USI Occasional Paper

No 3 - 2026

Major General Pawan Anand, AVSM, PhD (Retd)

Colonel Ashish Nagaich, SM

About the Occasional Paper

This paper examines the growing integration of Artificial Intelligence (AI) into military systems within an increasingly multipolar and contested global environment, highlighting both its transformative potential an d inherent risks. It argues that incidents such as misidentification and fratricide underscore the dangers of uncritical reliance on AI and the urgent need for 'Responsible AI in Military'. The paper analyses the unique operational, technical, legal, ethical, and political challenges associated with military AI, including complexity, adversarial environments, and accountability gaps. It traces the evolution of Confidence-Building Measures (CBMs) from Cold War precedents and advocates their adaptation to AI governance. While emphasising transparency, cooperation, and human control, it proposes capability- and behaviour-based CBMs to mitigate risks of escalation and instability. The paper concludes that, in the absence of enforceable global regulations, multilateral CBMs offer the most practical pathway to ensure responsible, secure, and trustworthy deployment of AI in warfare.

Introduction

As the world moves towards multi-polarity amidst contestations, conflicts, trade wars, and tariff regimes, one technology that has established itself as the most defining and transformational contributing factor impacting every field today is Artificial Intelligence (AI). In today's multi-domain environment where nation states are trying to impose their will on others, AI is all poised to shape the future of warfare in multiple ways. However, the civilian deaths due to alignment and identification errors in Israel's Lavender AI Systems[1], or losing costly equipment to fratricide by Ukraine's Bumblebee AI systems[2], have exposed the problems in blindly trusting the technology and a requirement of responsible fielding and governance.

The fast pace of advancements in this niche field, global AI race, and uncertainty around the development and integration of AI-based military systems preclude their effective governance and compliance to standards, which often trail their deployment. Presently, it is mostly left to voluntary adoption of certain governance principles, with little articulation of the details and methodology for their implementation, adherence, and compliance. The complexities involved in design, development, and fielding of Responsible AI in Military (ReAIM) systems and challenges thereof introduce a new stack of risks to the security environment. Traditional instruments dealing with risk prevention and incident management in the international security environment may not be equipped and capable of handling these emerging and evolving challenges and risks emanating from the use of AI in military domain. The

inevitable push for increasing use of AI in warfare makes it imperative for the international community to prioritise formulation, adoption, and adherence to Confidence-Building Measures (CBMs) to mitigate the risks of misuse of these systems, which may lead to unforeseen incidents and inadvertent escalation of situations and dangerous consequences.

Responsible Artificial Intelligence

In the military domain, the application of Responsible AI (RAI) assumes heightened significance due to the unique operational, ethical, and strategic implications involved. It is an approach to designing, developing, assessing, and deploying AI systems in a way that is safe, ethical, and trustworthy, keeping human goals and values central to the design.[3] It involves a set of principles that guide the design, development, deployment, and use of AI, ultimately building trust in AI solutions that empower organisations and all stakeholders. Key aspects of RAI[4] include the following.

In the military domain, the application of responsible artificial intelligence assumes heightened significance due to the unique operational, ethical, and strategic implications involved.

- **Alignment with Values**. Development and fielding of RAI systems need to consider the broader societal impact of the same and should be aligned to the

ethical standards, values, legal standards accepted by the society.

- **Focus on Outcomes**. RAI acts as a guide for AI decision makers, from defining the basic purpose of systems to user interaction and outcomes.

- **Mitigation of Risk**. The aim of RAI is to embed ethical principles into AI applications and workflows to mitigate the risks and negative outcomes associated with use of these systems, while maximising positive outcomes.

- **Framework for Challenges**. RAI serves as a framework for organisations to figure out how to address legal and ethical challenges concerning AI systems.

Responsible Artificial Intelligence for Military

Pending global consensus and standardisation of the principles of RAI, their adoption, and implementation frameworks have been left to various nations to decide upon. However, following peculiarities of military applications make it critically important for them to imbibe responsibility throughout their lifecycle.

High Impact. Military applications have a significantly high impact factor compared to other departments, which could even involve use of lethal force and take precious human lives.

Military applications have a significantly high impact factor compared to other departments, which could even involve use of lethal force and take precious human lives.

This risk of making an error while taking existential decisions differentiates it from other sectors or departments where the consequences may only lead to financial loss.

