

Nuclear Security in the Age of Artificial Intelligence

Muhammad Sadiq¹

Abstract

The concept of artificial intelligence is fast changing the principles of nuclear security in the world. Nuclear security no longer focuses on the protection of the physical and control of the material, but includes cyber resilience, algorithmic decision-making, and more autonomous systems. AI is already used to contribute to such important areas as predictive maintenance of reactors, anomaly detection in the fuel-cycle facilities, and data fusion to nuclear command, control, and communication networks. Such capabilities will be able to reinforce the early warning, verification, and non-proliferation efforts. Nevertheless, they are also becoming more vulnerable as they are being adopted. Algorithms' effectiveness, data manipulation, adversarial interaction, and automation bias endanger to reduce human control and enhance the chances of error or unintentional growth. With AI spreading among the state and non-state actors, the dynamics of deterrence are increasingly rapid, unpredictable, and unstable.

This paper argues that AI has the potential to serve as a force multiplier in nuclear security, its infiltration of sensitive nuclear systems as it currently stands with a weak ethical, legal, and technical frameworks can be counterproductive to strategic stability. The paper analyzes the impact of AI on nuclear governance, deterrence, and operational planning by means of comparative analysis of great nuclear powers and other significant actors in the region. It outlines the areas of critical governance deficiencies in what is known as the regimes, specifically the Nuclear Non-Proliferation Treaty, the Convention on the Physical Protection of Nuclear Material, and the United Nations framework in general. The conclusion made in the paper is that an efficient AI-nuclear governance has to focus on transparency, human-in-the-loop control, and algorithmic accountability to make sure that AI enhances, and does not disrupt the global nuclear order.

Article History

Received 22 September 2026

Accepted 14 October 2026

Published 20 October 2026

Corresponding Author*

 OPEN ACCESS

Keywords

Artificial Intelligence,
Nuclear Security,
Strategic Stability,
Nuclear Deterrence,
Cyber-Nuclear Risks, AI
Governance

1. Introduction

One of the most significant issues to international stability has been nuclear security. Previously focused on physical safety, state control, and international security, the nuclear security paradigm has been radically transformed by the advanced development of digital technologies, specifically, artificial intelligence (AI) [1]. Although the old method of deterring theft, sabotage, and unauthorised access was a priority, the modern meaning of nuclear security now has to deal with the cyber vulnerability and algorithmic decision-making, as well as the growing autonomy of a system.

In the field of security and governance, AI has become a revolutionary opportunity since it can be used to perform high-level data analysis, identify patterns, and provide automated

¹ Assistant Professor, Department of Defence & Strategic Studies, Quaid-i-Azam University Islamabad, Pakistan. msadiq@qau.edu.pk.

ORCID ID: <https://orcid.org/my-orcid?orcid=0000-0002-8699-0521>

decision-support. In nuclear systems, AI is used in reactor safety diagnostics, predictive maintenance, radiation monitoring and supply-chain optimisation and command-and-control. These capabilities promise efficiency, less human error and situational awareness [2]. Meanwhile, the increased dependence on interconnected data networks and opaque algorithms increases the cyber-attack surface and reduces the ability to exercise direct human control, which can be highly ethically, operationally, and strategically problematic, especially in nuclear command, control, and communications systems, where mistakes may be an irreversible factor.

The case of AI integration into nuclear security has also been seen as a manifestation of the wider geopolitics. The escalation in the strategic competition has pushed both the major and regional nuclear powers to seek AI-enabled military and deterrence competencies [3]. The higher the relationship of deterrence becomes, the more unpredictable and quicker it expands and the higher the possibility of miscalculation is. Meanwhile, the civilian nuclear sector is being digitalised, with AI being used to perform safety analysis, fault detection and emergency response. Nevertheless, the civilian safety technologies can be re-used to military/destabilising applications, which strengthens the dual-use dilemma of core AI-nuclear integration.

In spite of these changes, international regimes have failed to match. Current nuclear regulations and protection measures were designed in a pre-digital age and do not provide much advice regarding algorithmic autonomy, cyber-nuclear convergence, or AI-based decision-making. The lack of binding regulations and mutual principles exposes AI to most likely individualistic implementation, which has the opposite effect on stability, instead of strengthening it. This paper will discuss the role that AI is playing in transforming nuclear security, the role AI has on the concepts of deterrence and non-proliferation, and the governance measures that need to be implemented to address the emergent risks. The argument is that AI has become the biggest opportunity and the most severe threat to nuclear security since the introduction of nuclear capabilities. The dilemma facing policymakers is not about whether to use AI, but about the manner in which it should be introduced responsibly by ensuring that it is overseen by humans, subject to algorithmic transparency, and through enduring international collaboration.

2. Literature Review

2.1 The Evolution of Nuclear Security

Nuclear security has greatly changed since the atomic age due to the influence of technology, politics, and changing threat environments. It was originally concerned with preventing fissile and nuclear materials and facilities from being stolen, sabotaged, or otherwise compromised, and the role of this was mostly taken up at the national level during the period of the Cold War. Protection measures were focused on physical protection, limited access, and government control [4].

With the passage of time, the concept of nuclear security was institutionalised in the form of global structures that focused more on the prevention of malicious activities, detection, and response to them. Although these foundations are still relevant, the range of nuclear security has been enlarged to cover cyber resilience, information integrity, insider threats, and organisational governance as the digital and human vulnerabilities have continued to grow in importance. With the growth of the digitisation of nuclear sites, the application of automated systems and networked technologies brought new efficiency advantages but also new cyber threats [5]. This change proved the fact that any vulnerability in the software might have physical implications, which led to a more comprehensive security strategy that currently involves the management of data and the safeguarding of digital systems. The concept of

nuclear security today is perceived as a multidimensional system that can protect physical infrastructure and digital ecosystems that rely on it, which forms the preconditions of the introduction of advanced technologies like artificial intelligence and reveals the loopholes in the current governance frameworks.

2.2 Artificial Intelligence in Defence and Security Studies

Artificial intelligence has elicited a fierce controversy in the defence and security sector on whether it has strategic implications for military stability and deterrence. Its proponents think AI can enhance situational awareness, threat perception and decision-making in complex environments and reduce uncertainty, human fallibility and the risk of unintended escalation [6]. The decision-support systems based on AI can be used to strengthen deterrence through less misinterpretation and miscalculation of signals by sorting out false signals and recognizing patterns that cannot be perceived by humans. Nevertheless, opponents are alert to mention that AI is capable of disrupting the deterrence mechanism as it can make decisions more covertly and quickly. The short response times could also compel leaders to resort to opaque output of algorithms to process emergencies, which increases the probability of automation bias and lack of control. The sharing, responsibility, and delegation of existential decisions to machines are among the ethical and legal concerns that occur in nuclear because of this type of overdependence that creates the scepticism of blindly using AI in strategic decision-making [7].

