-of-its-floppy-disks/> [Accessed 1 June 2020].

International Committee of the Red Cross, 1977. *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I)*. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/7c4d08d9b287a42141256739003e636b/f6c8b9fee14a77fdc125641e0052b079> [Accessed 1 June 2020].

Isaev, A., Filatov, V., Fyodorov, V. and Grevkov, A., 2015. Model' avtomatizirovannoī sistemy upravleniya material'nym obespecheniem voinskih chastei i soedineniī RVSN v usloviyah razvitiya sistemy material'no-tekhnicheskogo obespecheniya VS RF. [A Model of an Automated Material Support Control System for Military Units and Formations of the Strategic Missile Forces]. *Nauka i voennaya bezopasnost',* 3(3), pp.59-65.

Joint Statement, 1990. Soviet-United States Joint Statement on Future Negotiations on Nuclear and Space Arms and Further Enhancing Strategic Stability. *The American Presidency Project,* 01 June [online]. Available at: <https://www.presidency.ucsb.edu/documents/soviet-united-states-joint-statement-future-negotiations-nuclear-and-space-arms-and> [Accessed 1 June 2020].

Karasev, P., 2018. Militarizatsiya kiberprostranstva [Militarization of Cyberspace]. In: A. Arbatov and N. Bubnova (eds.) *Security and Arms Control 2017–2018: Overcoming the Imbalance of the International Stability.* Moscow: IMEMO RAS, pp.247-259. Available at: <https://www.imemo.ru/files/File/ru/publ/2018/2018_31.pdf> [Accessed 1 June 2020].

*Kaspersky daily,* 2014. Stuxnet: Nachalo [Stuxnet: The Beginning], 18 November. Available at: <https://www.kaspersky.ru/blog/stuxnet-victims-zero/6119/> [Accessed 1 June 2020].

Kozyulin, V., 2018. Tri gruppy ugroz smertonosnyh avtonomnyh sistem [Three Groups of Lethal Autonomous Systems]. *RIAC—Russian International Affairs Council,* 1 November [online]. Available at: <https://russiancouncil.ru/analytics-and-comments/analytics/tri-gruppy-ugroz-smertonosnykh-avtonomnykh-sistem/> [Accessed 1 June 2020].

MacAskill, E., 2017. HMS Queen Elizabeth Could Be Vulnerable to Cyber-Attack. *The Guardian,* 27 June [online]. Available at: <https://www.theguardian.com/technology/2017/jun/27/hms-queen-elizabeth-royal-navy-vulnerable-cyber-attack> [Accessed 1 June 2020].

Matsakis, L., 2017. Researchers Fooled a Google AI into Thinking a Rifle Was a Helicopter. *WIRED*, 20 December [blog]. Available at: <https://www.wired.com/story/researcher-fooled-a-google-ai-into-thinking-a-rifle-was-a-helicopter/> [Accessed 1 June 2020].

National Science, 2015. *National Science and Technology Council Committee on Technology,* p. 71, footnote 15.

Nguyen, A., Yosinski, J. and Clune, J., 2015. Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images. In: *Computer*

*Vision and Pattern Recognition (CVPR '15).* IEEE [online]. Available at: <https://arxiv.org/abs/1412.1897> [Accessed 1 June 2020].

NSTC, 2016. Preparing for the Future of Artificial Intelligence. *National Science and Technology Council Committee on Technology,* October [online]. Available at: <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> [Accessed 1 June 2020].

RIA Novosti, 2017. Putin: lider v sfere iskusstvennogo intellekta stanet vlastelinom mira [Putin: The Leader in the Field of Artificial Intelligence Will Become the Ruler of the World]. *RIA Novosti*, 1 September [online]. Available at: <https://ria.ru/20170901/1501566046.html> [Accessed 1 June 2020].

Russian Defense Ministry, 2019. K 2020 godu RVSN polnost'yu pereīdut na tsyfrovye tehnologii peredachii informatsii [By 2020 the Strategic Missile Forces Will Completely Switch to Digital Information Transfer Technologies]. *Ministerstvo Oborony Rossiīskoi Federatsii,* 4 November [online]. Available at: <https://function.mil.ru/news_page/world/more.htm?id=12260401> [Accessed 1 June 2020].

Russian Ministry of Foreign Affairs, 2018. Statement of the Head of the Russian Federation Delegation, Director of the Department for Non-Proliferation and Arms Control of the Russian Ministry for Foreign Affairs V. Yermakov at the Meeting of the State-Parties of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons on Item 7 of the Agenda "General Exchange of Views", Geneva, 21 November. *Russian Ministry of Foreign Affairs* [online]. Available at: <https://www.mid.ru/web/guest/obycnye-vooruzenia/-/asset_publisher/MlJdOT56NKIk/content/id/3415655?p_p_id=101_INSTANCE_MlJdOT56NKIk&_101_INSTANCE_MlJdOT56NKIk_languageId=en_GB> [Accessed 1 June 2020].

President of Russia, 2013a. *Joint Statement by Presidents of the United States of America and the Russian Federation on a New Field of Cooperation in Confidence Building.* Available at: <http://en.kremlin.ru/supplement/1479> [Accessed 1 June 2020].

President of Russia, 2013b. *Osnovy gosudarstvennoī politiki Rossiīskoī Federacii v oblasti mezhdunarodnoī informacionnoī bezopasnosti na period do 2020 goda* [Fundamentals of the State Policy of the Russian Federation in the Field of International Information Security for the Period until 2020]. Available at: <http://www.scrf.gov.ru/security/information/document114/> [Accessed 1 June 2020].

