Techade: An International e-Journal on Emerging Technologies and Power Dynamics
Vol. 2, Issue 1 | June 2026

CONVERGENCE OF ARTIFICIAL INTELLIGENCE AND SECURITY: A DEFINING CHALLENGE IN 21st CENTURY

Prof. Ramakrushna Pradhan[i] and Miss. Lohita Pradhan[ii]

Abstract

Artificial intelligence is rapidly transforming the global security environment, reshaping the ways states project power, manage risks, and pursue national interests. Once confined to laboratory settings and specialized industrial systems, AI has evolved into a ubiquitous infrastructure that underpins military modernization, cyber operations, intelligence analysis, surveillance architectures, and political communication systems. This article argues that the convergence of AI and security constitutes a defining challenge of the twenty-first century because it simultaneously enhances strategic capability and introduces unprecedented systemic vulnerabilities. Unlike earlier technological advances, AI operates through autonomous decision-making, adaptive behaviour, and high-speed data processing, altering foundational concepts such as deterrence, sovereignty, and crisis stability. The dual-use nature of AI blurs the boundaries between civilian innovation and military application, accelerating a technological arms race while undermining traditional governance mechanisms. The article situates AI within international relations theory, drawing on realism, liberal institutionalism, and constructivism to analyze how AI's integration into defence and security institutions reshapes global power dynamics. Through a multidisciplinary approach integrating security studies, AI technical research, geopolitics, and digital ethics, the study examines the drivers of AI–security convergence, the militarization of autonomous systems, the risks associated with algorithmic surveillance and disinformation, and the emerging governance challenges in a fragmented international order. The findings demonstrate that while AI provides significant opportunities for enhancing predictive capabilities, operational efficiency, and strategic foresight, it also carries profound risks for civil liberties, democratic resilience, and geopolitical stability. The article concludes by proposing a set of normative and institutional recommendations grounded in transparency, inclusive global dialogue, AI-system security, and public awareness. Ultimately,

the convergence of AI and security will shape not only the future of warfare and statecraft but also the nature of global order and the moral boundaries of technological power.

KEYWORDS

Artificial Intelligence, Global Security, Autonomous Weapons, algorithmic Governance, AI Arms Race

Introduction

The role of artificial intelligence has dramatically changed over time and has now become one of the key elements of today's security scenarios. The presence of AI in the whole chain of military networks, intelligence agencies, monitoring systems and computer security infrastructures is such that it speeds up the processes of threat detection and forecasting. It also allows the making of decisions by machines that are at a scale which exceeds human capability (Horowitz et al., 2019; Russell & Norvig, 2020). This transition is a major turning point in world security, as algorithms reduce decision-making timelines, modify deterrence and even mix the movements of offensive and defensive operations (Falk, 2025). On the other hand, the same attributes that reinforce security towers seem to be the source of complicated vulnerabilities. AI is a double-edged sword: in both democratic and authoritarian contexts, it enlarges the challenges of surveillance, disinformation, and establishing public trust, thus exposing the inherent political aspect of algorithmic governance (Zuboff, 2019; UNHRC, 2021). As the case for the use of AI in national security has become a matter of how and to what extent the confluence of AI and security will shape the power balance, the question now rests on the international order's normative foundations. This article is an exploration of such dynamics through the lenses of various disciplines, showing that AI is at the same time a facilitator, a systemic risk and the major challenge of the geopolitics of the 21st century.

Conceptual and Theoretical Framework

Defining Security in the AI Era

The concept of security, which was originally portrayed as merely safeguarding state sovereignty and preserving the territorial integrity of states, is now greatly bonus the other way round in the case of international relations today. Deterrence, intelligence, and battlefield operations, which are military activities are the areas AI quandaries in the future. On the political side, it will

determine the way of the flow of information, the integrity of elections and the capacity of governance. Cyber security, which was a very technical area, has now become a central part of national security because AI, on the one hand, reinforces and on the other hand, weakens the digital infrastructures. In the AI period it is henceforth reasonable to talk of "security" as being of multi-faceted nature which not only requires the investment in physical defense but also in information and the keeping of society, economy and the so-called norms in a stable state.