2.3 AI in the Nuclear Domain: Emerging Applications

The convergence of artificial intelligence and nuclear security is becoming an increasingly popular topic; however, it is not thoroughly developed. AI is already supporting civilian nuclear activities and international verification and assists predictive maintenance, radiation measurements, and anomalies identification relying on the real-time analysis of sensor data [8]. The applications improve the security by detecting possible failures at an earlier stage before they develop. Non-proliferation efforts are also reinforced by AI since seismic data is examined to differentiate between nuclear explosions and natural occurrences and satellite images are processed to reveal undeclared actions, allowing quick and more precise verification. In spite of these benefits, the application of AI in the nuclear command, control, and communication systems is a very controversial one. Although AI is capable of enhancing data treatment and early prediction, it cannot eliminate human judgment, situational analysis, and moderation [9]. During the occurrence of crises, automation bias can make decision-makers apply algorithmic evaluations more than they should evaluate their risk of being misclassified, escalating, or responding unintentionally. Therefore, the question of integrating AI into NC3 systems is one of the most disputable ones in the modern nuclear security discourse [10].

Governance and Ethical Dimensions

Even though technical research on AI applications is growing, governance and ethical analysis have not kept up with it. Preexisting international nuclear systems were built amid algorithmic decision-making, autonomous systems, and cyber–nuclear convergence, and the AI-enabled nuclear systems remain in a regulatory void with undefined roles and accountability. The multilateral deliberations on autonomous weapons have covered the question of human control in warfare but have not gone as far as covering nuclear-specific cases and thus there have been no special international controls [11]. Ethical scholarship thus focuses on openness, responsibility, and effective human control, especially in the human-in-the-loop principle, where human control over key decisions is preserved. Even though the civilian models of AI governance emphasise the concepts of explainability, safety, and accountability, the use of

these models in nuclear settings is inconsistent, and people should be concerned that efficiency and technological superiority may be at odds with caution and responsibility [12].

2.5 Gaps and Theoretical Contributions

Although the importance of AI in security research is increasing, there are still a number of gaps in the literature about AI and nuclear security. Both of these are related to the first fact that most of the existing research is of a technical orientation, where the investigators focus on cybersecurity or automation but do not consider the overall strategic or governance consequences. The combination of technical vulnerabilities, the stability of deterrence, and normative regulation is depicted in few studies, and on a unitary analytical level [13]. Second, empirical studies are constrained by the fact that nuclear actions have been shrouded in secrecy and this makes them extensively dependent on hypothetical enterprise or biased information.

This limitation makes it difficult to measure actual risks and institutional practice. Third, ethical and legal aspects of algorithmic responsibility are not well developed, especially the command responsibility and liability aspects in AI-assisted nuclear decision-making. To fill these gaps, it is necessary to have an interdisciplinary approach that would bridge the technical expertise, strategic theory and normative analysis. The application of AI in nuclear security will require both technological evaluation and attention to the governance regimes, ethical limitations and the concepts of international order [14].

3. Artificial Intelligence Applications in Nuclear Security

The creation of artificial intelligence to control nuclear infrastructures is one of the major changes in which states, regulatory agencies, and operators think about the values of nuclear safety, security, and verification. AI is becoming a cross-cutting enabler in the nuclear industry, with the ability to improve analytical capability, predictivity and efficiency. Simultaneously, its dual-use character implies that the same technologies that can reinforce the nuclear security can also create new weaknesses [15]. This part reviews the key spheres where AI is already implemented in the nuclear field, outlines its benefits and the risks involved.

3.1 Facility Operations and Predictive Safety

The second area of AI use that is the oldest and most established in the nuclear industry is predictive maintenance and operational optimisation. Contemporary nuclear plants produce huge amounts of real-time data provided by sensors controlling temperature, pressure, vibration, and radiation levels. Machine-learning models process these continuous information streams to identify irregularities and anticipate component deterioration prior to faults growing into safety occurrences [16]. AI makes the work more reliable and more resilient to unplanned shutdowns or accidents, as it allows for preventing them in advance.

AI is also used to assist radiation mapping and emergency response planning. Historical incident-based models may simulate the evolution of accidents, predict the difficulties of containment, and assist in making decisions in the context of an emergency. By making nuclear safety more robust by minimizing the chances of technical faults being used towards sabotage or escalating into a more widespread crisis, these predictive capabilities enhance the security of nuclear safety [17]. Nonetheless, there are cyber-physical dependencies associated with greater dependence on digitally integrated systems as well, so the malfunction of AI or a cyber-attack on AI systems could have a direct physical impact on nuclear plants.

3.2 Physical Protection and Surveillance Systems

The use of AI-based surveillance technologies has changed the physical protection practice in nuclear sites. The video analytics, which are intelligent, are taking over the old closed-circuit television systems, which used human monitoring to a large extent. Machine-vision algorithms have the benefit of identifying abnormal behaviour, identifying unauthorised access, and identifying suspicious movement patterns in restricted regions faster and more consistently than their human counterparts [18].

There has also been the evolution of biometric access control, where AI has been used. Biometrics such as facial recognition and any other personnel-related databases, enhance the process of identity verification and decrease the threat of misusing credentials. Simultaneously, these systems increase issues regarding privacy, data protection, and system integrity. Attacks like spoofing or image manipulation can be described as adversarial, which poses difficulties for biometric reliability, has difficulties in and it is necessary to have strong encryption, redundancy and oversight to make sure that physical protection systems are secure and reliable.

3.3 Insider-Threat Detection and Behavioural Analytics

Insider threats are one of the most enduring risks to nuclear security since those that have the authorised access will knowingly or unknowingly breach sensitive systems or materials. The behavioural analytics offered by AI are new resources to handle this challenge as it will enable continuous monitoring of the patterns of access, the activity of the system, and anomalies in behaviour [19]. The AI systems are capable of detecting anomalies in set standards and notifying the business owner about the possibility of insider threats before they lead to security breaches. This change is a general trend towards reactive to proactive nuclear security where threats are foreseen as opposed to responding to them once they happen. Nonetheless, ethical and organisational issues are concerned with the application of AI in behavioural monitoring. Over-surveillance, false alarms and erosion of trust in the workplace can demoralize and invalidate legitimacy unless appropriately handled [20]. The best insider-threat programmes should thus be balanced in predictive capability, human monitoring, and privacy and proportionality.

3.4 AI in Nuclear Command, Control, and Communication (NC3)

The most debatable and sensitive realm of AI implementation is the implementation of AI in nuclear command, control, and communication systems. NC3 systems are the basis of strategic deterrence by making sure that there is reliable threat detection, safe communications and that a response is controlled in any circumstance. AI can be used to improve these systems, making them faster in data processing, enhance sensor fusion, and eliminate false alarms with early-warning data. Proponents state that NC3 with the help of AI might enhance deterrence, as it will be more reliable and less prone to human error [21]. Nevertheless, the opponents caution that automation can shorten decision-making times and make people more dependent on unclear algorithms in emergencies. It is especially dangerous in nuclear applications, where automation bias might arise, where human operators give too much importance to machine-generated advice. Moreover, AI software weaknesses may be used to issue inaccurate warnings about launches or interfere with communications, which increases the possibility of false or unintended attacks [22].