President of Russia, 2016. *Sovmestnoe zayavlenie Prezidenta Rossiīskoj Federatsii i Predsedatelya Kitaīskoī Narodnoī Respubliki ob ukreplenii global'noi strategicheskoi stabil'nosti* [Joint Statement by the President of the Russian Federation and the Chairman of the People's Republic of China on Strengthening Global Strategic Stability], 25 June. Available at: http://www.kremlin.ru/supplement/5098. [Accessed 1 June 2020].

President of Russia, 2019. *Ukaz Prezidenta RF ot 10 oktyabrya 2019 g. № 490 "O razvitii iskusstvennogo intellekta v Rossiīskoi Federatsii"* [Decree of the President of the Russian Federation as of 10 October 2019 No. 490 "On the Development of Artificial Intelligence in the Russian Federation"]. Available at: <https://www.garant.ru/products/ipo/prime/doc/72738946/> [Accessed 1 June 2020].

Sanger, D. and Broad, W., 2018. New U.S. Weapons Systems Are a Hackers' Bonanza, Investigators Find. *The New York Times,* 10 October [online]. Available at: <https://www.nytimes.com/2018/10/10/us/politics/hackers-pentagon-weapons-systems.html> [Accessed 1 June 2020].

Security Response Attack Investigation Team, 2018. Thrip: Espionage Group Hits Satellite, Telecoms, and Defense Companies. *Threat Intelligence—Symantec blogs,* 19 June [blog]. Available at: <https://www.symantec.com/blogs/threatintelligence/thrip-hits-satellite-telecoms-defense-targets> [Accessed 1 June 2020].

Spiegel Staff, 2013. Documents Reveal Top NSA Hacking Unit. *Spiegel Online*, 29 December. Available at: <http://www.spiegel.de/international/world/the-nsa-uses-powerful-toolbox-in-effort-to-spy-on-global-networks-a-940969.html> [Accessed 1 June 2020].

Stefanovich, D., 2019. Artificial Intelligence and Nuclear Weapons. *RIAC—Russian International Affairs Council,* 6 May [online]. Available at: <https://russiancouncil.ru/en/analytics-and-comments/analytics/artificial-intelligence-and-nuclear-weapons/> [Accessed 1 June 2020].

Wassenaar Arrangement, 1996. *The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies.* July. [online]. Available at: <https://www.wassenaar.org/docs/IE96.html> [Accessed 1 June 2020].

Wassenaar Arrangement, 2017. *The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies,* December. Available at: <https://www.wassenaar.org/app/uploads/2019/consolidated/WA-DOC-17-PUB-006-Public-Docs-Vol.II-2017-List-of-DU-Goods-and-Technologies-and-Munitions-List.pdf > [Accessed 1 June 2020].

Wassenaar Arrangement, 2019a. *The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies. Public Documents,* Vol. II, December [online]. Available at: <https://www.wassenaar.org/app/uploads/2019/12/WA-DOC-19-PUB-002-Public-Docs-Vol-II-2019-List-of-DU-Goods-and-Technologies-and-Munitions-List-Dec-19.pdf> [Accessed 1 June 2020].

Wassenaar Arrangement, 2019b. *The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies. Public Documents,* Vol. IV, December. Available at: <https://www.wassenaar.org/app/uploads/2019/12/WA-DOC-19-PUB-006-Public-Docs-Vol-IV-Background-Docs-and-Plenary-related-and-other-Statements-Dec.-2019.pdf> [Accessed 1 June 2020].

Unal, B. and Lewis, P., 2018. Cybersecurity of Nuclear Weapons Systems: Threats, Vulnerabilities and Consequences. *Chatham House,* 11 January [online]. Available at: <https://www.chathamhouse.org/publication/cybersecurity-nuclear-weapons-systems-threats-vulnerabilities-and-consequences> [Accessed 1 June 2020].

United Nations, 2015. Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, A/70/174. *United Nations, General Assembly,* 22 July. Available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N15/228/35/pdf/N1522835.pdf> [Accessed 1 June 2020].

United Nations, 2016. Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), CCW/CONF.V/2. *United Nations, Convention on Certain Conventional Weapons,* 10 June. Available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G16/117/18/pdf/G1611718.pdf> [Accessed 1 June 2020].

United Nations, 2017. Report of the 2017 Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS), CCW/GGE.1/2017/CRP.1. *United Nations, Convention on Certain Conventional Weapons,* 20 November. Available at: <https://www.unog.ch/80256EDD006B8954/(httpAssets)/B5B99A4D2F8BADF4C12581DF0048E7D0/$file/2017_CCW_GGE.1_2017_CRP.1_Advanced_+corrected.pdf> [Accessed 1 June 2020].

United Nations, 2018. Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, CCW/GGE.1/2018/3. *United Nations, Convention on Certain*

*Conventional Weapons,* 23 October. Available at: <https://undocs.org/ru/CCW/GGE.1/2018/3> [Accessed 1 June 2020].

White House, 2017. Vulnerabilities Equities Policy and Process for the United States Government. *The White House,* 15 November. Available at: <https://www.whitehouse.gov/sites/whitehouse.gov/files/images/External%20-%20Unclassified%20VEP%20Charter%20FINAL.PDF> [Accessed 1 June 2020].

White House, 2018. National Cyber Strategy of the United States of America. *The White House,* September. Available at: <https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf> [Accessed 1 June 2020].

Valentino-DeVries, J. and Yadron, D., 2015. Cataloging the World's Cyberforces. *The Wall Street Journal,* 11 October [online]. Available at: <https://www.wsj.com/articles/cataloging-the-worlds-cyberforces-1444610710> [Accessed 1 June 2020].

Vilovatykh, A., 2019. Iskusstvennyī intellekt kak faktor voennoī politiki budushchego [Artificial Intelligence as a Factor of the Future Military Policy]. *Problemy nacional'noi strategii,* 1(52), pp.177-192.