Enormous Complexity. The functions and decisions involved in military AI systems are often enormously complex and a combination of multiple factors, both tangible and intangible. Many of them are not easily measurable or clearly discernible, forcing decisions which considerably depend on intuitive reasoning and experience. The systems often comprise of multiple sub-systems, working on different sub-fields of AI.

Adversarial Environment. Unlike other deployments, military AI systems must operate against an adaptive, dynamic, and intelligent adversary, making it difficult to work in an intended manner by deceiving, manipulating, or neutralising the technology. Moreover, these adverse actions are not available to check or verify the systems in normal course of deployment and fielding. Thus, robustness against unknown cyber threats, electronic warfare, and deliberate deception are more important for military AI systems compared to other systems.

Operational Momentum. The pace of operations and events often necessitates the decisions to be taken in such short time spans that puts immense pressure on the decision makers and their team.

Military artificial intelligence systems must operate against an adaptive, dynamic, and intelligent adversary, making it difficult to work in an intended manner by deceiving, manipulating, or neutralising the technology.

This increases the probability of these decisions to be error prone.

Asymmetric Risks Involved. Risks involved with military systems and decisions are asymmetric in nature and weight of consequences and strategic ramifications could be very heavy, shrinking the leeway for errors and corrections.

Compliance to International Humanitarian Law (IHL). The legal and ethical obligations of compliance to IHL[5] and Laws of Armed Conflict, in addition to other local governance and legal bindings, far exceed these bindings in civil contexts.

Constantly Evolving Scenarios. Unlike commercial world, the military threats and operational scenarios are very dynamic and ever evolving. This enhances the degree of difficulty and challenges of building a system, which works equally effectively in all different scenarios.

Limitation Related to Training of the System. Training data to train the systems is often not readily available owing to the security classification and concerns or other technical reasons, precluding requisite training of the AI-based systems being planned or fielded.

Miscellaneous Aspects. Other miscellaneous aspects like non-availability of requisite hardware and software, integration with legacy systems, heterogeneity of equipment and systems, inadequately skilled human resource, long procurement

Unlike commercial world, the military threats and operational scenarios are very dynamic and ever evolving.

gestation periods, and policy restrictions, etc. considerably differentiate the fielding environment of AI systems in military[6], as compared to other organisations or departments.

Challenges to Responsible Artificial Intelligence for Military

The challenges to implementation of ReAIM can be divided into following categories:

- **Operational Challenges**. Immense complexity of the military AI systems and their dependence on various tangible and intangible factors make their fielding extremely challenging. The concerns around the warfare becoming impersonalised—devoid of compassion, empathy, reality, and adherence to IHL and Rules of Engagement (RoE) accentuate the challenges. Use of AI-based autonomous systems and platforms are bound to commence an AI arms race, to field better and more sophisticated systems into warfare, leading to enhanced lethality and precision.

- **Technical Challenges**. The technical challenges of designing, developing, and deploying secure ReAIM systems get pronounced due to

Immense complexity of the military artificial intelligence systems and their dependence on various tangible and intangible factors make their fielding extremely challenging.

constraints due to difficult service conditions and data availability.

- **Moral Challenges**. It is a significant challenge to make AI systems imbibe human moral qualities[7], as many military decisions are a mix of human emotions and professional knowledge and experiences. The challenge is to have AI systems also imbibe some emotions and use the correct, healthy mix of the same to take professional decisions. There is also a danger of AI systems falling into wrong hands or being used for inimical activities.

- **Legal Challenges**. These challenges emanate from the dissipation of responsibility while employing a complex AI system.[8] Who could be held responsible for an illegal behaviour of such a system—designer, developer, user, operator, or decision maker? The challenge is to program the principles of IHL and legal bindings into the AI systems and make them abide by the same in all conditions and situations.

- **Political Challenges**. Availability of new and relatively easier military technology may lead to a tendency lower than the threshold of conflict, thereby, violating the *Jus Ad Bellum* (Right to War) principle of the IHL and blaming the consequences to 'Deviations' by AI.[9]

It is a significant challenge to make artificial intelligence systems imbibe human moral qualities.

- **Social Challenges**. AI systems would influence a soldier's psychological state. The positive effect could be to reduce stress; however, use of autonomous systems would make the battle anonymous. They could make mistakes, with severe consequences and casualties, which could lead to severe mental and psychological stress and problems.

