# Nuclear weapons: towards the artificial intelligence bomb

Siham ELHANAFI
Docteur en études politiques et droit public
Morocco

**Abstract**

Among the most high-profile debates surrounding artificial intelligence, the one on defense seems crucial in this new context, as it raises the question of how artificial intelligence will be used in the field of military nuclear power, and how it could influence the art of warfare. The question of the distribution of military power and the resources it offers is an issue that nuclear states are embracing, and choosing to invest in the development of artificial intelligence to ensure the success of their nuclear missions. The convergence of artificial intelligence and nuclear weapons requires careful consideration of the ethical, security and strategic implications. International discussion and careful regulation are crucial to ensure that the benefits of AI are optimized while minimizing the associated risks.

**Key words:** Artificial intelligence, international security, nuclear weapons, nuclear deterrence.

## Introduction

Artificial intelligence is one of the major upheavals affecting our times. Rarely has technological development created so many opportunities to solve problems, so many changes in usage and so many fears[1],Cars can now drive without drivers, robots are becoming increasingly autonomous, doctors can make more refined diagnoses and lawyers can draw up more precise contracts. We must have the courage to look at things frankly and without fear: we are at the beginning of a new era that will spare no sector. As a result of some of the technological innovations of recent years, the artificial intelligence sector seems destined to occupy a considerable place in the military strategy of countries interested in it[2].Similarly, the development of artificial intelligence in the military field is a reality specifically envisaged by the USA, Russia and China. All three countries have invested in research and development projects in this field. "AI offers the ability to synthesize, analyze, see patterns, obtain information and make predictions in many, many more dimensions than a human can conceive"[3] .

A dozen researchers gathered in the summer of 1956 to attend this seminar devoted to this theme at Dartmouth College in New Hamsphir. One of its organizers, John However, there is no consensus on the definition of AI. Fundamentally, artificial intelligence is based on algorithms created specifically to solve specific problems. These algorithms are used to gather, structure, manipulate, analyze, transmit and exploit data sets of increasing size. The French Atomic Energy and Alternative Energies Commission (CEA) adds to this description, pointing out that "artificial intelligence (AI) consists of a set of algorithms that enable a machine to analyze and make decisions by adapting intelligently to different situations, based on previously collected data."[4]According to Stuart Russel, a professor of artificial intelligence, AI is the search for techniques that enable computers to behave intelligently.[5]

The first contemporary scientist to reflect on the subject of intelligence and computing was Mr. Alan Turing, who began his most famous articles, "Computing Machinery and Intelligence", written in 1950, with the question "Can machines think?" He considers this possibility plausible and answers a list of

---

[1] - Serge SOUDOPLATOFF, « L'intelligence artificielle : l'expertise partout accessible à tous », *Fondation pour l'innovation politique,* fondapol.org, février 2018. p.7.

[2] - Reuben STEFF, « Nuclear Deterrence in a New Age of Disruptive Technologies and Great Power Competition », *Deterrence: Concepts and Approaches for Current and Emerging Threats*, springer,*2020,* p.70.

[3] - William Henry Gates III, known as Bill Gates, born October 28, 1955 in Seattle, is an American computer scientist, entrepreneur and billionaire. He is known as the co-founder of Microsoft in 1975 and its main shareholder until 2014.

[4] - Norlain BERNARD, « les nouvelles technologies et la stratégie nucléaire », initiatives pour le désarmement nucléaire, paris, novembre 2021.

[5] - *Ibid.,*

objections. It was not until a few years later that artificial intelligence emerged as an independent field of scientific research.

McCarthy, first introduced the term "artificial intelligence" into the presentation. The aim was to distinguish himself both from the very fashionable studies of automata, whose theory he felt remained limited, and from research into cybernetics, dominated at the time by the mathematician Norbert Wiener, who reigned over the discipline. McCarthy explains that the challenge is to develop a machine that behaves in a way that would be called intelligent if a human being behaved in this way[6].Two schools of thought emerged: connectionists and symbolists. The former proposes to draw inspiration from nature and use computers to simulate the way the human brain works, thus copying neural networks in a very simplistic way. The latter prefer to stick to logical predicates to achieve their ends. In particular, they rely on "expert systems", which dominated the first developments in artificial intelligence. The reasoning to be followed to solve a problem is deconstructed, codified and transformed into a succession of relatively simple logical actions that the machine carries out very quickly. A final recommendation is proposed to the operator, who is expected to follow it[7].