Defining Convergence

The term “convergence” is used to depict the process of gradual integration of AI into the system of security governance comprising institutions, technologies and operational practices. It signifies a situation where AI becomes a component of security operations without which it would be very difficult to imagine strategic decision-making, threat assessment, border control systems, infrastructure of surveillance, cyber operations and platforms of autonomous weapons. Convergence also hints at a rethinking of the concept: in the future security agents will be more and more depending on computational models, algorithmic predictions, and machine-speed analytics to interpret complex environments and guide their future strategic choices.

AI as a Tool of Power Maximization

Realist thinkers consider AI as an internationally competitive resource that militarily, politically, and economically serves the states in their sharing-forth process through the continuous deterrence of the present opponents and the achievement of their respective national interests. The application of AI to military, intelligence and cyber operations is regarded as a way to gain a relative advantage over others in the anarchic world system (Allison, 2021). Artificial Intelligence is a “force multiplier” that not only doubling the power of the existing armies but also affecting the coalition, the competition, and the crises in bigger or smaller ways.

Cooperation, Norms, and Governance

Liberal institutionalism presents international cooperation, governance regimes, and multilateral institutions as the main players in the game that finally will overcome the AI-related risks. The liberal theorists did not ignore the competitive aspect of the situation but rather highlighted the presence of shared vulnerabilities such as cyber attacks, algorithmic bias, autonomous escalation,

and digital misinformation which power the cooperation (Keohane, 1984). Transparency mechanisms, standard-setting bodies, and normative agreements will be the three institutional frameworks, where the first two will mainly be for reducing uncertainty and the third one for building trust. The problem that needs a solution is how to create governance structures that can cope with the rapid technological change.

Norms, Ethics, and Discursive Power

Constructivist methods emphasize the importance of norms, concepts, and ethical discussions in influencing the application of AI. The terms "meaningful human control," "algorithmic accountability," and "responsible AI" are significant in the worldwide discussions on the topics of autonomous weapons, surveillance, and digital rights (Blanchard,2024). They points out that AI is not a neutral entity. The technology's design along with its use, both, mirror the social values, political cultures, and institutional practices that are at the core (Finnemore & Hollis, 2016).

These theoretical frameworks, when combined, show that AI is not just a tech innovation but a structural force that reconfigures international order. Realism accounts for the competition and arms races; liberalism points out the gaps in governance and difficulties in cooperation; and constructivism highlights the moral and ethical debate around the limits of AI. This multi-theoretical perspective allows us to grasp thoroughly the convergence of AI and security as a complex phenomenon that interacts at the same time through material, institutional and ideational levels.

Drivers of AI–Security Convergence

The fast merging of AI and security is not an unplanned occurrence but a combination of several factors that are interlinked. These factors include technological, geopolitical, economic, institutional, and social aspects among others. All these things rapidly push the use of AI in military, intelligence, and governance systems thus making the technology a part of the very essence of modern security infrastructures. It is very important to understand these powerful forces as, without them, it would be hard to explain why the interaction between AI and security has been one of the most notable and structural changes in the twenty-first century.

Accelerating Technological Innovation

AI has been developed in a very rapid and at an exponential pace because of the coming of age of deep learning, new neural architectures, and the massive availability of computing power. At the same time, various areas like natural language processing, computer vision, generative models, and reinforcement learning have caught up so much that AI is now capable enough to take part in decision-making in complex situations such as surveillance, border control, cyber defense, and battlefield (Russell & Norvig, 2020). The merging of the two worlds is also being pushed forward by the growing dependability of automation: on the one hand, there are machines such as drones, robots, and sensors that have been made totally by human beings, and, on the other hand, there is also a growing error, which machines are capable of handling faster than humans. All of these rapidly advancing technologies are motivating the states to bring AI into their national security systems in order to keep their operations relevant in an era that is gradually becoming highly data-driven and automated.

Strategic Rivalries and the Emerging AI Arms Race

The power rivalry between the two superpowers, the United States and China particularly, is the main reason for AI militarization. The United States, along with China, has started a global competition for technology that ultimately determines their respective geopolitical status. The People's Liberation Army (PLA) benefits from the fast acquisition of advanced AI skills due to China's policy of civil-military fusion which connects commercial innovation with military development (Kania, 2020). On the other hand, the U.S. using its mastery in semiconductors, top-level research centers, and control of the private sector in basic AI models to keep a strategic lead (Allison, 2021). Both countries assume that the introduction of AI in war scenarios is inevitable and thus their actions contribute to a wider technology arms race. The two rival powers, plus more who do not wish to be left out indirectly support the militarization process as they already perceive it as a security concern.