Background of Confidence-Building Measures

These measures were coined during the Cold War, when the United States and Soviet Union tried to establish communication channels to reduce the risk of nuclear conflict.[10] These comprised of the 'Hotline Agreement' (1963), 'Incidents at Sea' (1972), and 'Helsinki Final Act' (1975), which allowed the voluntary exchange of information and notification of major military activities. In the context of militaries and arms control, CBMs can be defined[11] as "Planned procedures to prevent hostilities, to avert escalation, to reduce military tensions, and to build mutual trust between countries". These measures are fundamentally voluntary and flexible, and can be unilateral, bilateral, or multilateral in nature, and are meant to clarify the intent and establish the basic understanding of military actions, exercises, and operational constraints under which the systems could be employed. They fall in the following two broad categories:

Confidence-building measures can be defined as "Planned procedures to prevent hostilities, to avert escalation, to reduce military tensions, and to build mutual trust between countries'.

- **Capability-based CBMs**. These refer to disclosures of capabilities of AI systems that a state possesses.

- **Behaviour-based CBMs**. These refer to the way a state would act in a specific operational situation.

Requirement of Confidence-Building Measures in Responsible Artificial Intelligence for Military

Integration of AI into critical and operational military systems introduces a complex landscape of multi-dimensional risks. These unique taxonomies of challenges and risks emanate from technical vulnerabilities and operational failures, which in military domain could have profound ethical and strategic implications. The severe nature of the damage that these systems can inflict necessitates a specialised trust and governance framework for building confidence between the states planning to use these systems. In the realm of responsible use of AI, CBMs[12] are flexible, voluntary, and mutually agreed tools designed to build transparency and trust in AI systems and give out rules of their utilisation. These developments of a shared understanding of goals, capabilities, limitations, risks, and challenges of use of military

Integration of artificial intelligence into critical and operational military systems introduces a complex landscape of multi-dimensional risks.

AI systems are aimed to prevent competitive pressures, which could lead to irresponsible and unsafe use of these systems. The need for these CBMs emanate from the following risks and challenges.

Conflicting Paradigms. Fundamentally, AI is a probabilistic technology[13], heavily dependent on the ability of machines to learn cognitive functions, mostly by iterative feedback, which keep increasing the probability of the machines taking 'Correct' decisions. In contrast, the desired outcome of all military operations is a "Definite and assured victory over the enemy". It is for the first time that a probabilistic technology is being employed for achieving definitive results.

Technical Vulnerabilities. AI systems have the following vulnerabilities, which could result into unique technical failures in complex and high-stakes applications:

- **Algorithmic Brittleness**. AI systems often perform sub-optimally or even fail when they encounter scenarios that deviate even slightly from their training data. In a complex military operational environment, which is a function of numerous factors, training a system in exactly similar environment is next to impossible. Hence, the behaviour of systems in such an environment is always suspected.

Artificial intelligence systems often perform sub-optimally or even fail when they encounter scenarios that deviate even slightly from their training data.

- **'Black Box' Problem**. Neural networks-based AI systems lack interpretability, making it difficult for human operators to understand the rationale behind a specific decisions or outputs. This makes it difficult for the user or commander to establish trust on the system.

- **Cybersecurity Risks**. AI systems are vulnerable to hostile or adversarial attacks, including hacking, data poisoning, spoofing, and manipulation of sensors.

- **Design Vulnerabilities**. Functionality is often prioritised over security requirements while initially designing a system, often leading to a software that is inherently insecure in hostile or contested environments.

Operational and Strategic Risks. The speed and autonomy of AI can lead to unintended outcomes that may be impossible to correct in time. These include the following:

- **Unintended Targeting and Inadvertent Escalation**. Incorrect logic or wrong target identification can result in use of force against unintended objects or persons, resulting in an unacceptable collateral damage and can spiral into inadvertent escalation.

The speed and autonomy of artificial. intelligence can lead to unintended outcomes that may be impossible to correct in time.

- **Strategic Instability**. The speed of AI systems creates 'First-Strike Instability', where actors are incentivised to strike pre-emptively to gain a tactical advantage before an adversary's system can react.

- **Magnitude of Failure**. Risks are significantly aggravated if AI is integrated with weapons of mass destruction or launch platforms.

Human–Machine Teaming Challenges. There are significant psychological and coordination hurdles in operationalising the man–machine teaming. These could be categorised as under:

- **Automation Bias**. This represents a condition when users or operators blindly trust the AI system, without checks and balances. They may under-trust it, resulting into repeated overrides, thereby, nullifying the advantages of the same.
- **Cognitive Overload**. The logic behind AI Systems' decision-making and opting for a specific course of action may be considerably difficult for human users or operators to understand. This may result into cognitive overload and trust deficit on the system itself.
- **Meaningful Human Control**. Maintaining appropriate control and judgement over critical decision-making and utilisation of

There are significant psychological and coordination hurdles in operationalising the man–machine teaming.

force, especially in Lethal Autonomous Weapons Systems (LAWS), remains a challenge.