However, in the strategic field, little research work has been devoted to it, the capabilities of developing AI can clearly find applications. RAND made up for this lack with the publication of a report in April 2018, this work is the first reference and should be the subject of further work to deepen its conclusions[8]. It questioned the ability of artificial intelligence to positively or negatively affect strategic stability. To do this, it looked to the future and proposed four scenarios for the evolution of artificial intelligence up to 2040. In the first case, it assumes that the current situation represents the current level of technology. According to this hypothesis, in the future, AI would logically have little impact on deterrence. The last case assumes the emergence of super-intelligence completely beyond human control.[9]

This subject therefore raises a number of issues that cannot be fully addressed, and we have focused on the most important in order to answer them: **To what extent could the introduction of artificial intelligence-based systems challenge our current nuclear weapons situation? How might advances in artificial intelligence influence international security, with the likelihood of**

---

[6] - Jean Christophe NOEL, « IA : vers une nouvelle révolution militaire ? », *laboratoire de recherche sur la défense,* 2018, *IFRI*, p.17.
[7] - *Ibid.,* p.18.
[8] - Emmanuelle MAITRE, « Dissuasion nucléaire et intelligence artificielle », Bulletin n°60, décembre 2018.
[9] - *Ibid.,*

**nuclear war[10]? What will be the role of human beings, especially leaders, in decision-making and the conduct of military operations? And finally, how can we reconcile artificial intelligence and ethics in the military sphere?[11]**

In this study, we will seek to provide a balanced analysis of the potential impact of the use of AI on the global strategic chessboard and nuclear policy, as well as its possible impact on international security. This topic will also highlight the ethical and security issues raised by technological advances.

In tackling this issue, we are taking a multidisciplinary approach to understanding how artificial intelligence (AI) influences international security, by studying the consequences of its use from both a strategic and geopolitical perspective.

To respond to the above-mentioned problematic, we have deemed it useful to treat this subject in two axes:

## I- AI and nuclear issues

Among the most high-profile debates surrounding artificial intelligence (AI), the one on defense seems to be of the utmost importance in this new context. It involves asking how AI will be used in the nuclear military field, and how it could influence the art of war. The question of the distribution of military power and the resources it offers is an issue which nuclear states are embracing and choosing to invest in the development of artificial intelligence (2), in order to succeed in their nuclear missions (1).

### 1- Successful nuclear missions in the AI era

Although their "excessive" nature is linked to the use of weapons of unrivalled destructive power, and despite specific procedures, rules of engagement and a chain of command and control (C2), the success of nuclear missions also relies on systemic and tactical knowledge, much the same as regular tasks, especially those on the largest scale. Indeed, the success of a nuclear air attack requires surviving enemy defenses, fighter aircraft or ground-to-air missile batteries, just like conventional attacks. Similarly, nuclear-powered ballistic missile submarine (SSBN) patrols draw on all the expertise of submarine missions: use of vessels, tracking and dilution techniques. Finally, for the success of the mission itself, strategic weapons systems rely on conventional means, either to open passage corridors, or to produce diversionary raids. The contribution of artificial

[10] - Michael C. HORWITZ, « Artificial intelligence and nuclear stability, the impact of intelligence on strategic stability and nuclear Risk », Euro-*Atlantic, volume I, perspectives, SIPRI*, May.2019, p.79.

[11] -Bernard NORLAIN , préambule – l'intelligence artificielle, une rupture stratégique, dans revue défense nationale 2019/5 (n° 820), p.20.

intelligence to conventional military systems can also enhance the effectiveness of nuclear missions[12].

Indeed, the effectiveness of nuclear deterrence relies on a triple credibility: political (the decision-maker, the President's affirmative will), technological and operational (the weapons must be reliable and their destructive potential guaranteed). Once the political decision has been taken, a chain of command is required. It also relies on the dialectic of the uncertainty of political information, guaranteeing the President's freedom of action in the face of potential opponents in the event of a crisis. The stable or unstable nature of artificial intelligence in the deterrence relationship therefore seems difficult to define. Artificial intelligence is indeed conducive to the implementation of nuclear strike capabilities, but it also enables other applications to improve the effectiveness of the defense system. In 1986, the Rafale A, a technological prototype aircraft, achieved an incredible industrial and operational adventure. Today, the aircraft is supported by the A330 MRTT[13]. It's a key element in protecting France's sovereignty and contributing directly to its economy[14]. However, enabling powers that do not invest in artificial intelligence, or refuse to integrate it into functions not related to nuclear weapons systems (C2, intelligence, missile defense systems, etc.), risk seeing certain tactics at risk. Its deterrent service loses credibility and risks weakening deterrence itself, particularly in the face of opponents with additional offensive capabilities in cyber and atmospheric space. In short, compared with human supervision, a poor understanding of the enemy's rules of engagement (principles of use, nuclear thresholds) and the effective authorization level of artificial intelligence can lead to self-dissuasion, in the face of an allegedly unstable system, leading to uncontrolled evolutions. In this case, the ambiguity of the actual role of artificial intelligence on the part of the adversary will play an additional role in deterrence.

## 2 - **Artificial intelligence and nuclear delivery platforms**

An autonomous nuclear delivery platform would be an autonomous version of a nuclear-armed fighter, bomber or submarine. A country might fear that its communication lines could be hacked in the event of a crisis. While programmed autonomous systems could potentially be invulnerable to such interference, in general, autonomous nuclear delivery platforms seem risky. The risks associated

---

[12] -Pierre REAL, "L'intelligence artificielle et ses applications : un défi stratégique pour la France," Un monde en turbulence - Regards du CHEM 2019 - 68e session.