3.5 AI for Non-Proliferation, Verification, and Nuclear Forensics

Non-proliferation monitoring and verification are two aspects that have been viewed as increasingly essential with the use of AI. High-tech algorithms are used to examine satellite images, environmental samples, and sensor measurements that help detect signs of undeclared

nuclear operations. The slight variations in construction patterns, thermal patterns, or emissions can be identified much faster and more precisely than with the help of manual analysis only. Nuclear test monitoring through AI will also assist in identifying natural seismic activities and those that are in line with nuclear tests. Nuclear forensics AI supplements isotope analysis and environmental sampling to achieve quicker attribution and increase confidence in the accuracy of the verification results [23]. Moreover, AI can also be used in export control by testing trading patterns and supply chains to determine the possible illegal transfer of dual-use goods.

Along with these benefits, AI-enhanced verification has political and legal issues. Advanced surveillance can be viewed as intrusive or biased by states, especially when they are in the hands of more technologically superior actors. Trust in the results of verification is also in danger of being undermined because of excessive reliance on proprietary or opaque algorithms, which, in turn, justifies the significance of transparency, explainability and multilateral validation.

3.6 Cross-Cutting Benefits and Constraints

In all these spheres of application, AI has evident advantages, such as rapidity, precision, predictability, or greater resilience of the system. The role of the continuous monitoring and quick detection of anomalies helps to decrease the risk of human error and enhance readiness to complex emergencies. Virtual stress tests and contingency planning are other benefits of AI-based modelling and simulation through which nuclear authorities can be better prepared to deal with a crisis. Nevertheless, a number of systemic limitations restrict these advantages. Intensive and quality data of large size and security is necessary in well-functioning AI systems, and many nuclear operators do not have a large amount, especially in the developing international states [24]. The adoption of AI in existing nuclear systems frequently involves the need to upgrade existing equipment at a significant cost, and also creates novel cybersecurity threats. Additionally, the lack of ethical and legal answers to the issue of data control, responsibility, and role of the private sector makes the implementation of AI in vulnerable nuclear contexts a complicated issue.

4. Emerging Risks in the AI–Nuclear Security Nexus

Even though the idea of artificial intelligence offers a lot of opportunities that can be used to foster the safety and efficiency of nuclear systems, it also introduces a new spectrum of risks to the stability and reliability of nuclear systems. Combination of AI, cyber technologies and nuclear plants creates openings in areas beyond the conventional challenges of cybersecurity [25]. Opaqueness of algorithms to human perception, adversarial manipulation, automatic bias, data poisoning and less humanness all lead to the heightened risk of accidents, sabotage or unintended escalation during crisis times.

4.1 Cyber Exploitation of AI Systems

Nuclear facilities, command networks are becoming highly dependent on the use of AI-driven decision-support systems, which require extremely large datasets to operate efficiently and must be digitally connected. This improves situational awareness, on the other hand, it increases the surface of cyber-attack. The system can be manipulated by methods such as data poisoning, adversarial inputs, and manipulation of the AI output to make systems misclassify threats or overlook actual anomalies [26]. There are historical examples of cyber-attacks resulting in physical effects, indicating the risk of cyber-attacks in nuclear settings with AI functionality.

These weaknesses are especially acute in the nuclear command, control, and communications systems. Early-warning architecture AI-assisted systems rely on massive sensor data to differentiate actual attacks and false signals. In case the enemies tamper or spoof this data, automated systems can raise false alarms or overrule genuine warnings, escalating the risks of high levels [27]. The increasing interconnection between cyber and nuclear space thus forms destabilising feedback mechanisms where comparatively small intrusions can disproportionately affect the strategy.

4.2 Algorithmic Opacity and the “Black Box” Problem

Algorithms are considered to be the defining problem of AI systems, which is “black box problem” or algorithmic complexity. Most sophisticated models can produce results without offering any clear and understandable explanations as to their choices. This non-transparency jeopardizes transparency and trust in nuclear settings [28]. The policymakers and the operators might have no technical ability to examine or question the AI-supported assessments, especially in cases of emergencies where time is of the essence. The issue of opacity is not only technical, but it is also a matter of governance. Once it becomes ambiguous about who is to take the responsibility in making decisions, it is challenging to hold anyone legally or politically accountable for errors. Confidentiality of proprietary or classified AI models can also undermine trust among the states and further breed suspicion and even worsen the arms race instead of decreasing instability.

4.3 Automation Bias and Human Disengagement

Automation bias occurs when people who operate algorithms trust in the results of the algorithms too much, regardless of the fact that there is conflicting evidence. This bias is deadly in the cases of nuclear operations where making a decision is life and death and the consequences are devastating. The leaders might yield to AI-delivered suggestions during a crisis because they believe that being a computer is being a correct. The significance of human judgement in accidental escalation prevention is demonstrated in the course of history. AI aggravates the risk of automation by heightening the complexity of the system and the sense of control of machine results [29]. Though retaining a human in the loop organization is highly recommended, speed and technological superiority pressure by political and military leaders may encourage more autonomy. The trend threatens to undermine human judgement and undercut crisis management systems.

4.4 Adversarial Manipulation and Deepfake Deception

Adversarial manipulation of AI systems is also susceptible to a carefully designed input that is meant to deceive algorithms. Minor information manipulation of satellite images, sensor data, or radar signals can result in AI models that consider civilian behaviour as an act of aggression. The cost of such attacks is comparatively cheap, can be scaled and hard to detect. Deepfake technologies also increase the risks [30]. False images, fake telemetry or fake evidence could be used to simulate the violation of treaties, and this can upset the verification processes, or possibly lead to military responses. In addition to technical deception, AI-created misinformation and propaganda can influence the perception in times of crisis and degrade trust in international institutions as well as blur the line between nuclear security and information warfare.

4.5 Systemic Complexity and Cascading Failures

The AI integration turns nuclear systems into closely interconnected, intricate networks where the failures may spread quickly within subsystems. When one of the AI-based functions malfunctions, the others can follow suit and will lead to unpredictable results. To illustrate, poor AI-based maintenance evaluations may reduce the quality of safety diagnostics and lead to the emergence of improper reaction of operations [31]. The real-time synchronisation between digital systems enhances this risk. Since the AI models are dynamic learners, they can change their behaviour in a manner that even designers cannot predict completely. This is further exposed to globalised supply chains, whose AI elements created by third parties can include additional dependencies or vulnerabilities that can be used to spy or sabotage.