Datafication and the Expansion of Global Digital Ecosystems

The tremendous amount of data generated by digital means has been the biggest factor in the adoption of AI. Security agencies have now moved to a realm where enormous and continuous streams of data are the basis for real-time monitoring, predictive threat analysis, and behavioral

modeling. AI is the one that makes it possible to process this information in a very large scale, spot the deviations, and help make decisions in the areas of intelligence, policing, and border management (UNHRC, 2021).

The Dual-Use Character of AI Technologies

The dual-use nature of AI, wherein civilian innovations could be very easily converted to military applications, so to say, hastens the convergence by making the boundary between the commercial and defense sectors less clear. For instance, no matter how different the intended application is (e.g., consumer analytics vs. targeting systems), the technology developed in one area can be adapted for the other with just a small amount of modification (Brundage et al., 2018). These technological dual-use dilemmas remove the barriers to states rushing to adopt the respective emerging technologies to be one step ahead of their rivals, thus contributing to open and widespread adoption of technologies even when no clear governance frameworks are in place.

Changing Nature of Warfare and Demand for Machine-Speed Decision Cycles

The modern battlefields have become excessively complicated, very fast, and full of information that is only partly usable by human operators. In cyber warfare, AI can not only detect intrusions but also implement countermeasures and even change its response depending on the enemy's behavior much quicker than human analysts (UNIDIR, 2023). In case of a conflict, AI takes part in real-time intelligence fusion, target recognition, logistics and even autonomously maneuvers. Hence, countries are more and more considering AI as the future necessary condition for winning and having operational supremacy in the wars of the future.

AI as a security enabler

The use of Artificial Intelligence has developed into a major factor enabling security, changing the operational capabilities of governments, and changing the methods used to find, control and eliminate threats. Mainly, AI technology is an all-in-one security solution. Unlike, past technologies, which collaborated with the security industry by providing single components, AI improves the whole process from intelligence gathering and analysis to strategic planning and tactical execution.

Early-Warning Systems and Predictive Security Analytics

AI is a major player in the transformation of early-warning systems, as it makes possible the real-time assessment of great and diversified datasets. The conventionally comprised intelligence processing was based on the heavy input of human analysts who had to work through and interpret fragmented data; unlike them, AI brings together data from numerous sources, such as satellites, sensors, social media, open-source intelligence, and even classified databases, to see through and detect the coming threats (Horowitz et al., 2019). Aimed at the future military mobilization, cyber-attacks and other such activities, extremer-wise or public health crises, machine learning models can already detect such events by uncovering their signs earlier than the events get observable.

Predictive technologies keep governments in the position of being able to anticipate crises instead of simply being reactive. Amongst the examples is the use of natural language processing that can spot the early signs of coordinated disinformation and also at the same time, anomaly detection systems can catch the unusual patterns of financial transactions which would be a signal for terrorism financing or organized crime networks.

Cyber security Enhancement and Automated Defence

AI is the biggest contributor to cyber security in the aspect of monitoring, threats to the system detection, and defensive actions being carried out by machines, all done continuously and at a very fast rate. While the traditional cyber-defence methods rely mainly on the identification of the known signatures or the heavily ruled manual updating, the latter ones driven by AI focus on the analysis of the network behaviour to discover the new threats, the so-called "zero-day" exploits, and the very sophisticated adversarial patterns, some of which have been recorded, and others not (Amodei et al., 2016). On the other hand, a machine-learning model that recognize and classify the malicious activity based on the very nature of each activity rather than the traditional way of using the static signatures, and thus becomes increasingly more effective in recognizing the sophisticated intrusions.

Intelligence, Surveillance, and Reconnaissance (ISR) Optimization

AI helps enormously the ISR systems by analyzing massive amounts of data from different media. The analysis of satellite imagery, which was a traffic-intensive job for human beings, can

now be carried out by machines using an accuracy-and-speed computer-vision-algorithms-based process that slices through the detections of troop movements, military buildup, or changes in infrastructure and does so with great precision (Kania, 2020). AI-powered geospatial applications are then bringing the military and intelligence planners such as the agencies now more than ever closer through awareness and hence better decision-making in fluid environments.