[13] - The Airbus A330 Multi Role Tanker Transport (MRTT) is a wide-body military transport and refueling aircraft developed and built by Airbus Defence and Space, and derived from the civil Airbus A330.

[14] - Bruno CLERMONT, « Le Rafale, un outil de souveraineté et de puissance au service de la France », *in Revue Défense Nationale*, 2022/2, N° 847, éditions, Comité d'études de Défense Nationale, p.59 -65.

with a nuclear-armed autonomous platform, which would eliminate positive human control over the use of nuclear weapons, seem obvious. Hacking or spoofing can render a system vulnerable to capture or malfunction, even before considering the risk that the fragile nature of an algorithm might lead to malfunction. In any case, it is those countries that feel relatively insecure about their nuclear arsenals that should be most likely to turn to unmanned nuclear delivery platforms. For example, the USA may feel relatively secure in having the most advanced nuclear arsenal in the world. For example, US military chiefs, despite their well-known reluctance to dismiss potential capabilities, have made clear their resistance to arming autonomous vehicles, or even unmanned vehicles, with nuclear weapons. For example, in 2016, General Robin Rand, head of Air Force Global Strike Command, said, "We expect on [the B-21 Raider long-range bomber] to be manned. . . I like the man in the loop; the pilot, the woman in the loop, a lot, especially since we're doing the dual-capable mission with nuclear weapons." However, Russia is generally safe, but fears the advantage of the US conventional military and its nuclear arsenal. Thus, in 2012, Lieutenant General Anatoly Zhikharev, commander of long-range aviation, declared that Russia could field an unmanned nuclear bomber in the 2040s. Russia has also disclosed its discussions on building the Poseidon unmanned nuclear-armed submarine vehicle. It's Russia's relative insecurity compared with the USA that has undoubtedly helped stimulate its interest in these systems. The way in which much weaker, nuclear-armed nations might be more likely to consider autonomous platforms armed with weapons of mass destruction is further demonstrated by discussions of the Democratic People's Republic of Korea (DPRK or North Korea) potentially using drones to deliver chemical or radiological weapons against the Republic of Korea (South Korea)[15].

The greatest risk of nuclear escalation from the use of AI systems may come from the way in which conventional military uses of AI could pressure nuclear powers to adopt unstable launch postures or even strike first in a crisis. However, one of the main advantages of AI is the ability of machines to make judgments faster than people, and operating at machine speed could have advantages on the battlefield, as countries using autonomous systems could outperform those with human operators, and the ability of one committee to potentially win a conflict at machine speed means that another country could also lose at machine speed, and the fear of losing at machine speed could encourage a weaker nuclear power, especially one that isn't confident in its second-strike nuclear capabilities.

In the worst-case scenario, the fear of losing quickly at the start of a conflict could even lead to the first strike of instability, as a country fearing decapitation decides to attack first, with weapons, in an attempt to avoid possible future defeat[16]. Note that there's nothing about systems necessarily being autonomous

[15] - Emmanuel MAITRE, « Dissuasion nucléaire et intelligence artificielle », *op.cit*. p.82.
[16] - *Ibid.,* p.*83.*

that generates instability in this case, it's the increasing speed of warfare. So other potential developments, such as hypersonic weapons, could also generate instability. These postures could lead to accidents and miscalculations as countries once again put their nuclear arsenals on a trigger because of the fear of rapid decapitation, as well as the impact of machine-speed combat at the conventional level on the risk of nuclear escalation, which could be exacerbated by uncertainty over the degree of autonomous operation of systems on the battlefield. Adding the idea of such an autonomous weapon that has already sparked intense ethical and legal debate around the world to determine the extent to which an algorithm can decide who should live or die.[17]

### 3- The race for artificial intelligence

The rise of AI today is similar to the Cold War-era nuclear arms race, or the space race, but not completely identical[18]. Competition in AI is characterized by its complexity, leading to the mobilization of a wide range of disciplines, including sociology and economics, as well as international relations, law and other fields. First of all, let's remember that the functional analysis of AI in international relations became important once the "big players" of this world became aware of its profound and lasting influence on the international order as we understand it. Vladimir Putin's declaration that the world will be ruled by a single AI leader demonstrates a shared mindset among the leaders of the great powers. Only two countries seem to have the potential to emerge victorious from the AI competition now being waged by the USA and China. The United States has officially introduced the idea of AI into the development of its military power since 2014, when the Third Offset Strategy was drawn up [19].

This strategy focuses on achieving technological superiority, particularly over China and Russia, in areas such as robotics, nanotechnology, Big Data and cloud computing, which are closely associated with AI and private sector companies. The integration of the AI and deep learning-based Maven project with Google is in line with the strategy to improve drone intelligence and target detection capabilities.

---

[17] - Ordy BETGA, « Les menaces majeures contre l'humanité en 2020 », *Une Synthèse du Rapport sur les risques catastrophiques globaux 2020 de Global Challenges Fondation*, Août 2020, https://www.ceides.org/wp-content/uploads/2020/08/Les-menaces-majeures-contre-lhumanit%C3%A9-en-2020.pdf , consulté le 30/06/2022.