Ethical and Legal Aspects. The undermentioned ethical and legal aspects merit consideration in respect of the AI adoption for critical functions:

- **Unclear Accountability**. The clarity with respect to the accountability of erroneous decision between the programmer, the system manufacturer or the user or commander is difficult to establish, especially in case of deep neural network[14] based systems.
- **Legal Compliance**. Legal compliance with respect to AI being able to perform the qualitative and subjective assessments required for the principles of IHL i.e., distinction, proportionality, and military necessity in complex combat zones, is difficult to achieve.
- **Dehumanisation**. Algorithms deciding over the use of force and life-and-death decisions may not be considered moral and deem to de-humanise warfare.

Governance and Transparency Challenges. Following governance challenges call for trust building through CBMs:

Algorithms deciding over the use of force and life-and-death decisions may not be considered moral and deem to de-humanise warfare.

- **Lack of Global Norms or Laws**. There is no international law, treaty, or even consensus on regulating or restricting use of AI and LAWS.
- **AI Arms Race**. AI arms race and associated pressures may propel nations to take shortcuts while testing, validation, and evaluation to deploy technology faster than adversaries. This, coupled with algorithmic opacity, Intellectual Property (IP) rights, trade secrets, and national security confidentiality may result in deployment of AI systems[15] getting into a dangerous spiral, laden with irresponsible and irrational behaviour.

Building Confidence in Responsible Artificial Intelligence for Military Systems

A combination of capability and behaviour-based measures are required to build credible confidence in utilisation of a complex technology like AI in critical military systems. Though it is certain that voluntary declaration of actual capabilities by belligerent nations may just be a utopian and theoretical construct, institutionalisation of behaviour-based CBMs is a definite possibility and is being attempted by various international organisations and forums like the United Nations, Organisation for Economic Cooperation and Development, and Institute of Electrical and

Artificial intelligence arms race and associated pressures may propel nations to take shortcuts while testing, validation, and evaluation to deploy technology faster than adversaries.

Electronics Engineers. These measures aim at achieving transparency and cooperation.

Measures for Transparency

Transparency involves disclosures of information to allow users or auditors to monitor internal functioning of the systems. This could be achieved through the following measures:

- **Voluntary Information Exchange on Policies**. Sharing of AI strategies, ethical principles, and legislative frameworks by nation states.
- **Creation of Algorithmic Registries**. Accessible repositories[16] of algorithms to help understand how various AI systems work and data used to train them.
- **Dialogues between Stakeholders**. Periodic exchange of ideas, policies, procedures, doctrines, and RoE between stakeholders including militaries to improve understanding of the capabilities and application of AI-enabled systems.

Measures for Cooperation

These measures aim at establishing common understanding[17] of the concepts and approach towards development and deployment of AI systems:

Transparency involves disclosures of information to allow users or auditors to monitor internal functioning of the systems.

- **Designate Points of Contact (PoCs)**. Specific PoCs for technical, policy, and diplomatic functions could enhance cooperation and ensure faster interaction during incidents.
- **Shared Terminology**. A common glossary or dictionary of technical terms could avoid mis-interpretations and enhance understanding of each other's actions.
- **Capacity Building**. Integration of the nations of Global South through capacity and capability development and building trust.
- **Sharing of Best Practices**. Best practices in the field of design, development and deployment of AI systems and use-cases may be shared between the nations to increase cooperation.
- **Legal Reviews (Article 36)**. Sharing of information on legal aspects of ReAIM systems and related technologies and methodologies for conducting legal reviews of AI-based weapon systems would enhance the cooperation in this field.

Best practices in the field of design, development and deployment of artificial intelligence systems and use-cases may be shared between the nations to increase cooperation.

Recommendations for Confidence-Building Measures

Policy and Governance.

- **ReAIM Doctrines and Red Lines**. It is recommended that nations should clearly lay down their ReAIM doctrines and policies, with specific mention of their red lines to establish accountability and trust.
- **Human Control**. It has been an established norm to have significant human control over critical decision-making by AI systems to include targeting and weapon release. Mechanisms to ensure its implementation are recommended to be formulated[18] and promulgated for ensuring the same.
- **Responsibility Frameworks**. Nations are recommended to not only give out strategic constructs but also their implementation frameworks[19] for ReAIM, including their oversight mechanisms for ReAIM systems.