Autonomous surveillance systems equipped with facial recognition, gait analysis, and behavioural detection algorithms expand state capacity to monitor borders, public spaces, and critical infrastructure. This capability is especially valuable in counterterrorism, border security, and urban policing. Automated drone patrols powered by AI extend surveillance reach while reducing operational costs and human risk.

Decision-Support Systems and Strategic Forecasting

AI serves as a helper for decision makers in both military and government sectors, giving them tools that make the whole process easier by giving them complex information combined and making predictions. These tools are capable of creating war scenarios, evaluating the different ways in which a conflict might escalate, and so on (Falk, 2025). Reinforcement-learning algorithms might go through millions of strategic variations, picking out the most operative plans for the field. It has a wider scope than just military and intelligence; it can also help in the domestic security field through algorithmic governance, which is a tool for policing, public safety, and crisis management.

AI as a security challenge

Even though artificial intelligence has a significant ability to improve national security, it still poses a challenge with extensive deep-seated vulnerabilities, operational uncertainties, and moral issues, which are all intertwined with the stability of nations and the unity of societies. Besides the fact that these challenges are already present because of the intentioned wrongdoers taking advantage of AI systems, the technical aspects of AI like the lack of transparency, unpredictability, dependency on data, and the risk of being manipulated by adversaries also contribute to this. As AI continues to play an increasing role in the critical areas of society such

as infrastructure, politics, and the military the consequences of a failure, fraud, or misinterpretation are going to be extremely high. This segment concentrates on the multilayered security risks caused by algorithmic bias, adversarial attacks, disinformation, mass surveillance, data vulnerabilities, and the instability of crises.

Vulnerabilities in AI Systems: Hacking, Adversarial Attacks, and Data Poisoning

AI systems possess a particular weak point that allows their manipulation to be done through their technical design. One of the potential threats is adversarial attacks where a minor change in input data results in a completely incorrect classification (Amodei et al., 2016). A tiny version change could deceive a person-recognition system that would wrongly identify an automobile, a human being, or a piece of equipment; such an inaccuracy might lead to lethal outcomes if it happens to an autonomous weapon system or a platform for automated targeting. These weaknesses are not only being taken advantage of by perpetrators but also being used by state-backed opponents to gain a strategic edge over others. As AI becomes more widely used, the attack surface grows along its whole lifecycle and it becomes more likely that the disruption caused by algorithmic manipulation will not be limited to specific sectors - transportation, energy, finance - but rather, entire nations.

Deepfakes, Disinformation, and the Weaponization of Information

The integration of AI into the media industry has made it easier for certain actors to fabricate, alter and manipulate entire digital content on a large scale, thus, such a practice becomes one of the main reasons for information integrity issues. The existence of deepfakes, which are composed of visual and audio recordings that are created by the use of adversarial networks, has raised doubts about the reliability of the visual evidence (Chesney & Citron, 2019). The misuse of the technology can lead to issuing fake military commands, creating fictitious evidence of atrocities, and even escalating diplomatic crises in security environments.

The combination of such technologies is referred to as an “epistemic crisis,” wherein both the citizens and the policymakers are having a hard time telling apart what is real information and what is not. The democracies, which are heavily relying on the public discourse being informed and on the trust being placed in the institutions, are the ones mostly affected. Besides that, the AI-based disinformation campaigns are such that their effects are felt through the undermining of

the social fabric, the aggravation of political crack-ups, and the raising of the whole society's vulnerability to the external influence (UNHRC, 2021). On top of that, the very existence of automated propaganda bots, artificial accounts, and manipulation of recommender systems on such a large scale allows the opponents to take control over the entire populations with their narratives. The adoption of AI in information warfare has led to a decrease in the cost of conducting influence operations while simultaneously increasing their reach, thus, misinformation becomes a tool that any non-state or even powerful state actors can easily access (Wagman,2025).

Opacity, Unpredictability, and the Black-Box Problem

One of the biggest problems in AI is opacity; the majority of machine-learning systems cannot provide a rationale for their decisions. This "black-box problem" leads to a dramatic lack of accountability. The security staff might not be able to grasp the AI system, for example, pointed out a threat, tagged a target, or suggested a certain course of action. In an emergency, politicians might rely too much on algorithmic predictions, confusing mathematical certainty with correctness (Horowitz et.al,2019).