[18] - Charles THIBOUT, l'intelligence artificielle, une géopolitique des fantasmes. Etudes digitales, 2019, Religiosité technologique, 5 (1), ff10.15122/isbn.978-2-406-09290-2. p.0105ff. Ffhal-03031223.

[19] - Jean-Patrice COL le saint, « Vers une troisième stratégie de compensation (Third Offset Strategy) ? » La note du CESA, Mars 2015. https://www.irsem.fr/data/files/irsem/documents/document/file/1880/hsVers_une_troisieme_strategie.pdf, consulté le 07/03/2024.

The Chinese point of view is quite clear, advances in AI technology, as Beijing has emphasized, are expected to bring about a significant change in the nature and conduct of warfare. Right now, we are essentially in the midst of an unprecedented shift from information warfare to intelligent warfare as defined by China. AI will be the backbone of all warfare and, by mastering its most advanced functionalities, it will be able to dramatically increase the decision-making capabilities of command and troops on the battlefield, while delegating to autonomous machines tasks that humans are currently able to perform with greater efficiency and speed…[20]

Beijing has also put itself in battle order; in July 2017, the Chinese State Council published the development plan for the new generation of artificial intelligence in which AI is presented as "a new focal point of international competition", it is even added that it will be "a strategic technology that will command the future"[21]. As for Vladimir Putin, he announced on September 1, 2017 that this technology hides unimaginable possibilities and threats. He even added that "whoever becomes the leader in this field will be the master of the world". In France, on February 14, 2017, Jean -Yves le Drian, Minister of Defense, asserted that "artificial intelligence is an element of national sovereignty". More recently, on March 16, 2018, Minister of the Armed Forces Florence Parly was present at the launch of the Man Machine Teaming study, which aims to explore the potential of artificial intelligence for combat aviation.

On this occasion, she unveiled a plan to promote the development of AI in the armed forces, including a progressive investment of 100 million euros per year and the recruitment of some 50 additional experts in data science and artificial intelligence. While the military applications of AI are already the subject of controversy, the prospect of AI being used in nuclear and, more generally, strategic weapons systems could fuel lively ethical debate. Works of fiction have extensively described the risks of entrusting a more or less autonomous machine with the task of launching weapons capable of annihilating mankind, as well as the fear that AI can carry out increasingly demanding tasks and, in the long run, outstrip human capabilities.

In short, a global rivalry has begun between the USA, China and Russia for the use of artificial intelligence as a weapon in advanced autonomous systems capable of teaching themselves to perform specific tasks. The advent of the nuclear age will have to take into account the advances and widespread presence of artificial intelligence, in May 2018, the RAND Corporation published a report that

[20] - Charles THIBOUT, « L'intelligence artificielle, une géopolitique des fantasmes », in Etudes digitales, 2019, Religiosité technologique.
[21] - Jean-Pierre DAMIANO, « Réflexions sur les enjeux de l'IA et les questions d'éthique », https://www.researchgate.net/profile/Jean-PierreDamiano , consulté le 17/3/2022.

highlighted the disruptive capacity of artificial intelligence on nuclear and strategic stability[22].

## II - AI: opportunity or threat

As artificial intelligence technologies evolve and become more advanced, they will revolutionize industrial sectors, simplify procedures and could even have an impact on the way we live. The nuclear field is no exception, and it is possible to see the beneficial effects of artificial intelligence in many processes and operations in nuclear and radiological facilities (2). However, it is important to bear in mind the risks involved in delegating decision-making to an autonomous machine (1).

### 1- AI and nuclear risk

AI applied to military nuclear power has recently reappeared in the strategic debate as one of the modalities of technological emergence. The main focus here is on the potential impact of machine learning and autonomy on future weapons systems, particularly command and control equipment, but not only. The important thing in combination with augmentation is that with the availability of ever larger databases in terms of scale and quality, this means that today's artificial intelligence is capable of performing increasingly complex tasks. The rapid advances in artificial intelligence also present a number of dangers. Indeed, ill-intentioned people can use artificial intelligence to develop more complex malware, automate cyber-attacks, exploit model errors and weaknesses, or bypass security features by imitating the behavior of legitimate users.[23]

Experts in artificial intelligence and nuclear security interviewed by researchers from the RAND Corporation consultancy concluded that by 2040, artificial intelligence could undermine the foundations of nuclear deterrence and undermine strategic stability. Some experts believe that, in order to upset the balance, it is not even necessary to integrate artificial intelligence into nuclear weapons; its existence could push any member of the nuclear club to a first strike[24]. Authorities in nuclear countries believe that the possession of nuclear weapons is already a guarantee of security, and they believe that the existence of such weapons can prevent attacks from another country, which may be worried about nuclear strikes on its territory: this is what is known as the principle of

---

[22] - Solène VIZIER, l'intelligence artificielle met en question la dissuasion nucléaire, https://www.idn-france.org/ ,consulté le 15/01/2024.