4.6 Data Security and the Problem of Trust

The quality, integrity, and security of the training data determine the performance of AI. Data sharing and testing in nuclear is frequently hampered by restrictions on national security, which results in many systems using few or synthetic data. This may make it more vulnerable and decrease the strength. Remote data storage and cloud computing make the protection of data even more challenging, exposing it to cyber intrusions even more [32]. On the global scale, distrust does not allow states to exchange information or algorithms to verify, which affects the lack of transparency and deteriorates collaboration in non-proliferation on an international scale.

4.7 Ethical and Legal Risks

It is accompanied by serious ethical and legal issues that are posed by the introduction of AI into the nuclear systems. The current legal systems that regulate the application of force presuppose the human judgment and responsibility when making decisions. Nuclear systems which are fully automated or highly automated pose a challenge to the principles of accountability, proportionality and distinction. Although various ethical frameworks are based on transparency, accountability, and human control, they are mostly voluntary and oriented towards civilian uses [33]. Without regulatory treaties at the international level, the adoption of AI into nuclear command systems is likely to blow out established moral and legal limits on the use of nuclear power and to confuse accountability when algorithmic decision-making leads to disastrous consequences.

5. Strategic Stability, Deterrence, and Non-Proliferation in the Age of Artificial Intelligence

5.1 Deterrence Stability in the Algorithmic Era

The ability of AI to process information at machine speed can be an incentive to use the launch-on-warning posture by the states, which will limit the time needed to verify the information. Although in theory rapid response may increase survivability, it also raises the chance of an accidental launch into space in case the automated systems give false alarms. With the continuous shortening of decision schedules, there is an increased use of algorithmic judgments that have been questioned due to the possibility of misinterpretation and unintentional growth.

It is also not clear where offensive and defensive capabilities are in AI [34]. The technologies that are designed to track missiles, which can be used in surveillance or counter the cyber-attacks can be used to attack or break in with the result being uncertainty over the intent. This dual-purpose nature complicates the process of sending signals in the case of a crisis and since AI-enhanced surveillance can be viewed by states as leading to pre-emptive action, it builds

to the escalation of arms competition. In addition to this, it can also have AI-assisted precision delivery that can be applied in decapitation strategy as the warning time can be reduced, particularly when combined with hypersonic delivery systems [35].

Such kind of development affects a loss of trust in the guarantee of the second strike and increases the incentives of the devolution of authority to the automated systems, which explicitly doubts the principles of human control. In the meantime, the deterrence can also be assisted with the help of AI, but only in a limited amount. False alarms can be reduced by improving sensor fusion and anomaly detection to gain more situational awareness, in order to make better decisions. Lastly, the factor of the AI in deterrence stability is decided by the integration of AI into the command systems and the human judgement in the nuclear decision-making.

5.2 The Emergence of Algorithmic Arms Races

The quest to gain military superiority through AI has led to what has commonly been referred to as an algorithmic arms race. Competition has come out of the measurement of traditional military power, but has been incorporated to include control over data, computing resources, semiconductors, and autonomous decision-support systems. The world powers are growing to position AI dominance as a strategic need, and have integrated it into national security and defence plans. These forces threaten to recreate the destabilising tendencies of previous arms races. States not only evaluate relative advantage based on model sophistication, the speed of calculation and data access, rather than based on the number of warheads. Fear of technological inferiority causes the transparency to decrease and secrecy to rise and erode mechanisms of confidence-building that decreased nuclear tensions historically [36]. The increasing involvement of private technology companies also makes arms control more difficult because most of the capabilities in the relevant AI are designed and disseminated beyond the standard state-controlled defence industries.

Competition threatens to detach deterrence with human psychology and restraint, making it a machine optimisation competition. It is not the intentional aggression but the unwanted amplification, which is predetermined by the interaction of the algorithms that are running at the speed of a machine. To overcome this obstacle, the arms-control frameworks must go beyond the scope of physical weapons to encompass the capabilities of digital nature as well as the explicit guarantees of human control.

5.3 Non-Proliferation and Technology Diffusion

AI has both opportunities and challenges to nuclear non-proliferation. On the one hand, AI minimizes technical barriers since it allows complex analysis by software instead of equipment. Algorithms digitally transferable are able to analyse isotopic data, model enrichment processes, or optimise reactor parts without necessarily considering the physical materials as a source of conventional types of export controls. There is also a risk of using commercially available AI tools in an illegal way to conduct research, such as simulate a reactor system or enrichment cascade [37]. The state and non-state actors might use AI-powered cyber activities to steal delicate nuclear data, or alter documents to avoid discovery. These threats make the current non-proliferation regimes, which are material control-based, to be more complex than digital diffusion-based.

AI, on the other hand, has great power in terms of verification. To a greater extent, increased imagery intelligence, anomaly detection, and environmental sensing will increase the capability of identifying undeclared facilities and activities that are suspicious. AI-assisted

seismic analysis can better detect nuclear tests, whereas better nuclear forensics can attribute more quickly. The main issue is the problem in balancing technological efficiency and political trust. Lacking transparency and capacity-building the AI-based protection will be seen as invasive or discriminative, which contributes to the technological inequity issue within the non-proliferation regime [38].

5.4 Crisis Management and Escalation Dynamics

The core feature of AI is the change in the crisis management in speeding up the information gathering, analysis, and distribution. The old system of deterrence was based on the time buffers that would enable diplomatic signalling and de-escalation. These buffers are compressed by the AI-enabled systems, putting more pressure on the decision-makers to make decisions when faced with uncertainty. This increase in speed increases the danger of pre-emptive attack, especially in areas with already low warning times. Surveillance and autonomous systems are becoming more common, thereby reducing the time the operator has to make crisis decisions, which makes it more likely that a mistake will be made [39]. Predictability and rational behaviour in times of stress is required to achieve strategic stability, but AI brings about non-linear and adaptive behaviours which overall can be hard to predict. The interaction of autonomous systems may create escalation loops of feedback that are impossible to control, and unprecedented challenges to the crisis management.

5.5 Legal and Normative Erosion

In addition to operational risks, AI is a threat to the normative sources of nuclear restraint. International legal systems are not used to algorithmic autonomy and AI-assisted decision-making or cyber–nuclear integration. Consequently, a question of accountability is becoming unclear as the automated systems affect the choices of the nuclear systems. In the event that AI systems are allowed to influence or perform nuclear activities, there are basic legal issues of accountability, proportionality, and coloured adherence to humanitarian guidelines [40]. Decentralizing power to black box systems may lead to the development of a morally vacuous situation where the responsibility is devolved among the programmers, operators, and political leadership. Moreover, the confidentiality of algorithms seems to be irreconcilable with the transparency of arms control and verification, which annihilates confidence and co-operation between states.