Crisis Instability and the Acceleration Problem

AI has the power to speed up tremendously the decision-making process in conflicts because it allows for machine-speed analysis, targeting, and response. The problem of acceleration decreases the chances of diplomatic de-escalation and at the same time raises the probability of crises expanding beyond the control of humans. An AI-enabled misinterpretation of radar data, social media signals, or satellite imagery could unleash premature military action, especially in tightly knit systems where human review is very limited.

States with nuclear warheads face especially dangerous situations: the use of AI in early-warning and command-and-control systems causes fear of false detections, misjudgment, or automated escalation during very tense times (UNIDIR, 2023). Nuclear deterrence combined with AI-powered automation is one of the largest security risks of the 21st century.

Moreover, AI is becoming an essential part of the government-operated electric grids, water plants, transportation systems, supply chains, and telecom domains. A problem such as a failure,

bug, or hack in any of these systems could affect communities and society as a whole. As we rely more and more on AI, the risk of systemic collapses rises concurrently. A cyberattack that is precisely directed to an AI-supported power grid or logistics system could put a country's economy down and render its military unprepared.

Geopolitical and Global Governance Dimensions

AI was the leading factor in the changing global geopolitics that made it problematic for nations to play their power-related cards, build alliances, and establish international norms. The role of AI as the global order determinant has been compared even to that of nuclear weapons during the Cold War or to industrialization in previous centuries. The implementation of an AI-like technology that provides universal access will re-map all the state capacities and their strategic behaviors across all the domains such as military operations, finance, manufacturing, surveillance, cyber conflict, and political communication (Allison, 2021; Kania, 2020). The techworld is not homogenous; a small number of states and private companies are controlling the lion's share. This kind of inequality comes up with the new power formations and exacerbates the existing geopolitical conflicts, as countries are more and more regarding AI capability as an indicator of national prestige, economic strength, and military power.

The escalating rivalry between the U.S. and China can be seen as the primary outline of the AI geopolitics focus. China's "civil–military fusion" strategy and its "Next Generation AI Development Plan" are entwined with military modernization, political control, and the fusion of the commercial sector's innovations, thus positioning AI as a tool for controlling both competitive power and society (Kania, 2020). China's vast biometric surveillance systems, the application of predictive policing algorithms in Xinjiang, and the funding of autonomous weapons are examples where not only the control of the population is said to be assisted by AI but also the power projection outside the borders (Zuboff, 2019; Allison, 2021).On the other hand, the U.S. is focused on an innovation-driven approach that relies mainly on the private sector, semiconductor leadership, and university research ecosystems.

Through the National AI Initiative, the U.S. government presents AI as a basic element of military preparedness, economic competition, and the struggle over ideas (Falk, 2025). The

Pentagon's Joint All-Domain Command and Control (JADC2) system, developed by the Pentagon and its allies, is one of the places where the integration of AI into defence networks is being done strategically to speed up decision-making and improve operations across multiple domains (Horowitz et al., 2019).

Technological competition is not the only aspect of this rivalry; it is a struggle over the future political order. China is advocating for a model of AI governance that takes state power into account and is heavily monitored, whereas the U.S. and its democratic allies argue for a model based on openness, privacy, and human rights (UNESCO, 2021). These contrasting paths show an emerging ideological division that not only affects the current state of international relations but also creates new alliances, controls over cyberspace, and the setting of global standards. Moreover, as both powers attempt to spread their governance models through international institutions and bilateral connections, AI gets deeply involved in the fight for the acceptance of certain norms.

The middle powers like India, Japan, South Korea, Israel, the UK, and the European Union's countries are opting for the path of technological self-reliance and thus trying to limit their dependency on either the US or China. The National AI Mission and other investments in digital public infrastructure by India show that the country is aspiring for "strategic autonomy" in AI development (Wadhwa, 2022). Similar to this, the European Union, through the EU AI Act, is trying to get the status of a world regulatory leader by advocating for ethical and rights-based governance thus having an indirect say in the global norms even if they are not technologically dominant. On the other hand, Japan and South Korea are utilizing their strengths in robotics and semiconductor manufacturing while having their security interests aligned with the US. Israel still continues to strengthen its cyber and defence AI industries which are already the most influential in the world thus it keeps the position of a technological hub. All these strategies display a collective turning point to an AI multipolar construct where regional powers intend to influence global governance while keeping their options open to switch between the great-power blocs.