[23] - Comment l'intelligence artificielle changera la sécurité de l'information et la sécurité informatique dans le monde nucléaire, https://www.iaea.org/fr/bulletin/comment-lintelligence-artificielle-changera-la-securite-de-linformation-et-la-securite-informatique-dans-le-monde-nucleaire , consulté le 29/01/2024.

[24] - « D'ici 2040, l'intelligence artificielle pourrait miner la stabilité nucléaire ». https://fr.sputniknews. Consulté le 18/06/2021.

nuclear deterrence. In addition, there is also the concept of nuclear equality (or nuclear stability): nuclear powers must not attack each other to avoid mutual destruction. Moreover, hybrid warfare, the development of high-precision nuclear weapons, the establishment of the nuclear triad (ballistic missiles, ballistic missile submarines and ballistic missile bombers) and the development of new missile defense systems have complicated these concepts.

Similarly, the experts interviewed by the RAND researchers were divided into several groups: moderates believe that artificial intelligence will enable nuclear weapons systems to be perfected without affecting nuclear deterrence strategies and nuclear parity. In their view, by 2040, the collection and processing of intelligence data on an adversary's readiness to use nuclear weapons, as well as the selection of priority targets, will become a serious problem for artificial intelligence.[25]On the other hand, alarmists claim that, in any case, artificial intelligence poses a threat to nuclear stability. They believe that it is enough for nuclear powers to decide that the artificial intelligence of the enemy's defense system can compensate for their retaliation in the event of nuclear conflict, to abandon the strategy of retaliatory strike and choose a strategy of pre-emptive strike, at the same time starting to increase its nuclear arsenal.

Another group of experts declares that artificial intelligence can be used in nuclear weapons systems, which is just as likely to reinforce nuclear stability as to destroy it. On the one hand, countries with nuclear weapons and "intelligent" defense systems can decide that the use of these weapons will result in mutually guaranteed destruction, which can have a deterrent effect. On the other hand, "smart" systems may give the government a sense of technological superiority to ensure victory in the event of war, in which case nuclear war has become more real.[26]

All the experts interviewed by the RAND Corporation agreed that by 2040, no nuclear-armed country would succeed in creating a doomsday weapon with artificial intelligence - in other words, even if national leaders are eliminated by the enemy's nuclear strike, they can still launch nuclear responses. While we are witnessing a renewed interest in tactical nuclear weapons, which are more likely to be used effectively, greater integration between conventional, nuclear and high-tech systems in defense systems could mean that, in the strict or traditional sense, favors a new concept of warfare that integrates all means in all domains. It is to be feared that the nuclear risk will thus be amplified. Similarly, the impact of directed-energy weapons on nuclear weapons is likely to become a reality, as are hypersonics, artificial intelligence and cyberdefense. In the new strategic context, with the intensification of confrontation in digital technology, artificial intelligence, cyberspace and outer space, major wars and so-called high-intensity

[25] - *Ibid.,*
[26] - *Ibid.,*

conflicts are once again becoming more likely. With the real risk of escalation, possibly involving nuclear issues.[27]

**2 - Artificial intelligence and nuclear stability**

The first step in the process leading to the possible use of nuclear weapons is how a nuclear-armed state tries to detect whether another country is launching nuclear weapons, and how it reacts Many countries already automate parts of their nuclear weapons infrastructure, particularly advanced nuclear powers like the USA. This includes early warning, command, control and missile targeting. Advances in artificial intelligence could lead to an expansion in the use of autonomous command and control systems. For example, states might decide to automate additional early warning components, as autonomous systems can detect patterns and pattern changes faster than humans. This could have potential benefits for nuclear safety and stability, as powerful algorithms could give decision-makers more time in a complex environment. In addition, autonomous systems could represent another form of redundancy, ensuring the dissemination of launch orders in the worst-case scenario.[28]

However, the Petrov incident[29] in 1983 clearly illustrates a drawback to fully automated command and control. In this case, the Soviet satellite early warning system reported a false alarm - the launch of five American intercontinental ballistic missiles. No missiles had been launched; Lieutenant-Colonel Stanislav Petrov was the watch officer on duty. It was his job to alert the Soviet leadership to an American attack when the automated systems reported the highest confidence that a missile strike was occurring, he reported a system malfunction rather than a nuclear attack. The risk is that a future incident could lead to escalation, instead of a malfunction report, for two reasons: Firstly, a decision to fully automate early warning would mean that there was no human operator - no Petrov - to prevent a false alarm from escalating. To be fair, however, it seems unlikely that any country would exclude humans from early warning altogether. Secondly, an automation bias could mean that a future Petrov trusts the algorithm and instead reports that an attack is underway.

---

[27] - Bernard NORLAIN, « Les nouvelles technologies et la stratégie nucléaire », *Initiatives pour le désarmement nucléaire, op.cit.*

[28] - Michael C. HOROWITZ, «The impact of intelligence artificial on strategic stability and nuclear risk», *op. cit.,*

[29] - In September 1983, when Soviet surveillance satellites triggered an alert, Petrov decided to inform his superiors that it might have been a false alarm and not a missile attack against the Soviet Union, as indicated by the computerized missile warning system. Consequently, its management decided not to retaliate. The crisis, however, comes at a time of extreme tension between the Soviet Union and the United States, just three weeks after Korean Air Flight 007 was shot down by a Soviet fighter jet and potentially triggered a missile exchange. The fact that Petrov detected this alert as false would avoid a response from both opposing camps.