5.6 Potential Stabilizing Effects

These risks notwithstanding, AI can be used to provide strategic stability when it is managed in a responsible manner. Improved early-warning system which is a multisource of data can help to minimize false alarms and increase the reliability of threat evaluation. The use of AI-driven simulations and digital models enables policymakers to test the scenarios of the escalation and experiment with crisis-management plans prior to the emergence of the real crises [41]. The use of AI to conduct analysis also has the advantage of providing faster and more accurate verification by regimes. Artificial intelligence can also help in terms of transparency, by means of open-source tools and increased involvement in monitoring activities, in addition to the state-specific verification processes. Nevertheless, the stabilising effects are subject to high ethics, institutional controls, and international collaboration. Without this kind of protection, the very technologies that can offer stability can increase the threats of inaccuracy and escalation.

6. Comparative Case Studies: Global and Regional Approaches to AI–Nuclear Integration

Artificial intelligence has been introduced in nuclear security and strategic infrastructures in an uneven manner at the state level. Although AI may offer the benefits of deterrence and modernisation and safety to all nuclear-armed countries, they vary significantly in their approaches toward their adoption and regulation [42]. These disparities indicate the differences in technological capability, strategic culture, political priorities and availability of resources. The analysis of great players depicts that AI -nuclear integration is more about great powers pursuing the more overarching geopolitical goals than about an international strategy.

6.1 The United States: Responsible Innovation and Ethical Restraint

The United States has been a worldwide leader in the advancement of AI and has taken a safer stance in the implementation of AI in its nuclear business. The main applications of AI are to modernise the infrastructure, improve predictive maintenance, better cyber defence, and provide greater material accountability. The AI in national laboratories is also used to monitor the behaviour of the reactors, identify anomalies, anticipate equipment failures and the governance systems aim at safety, controls and ethical standards [43]. Artificial intelligence-driven data fusion and analytics may support intelligence, monitoring, and early-warning systems that are applicable in nuclear command in the military.

Nevertheless, the U.S. policy directly limits complete autonomy in decision-making in nuclear operations, keeping the human control over critical decisions. This is aimed at balancing technological innovation and containment of escalation, yet critics state that the U.S. dominance in the field of algorithmic deterrence can encourage competitors to develop the capability faster which adds to the competition on the global stage.

6.2 China: Intelligentsia and Civil-Military Fusion

In its AI-nuclear integration strategy, China is part of an overall strategy of the so-called intelligentsia warfare, making AI the core of future military strength. Beijing emphasizes the adoption of AI, big data, and autonomous systems in the command, control, and surveillance networks. In the nuclear domain, AI is the aid to early warning systems, missile control, and decision-support vehicles, as well as a large-scale integration of civil and military research. Another area where China has invested heavily is the AI-powered civilian nuclear technologies, such as so-called smart reactors with predictive diagnostics and robotic maintenance [44]. Although the given applications enhance efficiency and safety, they also create the dual-use synergies that dilute the border between civilian and military AI. The strong interaction between the state and the private technology companies will make the process of innovation faster, but at the cost of international efficiency concerns of transparency, information privacy, and strategic direction. China emphasizes more on the development of capability compared to the formal ethical restriction that is emphasized in the Western models.

6.3 Russia: Automation and Strategic Ambiguity

The AI-nuclear approach by Russia is an amalgamation of ambitions and ambiguity in institutions. Moscow considers AI as the tool to balance out economic limitations and maintain the strategy level with more technologically sophisticated competitors. The defence-industrial policies it has focus on automation, simulation and command system resilience. Traditionally, Russia has expressed readiness to make vital functions be automated, which speaks of a long-term focus on the establishment of retaliatory capabilities in extreme circumstances [45].

Current efforts are directed at the implementation of AI in detecting threats, electronic warfare, cyber operations, and the reliability of command systems. Nonetheless, the lack of transparency and low involvement in the debate on the international AI governance makes it hard to determine the scope of protection or ethical limitations to the integration of AI in Russia, which adds to the lack of strategic clarity.

7. Policy Frameworks and Governance Gaps

Nuclear governance globally has failed to abreast with the development of artificial intelligence. The current structures, which were established on the basis of treaties of the mid-twentieth century, were not created to take into consideration algorithmic autonomy, data-based surveillance, or cyber-nuclear convergence. With the incorporation of AI into the nuclear systems, these constraints have resulted in the creation of fragmented oversight, normative uncertainty, and institutional stasis [46]. This part explores significant structures of governance, and discerns regulatory loopholes which are increasingly becoming a menace to world peace.

7.1 The International Atomic Energy Agency (IAEA) and the Limits of Mandate

International Atomic Energy Agency is the main institution of world nuclear safety, security and protection. The Agency provides guidance and nuclear security recommendations, which offer best practices in the area of physical protection, material accounting, and information security. It has also started working on the opportunity of using AI to assist in managing a reactor and conducting verification processes. Nonetheless, IAEA mandate is mostly advisory and it was in place even before the digitalization of nuclear systems. Although its cybersecurity recommendation creates awareness, the Agency does not have legal power to control AI algorithms, data governance, or autonomous decision-support systems [47]. Its advice is not compulsory, so its implementation is not even across states, especially the states with weak technical capability. The lack of resources also means that the Agency will be unable to meet the fast-changing AI technologies and the emerging cyber threats, which could limit its performance in regulating the use of AI-enabled nuclear systems.

7.2 The Nuclear Non-Proliferation Treaty (NPT) and Technological Obsolescence

The Nuclear Non-Proliferation Treaty has been the keystone of nuclear order in the world which is based on the principles of non-proliferation, disarmament and peaceful use. Nonetheless, it has its provisions based on a material-based concept of proliferation and fails to consider software-based and intangible technologies. All three pillars of the treaty are challenged by AI since it facilitates diffusion of the capabilities of the duo of uses digitally, promotes the modernisation of the nuclear arsenals with the assistance of automation, and increases the technological disparity between developed and developing countries [48]. Despite the recent review mechanisms that have recognised new technologies, the new norms have remained elusive against concerns of military transparency and sovereignty. Consequently, the governance functionality of the treaty is still linked to an analog mode of operation that is ill suited to handle algorithmic risks.

7.3 The Convention on the Physical Protection of Nuclear Material (CPPNM) and Cyber Deficiencies

The revised Convention on the Physical Protection of Nuclear Material compels states to ensure that the protection of nuclear materials against theft and sabotage, but it is more physical. Although the amendment broadened the meaning of sabotage, it failed to cover the

cyber threats or AI-assisted control systems. With AI taking over more and more of the control power in access control, surveillance, and anomaly detection at nuclear plants, digital interference can compromise the physical security completely [49]. The lack of binding demands in software assurance, algorithmic auditing and cyber -sampling assessment of risk has led to inconsistent national standards. The CPPNM is at risk of losing its relevance unless it is revised to incorporate the digital acts of sabotage, since nowadays, attackers can do so via code and not necessarily by physical intrusion.