The Global South encounters a distinctly different list of problems to tackle. The lack of access to top-notch computing, skilled personnel, and investment capital is the case in many developing

countries, leading to a technological gap that is getting wider and might end up keeping the latter forever depending on the former, while, at the same time, multinationals would be controlling the global flow of technology through such disparity. In fact, the scenario created is termed "AI peripheries," where countries have to import AI systems mainly developed in political and cultural contexts too different from theirs and are also unable to handle the underlying data infrastructures (UNESCO,2021). This makes the security of the country vulnerable and at the same time, policy autonomy is reduced, and the government is forced to listen to and even accommodate the demands of the foreign players. Besides, the influx of cheap AI surveillance technologies from China and private companies is a factor that increases the risk of digital authoritarianism in unstable democracies; thus, the violation of human rights is further aggravated. Moreover, as dependency grows, the Global South's participation in the discourse over global AI norms gets throttled which only serves to perpetuate the already existing structural inequalities.

The global alliance scenario is also being reshaped by AI. Recent AI-oriented partnerships such as the Quad's Critical and Emerging Technologies Working Group, NATO's AI governance frameworks, and the transatlantic cooperation through the EU–U.S. Trade and Technology Council are testimony of how states take advantage of alliances to gain control over the supply chains, synchronize standards, and counter the technological influence of China (UNIDIR, 2023). These partnerships integrate AI into intelligence sharing, military interoperability, and ethical standard-setting, reflecting the technology's centrality to strategic coordination. At the same time, many regional organizations, including ASEAN and the African Union, seek cooperative mechanisms for managing cross-border data flows and mitigating technological dependency, though capacity constraints limit their influence.

Nevertheless, the global governance institutions still have difficulty keeping up with the fast pace of AI's development. Although places like UN, ICRC, and UNIDIR have not only advocated for the formation of norms directing the use of autonomous weapons but also committed to discussing proper accountability of algorithms and data protection, the power struggle among the major states has been a big obstacle to the process of getting through the opinions of the various groups (ICRC, 2020; UNIDIR, 2023). The failure of the UN Convention on Certain Conventional Weapons (CCW) to agree on any binding agreements regarding lethal autonomous

weapons systems (LAWS) is an indicator of the difficulty. There is no clear material signature of AI like that of nuclear or chemical weapons which makes it impossible for traditional verification regimes to work effectively. It is very easy for software-based systems to be hidden, changed, or copied which makes monitoring hard and reduces trust between states. Scholars therefore suggest new governance instruments such as algorithmic auditing, transparency obligations, incident reporting, safety certification, and global registries for high-risk AI systems which would work alongside rather than replace traditional arms-control approaches (Brundage et al., 2018). A further governance challenge arises from global regulatory fragmentation. The EU's rights-based approach contrasts sharply with China's regulatory algorithm framework, and both differ from the US' innovation-first model. These competing regimes risk establishing standards that are not only incompatible but also interoperability hindering, increasing regulatory arbitrage, and producing safety environments with different levels across regions (UNESCO, 2021). With the countries going for different rules on facial recognition, data transfers, autonomous weapons, content moderation, and algorithmic transparency, the world is getting divided into a governance landscape that is not only shaped by geopolitical competition but also by a lack of cooperation.

AI geopolitics might unfold in different ways, looking ahead. The world may witness, in one scenario, a confrontation between the U.S. and China where countries adopt such different technologies that they split into two opposite sides. Another possibility has the regional powers creating their own separate AI worlds and the whole thing turning into a very complicated and possibly unstable governance system. Yet another scenario speaks of the emergence of an international cooperative governance but with the major global players needing to build up unprecedented trust levels first. Current trends, however, point to a competitive multipolarity, which is marked by nationalistic industrial policies, the arms market race, and different ethical frameworks.

Looking forward, the geopolitical and governance aspects of AI show that the technology is much more than a new statecraft tool; it is a reshaping force of the global political architecture. By speeding up the debate between the superpowers, intensifying the unequal situation, splitting governance, and changing the power base of alliances, AI has already become a decisive factor in the future of the international order. If states don't find a way to create inclusive, enforceable,

and adaptive governance mechanisms, AI will continue to be a trigger for geopolitical instability rather than a pillar of collective security**.**