Although also unlikely, academic research on automation bias suggests that this is a real risk. Automation bias is when humans, lulled into a false sense of security by the repeated success of algorithms, will stop questioning them and trust them entirely. This generates cognitive unloading, where humans become unlikely to question an autonomous system, even in a scenario where an impartial human might recognize that an algorithm-based system is reporting incorrect information[30].

Used to analyze intelligence and observation data, control autonomous observation platforms or automatically identify objects, artificial intelligence will pose a threat to minimal deterrence strategies. It could also encourage the use of conventional precision weapons to target strategic systems. In a crisis context, the existence of such capabilities could encourage an adversary worried about the survivability of its weapons to cross the nuclear threshold first, in a pre-emptive logic. In peacetime, it could at least encourage an actor who believes himself threatened to modernize and expand his arsenal, first to abandon the non-use posture, or to increase his level of vigilance. The automation of early warning and intervention procedures can also be envisaged.[31]

Consequently, the destabilizing nature of artificial intelligence used for such functions can be partially offset by the negotiation of arms control measures.

Most of the experts interviewed by RAND seem to indicate that these new technologies have a negative strategic stability record. As far as AI's contribution to decision support is concerned, opinions differ as to whether humans are prepared to delegate such important decisions to machines in 2040, given the risk of hacking. For most, the initial introduction of these systems seems to be the most dangerous, with the greatest risk of errors and uncontrollable upgrades. Some have suggested possible positive effects: artificial intelligence can be more reasonable and limit the risk of errors in the analysis of early warning system data. Others believe that the value of such intelligence lies in providing more reliable means of verification on arms control and even non-proliferation issues. However, the report highlighted the danger of this mechanism in this case: if they think they can control escalation, then actors can be over-confident and accurately trigger a crisis.

This report provides a useful starting point for thinking about the consequences of the development of AI in the nuclear field. These must take several parameters into account: Firstly, it cannot be immediately determined that the consequences are negative, so if we consider the use of artificial intelligence to assist the

[30] - Michael C. HOROWITZ, « The impact of intelligence artificial on strategic stability and nuclear risk », *op. cit.,* p.80.
[31] - Emmanuel MAITRE, « Dissuasion nucléaire et intelligence artificielle », *op.cit.*

offensive force, we must also consider the role it can play and that the system must be protected[32].

Such observations were made by a working group set up on the same theme by Lawrence Livermore National Laboratory in September 2018. The experts called asked similar questions about the ability to play a major role in decision support. On the other hand, they expressed concern about the impact on strategic stability, as the increased vulnerability of these systems may also pose a threat to them. In this case, cyber protection remains a major challenge for nuclear powers. Indeed, the short- and medium-term vulnerability of strategic capabilities to cyber risk seems established. With the increasing use of AI to implement systems, this threat can only increase. Consequently, it is necessary to integrate this issue right from the architectural design stage, and to continue discussing their possible inclusion in arms control work.[33] Although very difficult in practice, the inclusion of cyber and artificial intelligence in international treaties remains a regular recommendation, so as not to reduce current strategic stability.

In short, stepping up the use of artificial intelligence technologies to enhance cyber security measures at nuclear sites could bring significant benefits, such as improved hazard detection, preventive security measures, reduced human involvement and more effective incident response. By using the benefits of artificial intelligence while managing the associated dangers, companies can significantly strengthen their protection in the face of growing cyber threats.

### 3- The AI and ethics dialectic

"For these laws are either laws of nature or laws of freedom. The science of the former is called physics, that of the latter is called ethics" Immanuel Kant, Foundation of the metaphysics of morals,1785.

The use of artificial intelligence offers many advantages in various fields, but if ethical standards are not put in place, it can reproduce the biases and discriminations present in the real world. What's more, it can fuel divisions and pose a threat to human rights and fundamental freedoms. These versatile technologies are changing the way we work, socialize and live. However, these rapid changes are also giving rise to serious ethical concerns. These stem from the ability of artificial intelligence systems to incorporate bias, be harmful to the climate, pose a threat to human rights, and other negative impacts[34]. AI

---

[32] - *Ibid.,*

[33]- *Ibid.,*

[34] - Gabriela Ramos, « Éthique de l'intelligence artificielle », https://www.unesco.org/fr/artificial-intelligence/recommendation-ethics ,consulté le 15/04/2024.

[35]- Marie des-Neiges RUFFO DE CALABRE, « Avons-nous le choix d'utiliser l'IA en temps de guerre ? »,https://www.ihemi.fr/articles/avons-nous-le-choix-dutiliser-lia-en-temps-de-guerre, consulté le 17/01/202

applications in the military field are already generating controversy, and the idea of AI being used in nuclear weapons and other strategic systems could provoke intense ethical discussions. Numerous fictional works have detailed the potential dangers of delegating the responsibility for launching weapons destructive to humanity to machines that are more or less autonomous. Should we use AI in the military domain without first carrying out an ethical analysis, given the many employment opportunities it offers? Even if there are strategic and legal issues at stake, it remains essential to respect fundamental principles of military ethics such as discrimination and proportionality, as well as the tradition of just war.