7.4 The Role of the United Nations and Multilateral Dialogues

On the multilateral level, AI governance debates are mainly held at United Nations forums dealing with autonomous weapons and the security of cyber space. The development has been hampered by political differences whereby some states prefer binding regulations whilst others use voluntary controls based on human control. These debates have mostly omitted nuclear weapons, even though this is the most dangerous use of AI. Other UN efforts in the area of digital ethics do not have enforcement mechanism and poor coordination of the processes of technology regulation and disarmament distributed control. Consequently, AI-powered nuclear systems are in a regulatory Gray area, which is between institutional requirements and policy areas [50].

7.6 National Governance and Private-Sector Influence

Different nuclear-related AI governance approaches are practiced in different countries. States tend to focus on the ethical approaches that encourage the transparency and human control, whereas others make technological sovereignty and control of information a priority. The developing countries do not have a detailed AI policy and are dependent on external support on a case-by-case basis. The prevailing role of the business world in the development of AI also makes governance harder. Technology companies are also providing algorithms in the energy and defence markets, but they are informally regulated by the current nuclear security and arms-control systems [51]. Algorithms are also not clear as to who is responsible in case of failure in such sensitive nuclear environments especially as open-source models continue to spread and can be abused by non-state actors. Innovation and non-proliferation should be thus balanced in terms of closer coordination between the public and the private, and better-defined accountability structures.

7.7 Key Governance Gaps

Collectively, the available frameworks demonstrate governance loopholes. Nuclear decision-making on autonomous or AI-driven basis is not regulated by binding international rules, there is no unified institutional oversight, and no universal criteria regarding the AI systems verification in nuclear installations. Capacity asymmetries place lots of states in a situation where they are unable to effectively enforce AI safeguards when the developers of AI work outside the old regulatory framework. All of these weaknesses add up to a lack of faith in the nuclear order and the normalisation of automation in strategic decisions, which erodes years of deterrence stability.

7.8 Towards Integrated Governance

As a solution to these challenges, integrated governance involving institutional reform and normative innovation is needed. The current treaties should be modified to incorporate

algorithmic accountability and digital risk management, whereas the international institutions are to enhance technical coordination and supervision. The measures of confidence-building concerning AI-related incidents, more transparent responsibility frameworks between the public and the private, and long-term capacity-building of developing countries are necessary [52]. The international community will be left with no proper protection against an algorithmic nuclear crisis without political will and multilateral cooperation.

8. Recommendations and Future Outlook

8.1. Embedding Human Control and Algorithmic Accountability

The most urgent need of AI integration is keeping a significant human control over nuclear decision-making. It is important that the human authority is kept to the core of command, control, and communication systems, especially in high-stakes situations. Human in the loop principle must be entrenched in the laws of the country and should be enshrined in international standards to make sure that it is automated systems that assist not substitute human judgement. This principle can only be operationalised by having clear accountability mechanisms such as audit trails, explainable system outputs and unceasing monitoring of algorithmic performance. Artificial intelligence deployed in the nuclear environment must be of a minimum transparency and traceability to ensure that the actions taken by the system can always be traced back to the human participants [53]. Explainable AI should be invested in to minimize the use of non-transparent systems and provide operators the ability to comprehend, challenge, and go against the advice of an automated system when needed.

8.2. Establishing International AI-Nuclear Governance Mechanisms

Since AI-based nuclear risks do not have a sovereign state, it is not possible to manage them unilaterally. There must be a special international system of AI in strategic systems in order to bring a system of harmony in the standards, transparency, and cooperation. A framework like this can be organized by using the exiting international institutions, and it should be made to complement the already existing nuclear treaties and not to substitute them. In the short run, although binding commitments can be politically challenging, small steps, including voluntary codes of conduct, common technical standards and reporting systems can foster trust and provide the basis of subsequent formalisation [54].

8.3. Integrating AI into Existing Arms Control and Verification Regimes

The current arms control and non-proliferation treaties need to be revised to include AI-controlled technologies. Additional procedures may help make clear the way AI-based verification tools are created, assessed, and audited to make them reliable and just. Nuclear plants are supposed to be advised to undertake routine cybersecurity and AI-risk testing considering the digital interference as a type of sabotage. An arms control negotiation between nuclear-armed countries must directly deal with the AI-related risks, such as data exchange, joint simulation, and confidence-building agreements that are aimed at mitigating the risk of misinterpretation or escalation by automated systems [55].

8.4. Enhancing Cyber Resilience and Technical Safeguards

AI implementation in nuclear settings requires a strong cyber resiliency. Multitiered defence techniques must be implemented and these include redundancy, encryption, anomaly detection and isolation of the network. The technical protection should be included in the

procurement and licensing procedure to ensure that the deployment of AI systems in the sensitive nuclear purposes is limited to certified systems. Stress testing and adversarial simulations should be conducted regularly to determine the vulnerabilities and fail-safe designs and human override features should become a standard of AI-enabled nuclear systems.

8.5. Promoting Ethical and Normative Restraint

In order to ensure the validity of nuclear decisions, there must be ethical governance. The ethics frameworks in AI developed by states and organisations must focus on accountability, fairness, transparency, and human supervision and be modified to address the distinct dangers of strategic infrastructure. Ethical restraint is maintained through education and training [56]. The result of the interdisciplinary programs, which integrate nuclear engineering with computer science, ethics and international law can be used to develop a new generation of professionals who will be able to balance innovativeness and responsibility.

8.6. Engaging the Private Sector and Civil Society

With the innovation of AI driven mainly by the efforts of the private companies, appropriate nuclear governance should incorporate systematic interaction with industry. The governments are supposed to institutionalise the partnership between the governments and the businesses, which lays down the safety responsibilities, export restrictions, and the transition disclosures to those AI developers in nuclear-proximate fields. The voluntary responsibility commitments and cooperation in terms of the open verification tools can enhance accountability and decrease reliance on the proprietary systems [57]. The civil society, academic institutions, and think tanks are also crucial since they keep track of the progress in technology, facilitate discussions, and make contributions to the informal confidence-building processes.

8.7. Fostering Inclusivity and Global Equity

AI poses a threat of increasing the inequality that is already present in the world nuclear order because few states have developed technological infrastructure. It is a normative and strategic imperative therefore that safe AI applications are made available to everyone. The developing states will be able to implement AI in a responsible way through the aid of the capacity-building initiatives, such as training programs, joint research, and technology-transfer mechanisms. Inclusive governance will minimize motivation towards secret procurement and confidence in multilateral institutions [58].

8.8. Anticipating Future Threats

AI's interaction with other emerging technologies will further complicate nuclear security. Autonomous weapons, hypersonic delivery systems, and synthetic media may compress decision timelines and undermine trust in information environments. Preparing for these risks requires anticipatory governance rather than reactive regulation.

States should establish permanent foresight units to integrate scientific expertise into strategic planning, while international institutions should provide regular forums for assessing emerging technologies and updating risk-reduction measures [59]. Experience from the nuclear age demonstrates that institutions created after crises rarely prevent future ones; proactive governance is therefore essential.