POLICY AND GOVERNANCE RECOMMENDATIONS

Tackling the risks caused by the merging of artificial intelligence with security needs a governance system that is principled, multidimensional, and characterized by transparency, global cooperation, public awareness, technical safeguards, and ethical oversight. Governments need to create conditions for independent auditing, have a system of clear accountability for AI-driven actions, and increase the transparency of the algorithms in order to gain public trust. A global forum on AI and security could be created based on the IAEA model that would have the capability to set standards, deal with the risks of autonomous weapons, and stop the escalation caused by AI. Increasing the society's understanding is just as important as AI is the one that will mostly influence the areas of policing, administration, and information ecosystems. The governments should apply the measures of safety-by-design, adversarial testing, protected datasets, and specialized response units not only to AI systems but also to their security. In the end, the support of human supervision and the incorporation of ethical principles in the high-stakes areas will guarantee that AI will be the ally of collective security and the opposite of the undermining of democratic norms or global stability.

Conclusion

“In the final reckoning, the convergence of AI and security will define not just how nations protect themselves, but how they shape the world that follows.”

AI has become embedded in core security functions from military modernization and cyber operations to intelligence processing and surveillance, fundamentally reshaping deterrence and speed up the decision-making processes at a very high level. The integration of AI, however, came along with some major vulnerability, such as system manipulation, unintended escalation of hostilities, and democratic risks arising from the use of biased algorithms, surveillance, and political influence. The dual-use aspect of AI makes it difficult to regulate the technology, as civilian tools can be rapidly militarized (Pauwels, 2018). From a geopolitical standpoint, AI is a rivalry among the great powers that gets more intense, is a factor in the fragmentation of global

governance, and demonstrates the incapacity of the existing global institutions like the UN and CCW to deal with the debate on autonomous weapons at the same level as it is conducted. The measures for governance suggested transparency, a global AI-security forum, public awareness, AI-system protection, and ethical oversight to underscore the truth that AI is not only a technological challenge but also a political and societal one. The question of whether AI will be an asset or liability in security, therefore, lies not only in the hands of the innovators but also on the politicians who will decide whether to cooperate or compete and whether to regulate for peace or arm for a potential conflict.

References

Allison, G. (2021). *AI and national power: The geopolitical implications of artificial intelligence*. Belfer Center for Science and International Affairs.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). *Concrete problems in AI safety*.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, J., & Amodei, D. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute.

Blanchard, A. (2024, April 25). *The road less travelled: Ethics in the international regulatory debate on autonomous weapon systems*. International Committee of the Red Cross (ICRC)

Buzan, B., Wæver, O., & de Wilde, J. (1998). *Security: A new framework for analysis*. Lynne Rienner.

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1819.

Falk, R. (2025). *AI, strategic stability, and the future of deterrence*. Oxford University Press.

Finnemore, M., & Hollis, D. B. (2016). Constructing norms for global cybersecurity. *American Journal of International Law, 110*(3), 425–479.

Horowitz, M. C., Scharre, P., Payne, K., & Vavrečík, T. (2019). *Artificial intelligence and international security*. Center for a New American Security (CNAS).

ICRC. (2020). *Autonomy, artificial intelligence and robotics: Technical and legal challenges*. International Committee of the Red Cross.

Kania, E. (2020). *Battlefield singularity: Artificial intelligence, military revolution, and China's future military power*. Center for a New American Security (CNAS).

Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton University Press.

Pauwels, E. (2018). *The new geopolitics of artificial intelligence*. United Nations University Centre for Policy Research.

Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

UNHRC. (2021). *The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights*. United Nations Human Rights Council.

UNIDIR. (2023). *AI and strategic stability: Risks, governance, and global security*. United Nations Institute for Disarmament Research.

Wadhwa, V. (2022). *India's AI mission and strategic autonomy*. Observer Research Foundation.

Wagman, J. (2025). *Synthetic influence and the future of information warfare*. Cambridge University Press.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

[i] **Prof. R. K. Pradhan** is the Dean of School of Social Science, and Head of the Department, Department of Political Science, Guru Ghasidas Vishwavidyalaya (A Central University), Bilaspur, Chhattisgarh. A Ph.D. from JNU, Prof. Pradhan also served as the Visiting Professor at L. N. Gumilyov Eurasian National University, Astana, Kazakhstan. He can be contacted at: rkpradhanjnu@gmail.com.

[ii] **Miss. Lohita Pradhan** is Post-Graduate scholar in the Department of Political Science, Guru Ghasidas Vishwavidyalaya (A Central University), Bilaspur, Chhattisgarh.