What's more, the integration of artificial intelligence raises many existing problems, such as the arms race, proliferation, deterrence, the risk of reusing nuclear weapons, total war, pre-emptive attacks and weapons of mass destruction. The human being must maintain both effective control over the machine and enlightened surveillance based on technical skills and democratic ethical values. In the specific case of military operations, what is the ethical impact of new capabilities linked to automation, autonomy, the use of Big Data and artificial intelligence? Despite their ethically "neutral" potential, is it still possible not to employ them, given the scale of these digital revolution[35]? Are we compelled to use them, even if it's unethical to do so? Fortunately, legal developments have not been neglected. This can be seen in the legal principles that have gradually been put in place to regulate conflicts, including the latest developments in international humanitarian law (IHL), the Geneva Conventions. So, what would be the major risks to ethical compliance associated with automation, autonomy, Big Data and AI? In all these areas, the threat would be all the more serious if they were used to enable lethal capabilities. Overall, the first aspect, shared by Jus in Bello[36] ,is that the verification of a technology's compliance with the constraints of just war theory and IHL, from a legal point of view, will be based on two traditional criteria: discrimination and proportionality. Consequently, if these systems commit targeting errors and kill civilians rather than combatants, this would constitute a definite violation of the principle of discrimination. As for the second criterion, if these systems cause civilian deaths not because of misaiming, but because of excessive strike force, and if collateral damage increases, the system's compliance with the proportionality requirement could be called into question. At this stage, this is all conjecture. However, complying with these criteria alone will not be enough to satisfy other ethical issues[37].

Although the ethical issues of entrusting software with the decision to fire and target pose a risk, the type of weapon that could be used has not been specifically discussed. Recently, however, the USA seems to be seriously considering the use of autonomous artificial intelligence to make decisions about launching nuclear

---

[36] - Law governing the conduct of hostilities (jus in bello).
[37] - Marie-des-Neiges RUFFO DE CALABRE, *op.cit*.

missiles, since the development of nuclear-capable hypersonic missiles. In the field, its main objective will be to align itself with the integration of artificial intelligence into the NC3 macro system. It is assumed that the arrival of artificial intelligence will exclude command and control center actors such as the President of the United States and the National Command Authority. According to this hypothesis, nuclear codes would thus be entrusted to artificial intelligence, according to some, this would improve security by reducing the risk of codes being lost due to the distraction often encountered in presidents. Some will argue that it's crucial to react more quickly to attacks in order to guarantee the possibility of a "second strike". Some will recall a historic event that could have gone unrecognized by all, namely when Lieutenant-Colonel Stanislav Petrov decided to oppose the Oko alert system[38].

[38] -*Ibid.,*

## Conclusion

Today, AI is emerging as a competitor to nuclear technologies, creating a paradox: while it makes nuclear deterrence increasingly obsolete, it also makes it increasingly risky[39]. Artificial intelligence should not be used to suppress human thinking; it would be like taking the soldier away from the battlefield, and then also taking the intelligence analyst and the general away from strategic thinking. As well as being the people who have to take responsibility, which is at the root of the word "responsibility", their individual abilities to perceive, imagine, cope with the unexpected, be creative and surprise will be all the more valuable in future wars, if they are fought by machines that analyze our predictable behaviors[40].

In contrast, artificial intelligence represents a new area of investment and competition between powers, as well as a means for other players to catch up in all areas of conflict. However, it could also lead to a dangerous arms race that would compromise our security. While we wait for robots and super-intelligent computers to enslave us or replace us one day, artificial intelligence is recognized at the highest level of the defense industry as having the potential to offer military superiority to those who can best master its development[41]. Furthermore, ethics are of great importance in the management of complex systems, and are indispensable in the development of intelligent systems. It is essential to ensure that AI is used responsibly and ethically. The incorporation of artificial intelligence (AI) in the nuclear sector offers significant opportunities and challenges:

- For the safety of nuclear facilities, this can be enhanced through constant monitoring and data analysis by AI, to detect any suspicious activity or risk of sabotage.

- The use of artificial intelligence in nuclear weapons raises ethical and security concerns, particularly with regard to the trust placed in automated decisions in the event of a crisis.

- Strategic stability may be called into question by the introduction of artificial intelligence, as it can speed up decision-making processes and pave the way for new forms of offensive strategy. This could increase the chances of conflicts escalating. However, artificial intelligence can be essential for monitoring compliance with non-proliferation commitments by examining vast volumes of data from satellites, sensors and other sources.

---

[39] - Bernard NORLAIN, "Les nouvelles technologies et la stratégie nucléaire", *op.cit.*
[40] - Marie-des-Neiges RUFFO DE CALABRE, « Avons-nous le choix d'utiliser l'IA en temps de guerre ? », *op.cit.*
[41] - Jean- Christophe NOEL, « IA vers une nouvelle révolution militaire », *op.cit.,* p.13.