9. Conclusion

Artificial intelligence is becoming one of the characteristics of the nuclear order of the twenty-first century. It provides strong means of improving safety, monitoring, verification and decision support, but also presents new threats in terms of cyber vulnerabilities and automation mistakes and normative erosion. This two-sided nature of AI as an enabler and a threat is the very center of modern nuclear security issues. Technological change has greatly exceeded the processes of nuclear governance. The fundamental structures of the nuclear order were designed to deal with material control and physical protection and they were unprepared to deal with algorithmic obscurity, manipulations with data, and autonomous decision-making. Without any renewed rules and common norms, states become more and more competitive and fragmented in their pursuit of AI-enabled capabilities, which increases the risk of error of calculation and unintentional growth. Devoid of legal controls over algorithmic control, deterrence stability is weaker, not stronger [60].

AI can reduce the time taken to decide and create a Gray area in terms of offensive and defensive operations, which is a strategic move. Although these technologies have the potential to enhance verification and crisis management, they also destroy predictability and cause overdependence on automated systems whose behaviour can be misinterpreted or manipulated. The very algorithms that will help identify the breaches of treaties or improve the early warning may be used, if abused, as the fuel to the escalation instead of the tools of restraint. Whether AI should be incorporated into the nuclear systems or not is not the policy challenge, but the means of its regulation. Human control should be very meaningful, as the basis of nuclear decisions. The global collaboration is necessary to align the technical standards, exchange risk-cutting experience and assist states, which lack the capacity. The participation of the private industry and civil society is equally important, as their experience and responsibility will determine the confidence in AI-driven security.

Finally, ensuring nuclear security in AI will less rely on technological advanced than the political decision making. The same eternal truth of the nuclear era still remains: technology changes quicker than conventions, yet innovation is based upon human decisions: it will stabilize, or it will destroy everyone. The penetration of AI into the nuclear sector is irreversible yet the outcomes of this process are not pre-programmed. With the principles of transparency, ethical moderation and responsible governance, AI can strengthen the very norms it currently is undermining and assist in transforming the nuclear order of a precarious equilibrium of fear into a more resilient system of cooperative security.

References

- 1 Warren, A.: 'The Reemergence of Nuclear Competition and Global (In) Security': 'Global Security Reimagined: A Multi-Disciplinary Exploration' (Springer, 2025), pp. 113-133
- 2 Moran, D., Russell, J.A., Leavitt, S., Kesler, B., Blanken, G., and Dearing, M.: 'Global Trends and the Future of Warfare 2025, Strategic Insights, Volume 10, Special Issue, October 2011', 2011
- 3 Verbruggen, M., and Boulanin, V.: 'Mapping the development of autonomy in weapon systems', 2017
- 4 Herzog, S.: 'The Nuclear Fuel Cycle and the Proliferation "Danger Zone"', *Journal for Peace and Nuclear Disarmament*, 2020, 3, (1), pp. 60-86

- 5 Kazi, R.: 'The Global Centre for Nuclear Energy Partnership: An Indian Answer to the Nuclear Security Challenge': 'India in Global Nuclear Governance' (Routledge, 2019), pp. 83-106
- 6 Ali, I., and Sadiq, M.: 'The Perils of Non-State Actors in Pakistan: Assessing the Risks of Nuclear Safety and Security', *International Journal of Nuclear Security*, 2023, 8, (1), pp. 5
- 7 Cox, J., and Williams, H.: 'The unavoidable technology: How Artificial Intelligence can strengthen nuclear stability', *The Washington Quarterly*, 2021, 44, (1), pp. 69-85
- 8 Johnson, J.: 'Artificial intelligence in nuclear warfare: a perfect storm of instability?', *The Washington Quarterly*, 2020, 43, (2), pp. 197-211
- 9 Weitoish, T.: 'US Cybersecurity and Enhancing SCADA Systems and Critical Infrastructures', *Capitol Technology University*, 2025
- 10 Greiman, V.A.: 'Nuclear Cyber Attacks: A Study of Sabotage and Regulation of Critical Infrastructure', in Editor (Ed.)^(Eds.): 'Book Nuclear Cyber Attacks: A Study of Sabotage and Regulation of Critical Infrastructure' (Academic Conferences International Limited, 2023, edn.), pp. 103-XII
- 11 Bauer, S., and O'Reilly, C.: 'The Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO): current and future role in the verification regime of the nuclear-test-Ban Treaty': 'Nuclear Non-Proliferation in International Law: Volume II-Verification and Compliance' (Springer, 2015), pp. 131-150
- 12 Pokharel, P.: 'Artificial Intelligence Modernization in the US Air Force: Strategic Implications, Readiness, and National Security Impact', *Readiness, and National Security Impact* (November 17, 2025), 2025
- 13 Schmidt, E.: 'US National Security Commission on Artificial Intelligence': 'Augmented Education in the Global Age' (Routledge, 2023), pp. 234-244
- 14 Taddeo, M., McNeish, D., Blanchard, A., and Edgar, E.: 'Ethical principles for artificial intelligence in national defence': 'The 2021 Yearbook of the Digital Ethics Lab' (Springer, 2022), pp. 261-283
- 15 Shehabi, A., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., Koomey, J., Masanet, E., and Sartor, D.: '2024 united states data center energy usage report', 2024
- 16 Group, T.I.E., Branson, W., Laffont, J.-J., Solow, R., Ulph, D., von Weizsäcker, C., and Kyriakou, D.: 'Economic dimensions of prospective technological studies at the Joint Research Centre of the European Commission', *Technological Forecasting and Social Change*, 2002, 69, (9), pp. 851-859
- 17 Horowitz, M.C., Allen, G.C., Saravalle, E., Cho, A., Frederick, K., and Scharre, P.: 'Artificial intelligence and international security' (Center for a New American Security., 2022. 2022)
- 18 Chauvet-Maldonado, C.: 'The Protection of Nuclear Installations in an Armed Conflict: Nuclear Security Concerns', *Journal of Nuclear Materials Management*, 2024, 51, (3), pp. 20-30
- 19 Findlay, T.: 'Safeguarding a Nuclear Energy" Boom": the IAEA's Needs by 2050', 2025
- 20 Pollard, S.D., Aytac, J.M., Kellison, A.E., Laguna, I., Nedunuri, S., Reis, S.A., Sottile, M.J., and Thornquist, H.K.: 'The First Tri-Lab Workshop on Formal Verification: Capabilities, Challenges, Research Opportunities, and Exemplars', 2024
- 21 McInnis, J.M.: 'Russia and China look at the Future of War', *Washington, DC: Institute for the Study of War*, 2023
- 22 Andrade, P., Laadjal, K., Alcaso, A.N., and Cardoso, A.J.M.: 'A Comprehensive Review on Condition Monitoring and Fault Diagnosis in Fuel Cell Systems: Challenges and Issues', *Energies*, 2024, 17, (3), pp. 657