- The growing use of AI systems raises concerns about their reliability, the need for adequate human control and their vulnerability to cyber-attacks. The technological arms race could be reignited by the use of artificial intelligence (AI) in the military sector, including nuclear weapons, with potentially destabilizing consequences.

In short, the convergence of artificial intelligence and nuclear weapons calls for careful consideration of the ethical, security and strategic implications. International discussion and careful regulation are crucial to ensure that the benefits of AI are optimized while minimizing the risks involved.[42]

[42] - Ahmed BOUTAMO, « L'IA au Carrefour des Géopolitiques Modernes : Naviguer dans les Eaux Troublées des Conflits et des Risques Émergents ; CAS France-Russie », mars 2024, https://www.linkedin.com/in/ahmed-boutamo/recent-activity/articles/ ,consulté le 16/04/2024.

## References

- BERNARD Norlain, « les nouvelles technologies et la stratégie nucléaire », initiatives pour le désarmement nucléaire, paris, novembre 2021.

- BETGA Ordy , « Les menaces majeures contre l'humanité en 2020 », *Une Synthèse du Rapport sur les risques catastrophiques globaux 2020 de Global Challenges Fondation*, Août 2020, https://www.ceides.org/wp-content/uploads/2020/08/Les-menaces-majeures-contre-lhumanit%C3%A9-en-2020.pdf , consulté le 30/06/2022.

- BOUTAMO Ahmed, « L'IA au Carrefour des Géopolitiques Modernes : Naviguer dans les Eaux Troublées des Conflits et des Risques Émergents ; CAS France-Russie », mars 2024, https://www.linkedin.com/in/ahmed-boutamo/recent-activity/articles/ ,consulté le 16/04/2024.

- C. HORWITZ Michael, « Artificial intelligence and nuclear stability, the impact of intelligence on strategic stability and nuclear Risk », Euro-*Atlantic, volume I, perspectives, SIPRI*, May.2019.

- CLERMONT Bruno , « Le Rafale, un outil de souveraineté et de puissance au service de la France », *in Revue Défense Nationale*, 2022/2, N° 847, éditions, Comité d'études de Défense Nationale.

- COL Jean-Patrice le saint, « Vers une troisième stratégie de compensation (Third Offset Strategy) ? » La note du CESA, Mars 2015.

- Comment l'intelligence artificielle changera la sécurité de l'information et la sécurité informatique dans le monde nucléaire, https://www.iaea.org/fr/bulletin/comment-lintelligence-artificielle-changera-la-securite-de-linformation-et-la-securite-informatique-dans-le-monde-nucleaire , consulté le 29/01/2024.

- « D'ici 2040, l'intelligence artificielle pourrait miner la stabilité nucléaire ». https://fr.sputniknews. Consulté le 18/06/2021.

- DAMIANO Jean-Pierre, « Réflexions sur les enjeux de l'IA et les questions d'éthique », https://www.researchgate.net/profile/Jean-PierreDamiano , consulté le 17/3/2022.

- GABRIELA Ramos, « Éthique de l'intelligence artificielle », https://www.unesco.org/fr/artificial-intelligence/recommendation-ethics ,consulté le 15/04/2024.

- NORLAIN Bernard, préambule – l'intelligence artificielle, une rupture stratégique, dans revue défense nationale, N°820,2019/5.

- NOEL Jean Christophe, « IA : vers une nouvelle révolution militaire ? », *laboratoire de recherche sur la défense,* 2018, *IFRI*.

- REAL Pierre, « L'intelligence artificielle et ses applications : un défi stratégique pour la France », Un monde en turbulence - Regards du CHEM 2019 - 68e session.

- RUFFO DE CALABRE Marie-des-Neiges, « Avons-nous le choix d'utiliser l'IA en temps de guerre ? », https://www.ihemi.fr/articles/avons-nous-le-choix-dutiliser-lia-en-temps-de-guerre, consulté le 17/01/2024.

- SOUDOPLATOFF Serge, « L'intelligence artificielle : l'expertise partout accessible à tous », *Fondation pour l'innovation politique,* fondapol.org, février 2018.

- STEFF Reuben, « Nuclear Deterrence in a New Age of Disruptive Technologies and Great Power Competition », *Deterrence: Concepts and Approaches for Current and Emerging Threats*.

- TONIN Matej, *intelligence artificielle : impact sur les forces armées de l'OTAN*, Rapport, 149 STCTTS 19 F rév. Original : anglais | 13 octobre 2019

- THIBOUT Charles, l'intelligence artificielle, une géopolitique des fantasmes. Etudes digitales, 2019, Religiosité technologique, 5 (1), ff10.15122/isbn.978-2-406-09290-2. p.0105ff. Ffhal-03031223.

- VIZIER Solène, l'intelligence artificielle met en question la dissuasion nucléaire, https://www.idn-france.org/ ,consulté le 15/01/2024.