- 23 Heller, K.J.: 'The Concept of "the Human" in the Critique of Autonomous Weapons', *Harv. Nat'l Sec. J.*, 2023, 15, pp. 1
- 24 Lindsay, J.R.: 'Stuxnet and the limits of cyber warfare', *Security studies*, 2013, 22, (3), pp. 365-404
- 25 Clement, S.: 'NATO and Artificial Intelligence: Navigating the challenges and opportunities', *Science and Technology Committee (STC)*, 2024, 58, pp. 1-17
- 26 Canton, H.: 'Organisation for economic co-operation and development—OECD': 'The Europa Directory of international organizations 2021' (Routledge, 2021), pp. 677-687
- 27 Payne, K.: 'I, warbot: The dawn of artificially intelligent conflict' (Oxford University Press, 2021. 2021)
- 28 Perrow, C.: 'Normal accidents: Living with high risk technologies-Updated edition', 2011
- 29 Schneider, M.B.: 'How many nuclear weapons does Russia have? The size and characteristics of the Russian nuclear stockpile', *Comparative Strategy*, 2024, 43, (4), pp. 305-433
- 30 Johnson, J.: 'AI and the bomb: Nuclear strategy and risk in the digital age' (Oxford University Press, 2023. 2023)
- 31 Favaro, M., and Williams, H.: 'False sense of supremacy: Emerging technologies, the war in ukraine, and the risk of nuclear escalation', *Journal for Peace and Nuclear Disarmament*, 2023, 6, (1), pp. 28-46
- 32 Khademizadeh, S., Dakhesh, S., and Lund, B.: 'Characteristics of Global Retracted Publications in Engineering Sciences: A Bibliometric Analysis', *Journal of Academic Ethics*, 2025, pp. 1-16
- 33 Aziz, L.A.-R., and Andriansyah, Y.: 'The role artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance', *Reviews of Contemporary Business Analytics*, 2023, 6, (1), pp. 110-132
- 34 Spindel, J.: 'Arms for influence? The limits of Great Power leverage', *European Journal of International Security*, 2023, 8, (3), pp. 395-412
- 35 Willardson, S.L.: 'Under the influence of arms: The foreign policy causes and consequences of arms transfers' (The University of Iowa, 2013. 2013)
- 36 Johnson, J.: 'Inadvertent escalation in the age of intelligence machines: A new model for nuclear risk in the digital age', *European Journal of International Security*, 2022, 7, (3), pp. 337-359
- 37 Johnson, J.: "Catalytic nuclear war" in the age of artificial intelligence & autonomy: Emerging military technology and escalation risk between nuclear-armed states', *Journal of Strategic Studies*, 2021, pp. 1-41
- 38 Maas, M.M., Lucero-Matteucci, K., and Cooke, D.: '10. Military Artificial Intelligence as a Contributor to Global Catastrophic Risk', 2023
- 39 Baker, J.E.: 'Promise and peril: the next ten years of national security AI and its regulation': 'International Security Studies and Technology' (Edward Elgar Publishing, 2024), pp. 180-195
- 40 Ahmed, M.: 'Artificial Intelligence, Nuclear Warfare and Towards Laws: A Study on Limitations of Existing International Laws and Treaties', 2025
- 41 Horowitz, M.C., Scharre, P., and Velez-Green, A.: 'A stable nuclear future? The impact of autonomous systems and artificial intelligence', *arXiv preprint arXiv:1912.05291*, 2019
- 42 Jensen, B.M., Whyte, C., and Cuomo, S.: 'Algorithms at war: the promise, peril, and limits of artificial intelligence', *International Studies Review*, 2020, 22, (3), pp. 526-550
- 43 Haney, B.S.: 'The perils and promises of artificial general intelligence', *J. Legis.*, 2018, 45, pp. 151

-
- 44 Laton, D.T.: 'Manhattan_Project. exe: A Nuclear Option for the Digital Age', Cath. UJL & Tech, 2016, 25, pp. 94
- 45 De Spiegeleire, S., Maas, M., and Sweijs, T.: 'Artificial intelligence and the future of defense: strategic implications for small-and medium-sized force providers' (The Hague Centre for Strategic Studies, 2017. 2017)
- 46 Horowitz, M.C., Kahn, L., and Mahoney, C.: 'The future of military applications of artificial intelligence: A role for confidence-building measures?', *Orbis*, 2020, 64, (4), pp. 528-543
- 47 Zaidi, W., and Dafoe, A.: 'International Control of Powerful Technology: Lessons from the Baruch Plan for Nuclear Weapons', Centre for the Governance of AI, 2021
- 48 Chesterman, S.: 'Artificial intelligence and the problem of autonomy', *Notre Dame J. on Emerging Tech.*, 2020, 1, pp. 210
- 49 Calo, R.: 'Artificial intelligence policy: A primer and roadmap', *U. Bologna L. Rev.*, 2018, 3, pp. 180
- 50 Amigud, A.: 'The age of the Intelligent machine: Singularity, efficiency, and existential peril', *Philosophy & Technology*, 2024, 37, (2), pp. 49
- 51 Calo, R.: 'Artificial intelligence policy: a primer and roadmap', *UCDL Rev.*, 2017, 51, pp. 399
- 52 Mazzucato, M., Schaake, M., Krier, S., and Entsminger, J.: 'Governing artificial intelligence in the public interest', 2022
- 53 Jin, G.Z.: 'Artificial intelligence and consumer privacy': 'The economics of artificial intelligence: An agenda' (University of Chicago Press, 2018), pp. 439-462
- 54 Johnson, J.: 'Delegating strategic decision-making to machines: Dr. Strangelove Redux?', *Journal of Strategic Studies*, 2022, 45, (3), pp. 439-477
- 55 Nair, M.M., Deshmukh, A., and Tyagi, A.K.: 'Artificial intelligence for cyber security: Current trends and future challenges', *Automated secure computing for next-generation systems*, 2024, pp. 83-114
- 56 Biden Jr, J.R.: 'Interim national security strategic guidance', 2021
- 57 Gilli, A., Gilli, M., Leonard, A.-S., and Stanley-Lockman, Z.: '" NATO-Mation": Strategies for Leading in the Age of Artificial Intelligence' (NATO Defense College., 2022. 2022)
- 58 Samala, A.D., Rawas, S., Wang, T., Reed, J.M., Kim, J., Howard, N.-J., and Ertz, M.: 'Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects', *Education and Information Technologies*, 2025, 30, (3), pp. 3239-3278
- 59 Rudolph, J., bin Mohamed Ismail, M.F., and Popenici, S.: 'Higher education's generative artificial intelligence paradox: The meaning of chatbot mania', *Journal of University Teaching and Learning Practice*, 2024, 21, (6), pp. 14-48
- 60 Allen, G., and Chan, T.: 'Artificial intelligence and national security' (Belfer Center for Science and International Affairs Cambridge, MA, 2017. 2017)