It has been an established norm to have significant human control over critical decision making by artificial intelligence systems, to include targeting and weapon release.

Transparency and Communication.

- **Articulation of Principles**. Nation states are recommended to not only clearly articulate their ReAIM principles[20], but also to give out the

procedural implementation aspects with respect to the same.

- **Operational Communication Channels**. Establishment of effective multilateral communication channels between nations to clarify capabilities, doctrines, and code of conduct. Crisis communication hotlines[21] for AI-related incident reporting and management may be thought of.
- **Voluntary Disclosures**. As a norm, high-level military AI applications and breakthroughs that may engender autonomous warfare and escalation should be voluntarily disclosed.

Technical and Design Aspects.

- **Responsible Design and Deployment**. The designing and deployment of RAI systems with a thorough understanding of requirement and constraints is an operational imperative.
- **Explainable and Auditable Systems**. Responsible by design AI systems for mission-critical applications must be complied with, to keep them explainable and auditable.

As a norm, high-level military artificial intelligence applications and breakthroughs that may engender autonomous warfare and escalation, should be voluntarily disclosed.

- **Graceful Degradation**. ReAIM systems could incorporate features like fail-safe and graceful degradation modes to counter cyber threats.

Operational CBMs.

- **Gradual Enhancement of Roles**. As the technology matures and systems get tested and validated, enhancement of roles like decision support to semi-autonomous weapons systems could be considered.
- **Operator Training and Certifications**. Training on responsible usage, limitations, ethical and legal considerations of AI systems is an operational imperative for their successful deployment. Periodic reviews and certifications are necessary for training on updates and improvements in the systems.
- **Sharing of Experiences**. Sharing of experiences to include incident reports, RoEs, and after-action reviews with other nations to take corrective actions and improvements.

Testing, Evaluation, and Assurance CBMs.

- **Simulation and Testing**. Nations must undertake simulation-based testing for maximum possible operational conditions prior to fielding of AI systems to minimise errors.

Periodic reviews and certifications are necessary for training on updates and improvements in the systems.

- **Red-Teaming and Testing**. Red-teaming based on adversarial testing and audits to improve the system and ensure traceability and forensics must be made compulsory for deployment.
- **Lifecycle Governance**. Governing AI system across their lifecycle, from design to deployment and exploitation, must be considered with re-calibration after major model and data updates.

International and Multilateral CBMs.

- **Dialogues**. Multilateral dialogues on ReAIM systems (Track 1, 1.5, and 2)[22] to build confidence.
- **Sharing of Norms**. Sharing of various norms on human and escalation control and steps to align the code of conduct with IHL.
- **Joint Exercises and Information Sharing**. Joint table-top exercises on AI-related crisis scenarios and exchange of information on safety, on the lines of aviation and cyber safety, would go a long way in building confidence.

Legal and Ethical CBMs.

- **Legal Alignment**. Explicit alignment with IHL, constitution, and Laws of Armed Conflict and periodic legal reviews would facilitate confidence building.

Governing artificial intelligence system across their lifecycle must be considered with re-calibration after major model and data updates.

- **Ethical Review**. Formulation of ethical review boards within the organisation to institute mechanisms for investigation and correct AI-related issues would build confidence.

Trust within Military Ecosystem.

- **Sovereign Model and Systems**. Development of sovereign AI models and systems to ensure training, testing, and deployment aligned to own problems is the need of the hour in order to instil trust.
- **Collaborations**. Civil–military–industry–academia collaboration for working towards AI safety and adherence to standards.
- **Interoperability and Standardisation**. Common standards for interoperability and testing across the three services and with trusted partners would enhance confidence.

Challenges in Implementing CBMs in ReAIM. Despite the immense benefits of implementing CBMs in ReAIM, it faces following important challenges:

- **Design and Development**. Understanding the exact requirement, design, and development of accurate sovereign models, availability of accurate training data thorough

Common standards for interoperability and testing across the three services and with trusted partners would enhance confidence.

testing and unbiased deployment of AI military systems is a very complex and difficult task, which gets accentuated by limited understanding and short tenures of commanders or users.

- **Technical Gaps**. Lack of common terminology regarding what constitutes AI or 'Autonomy' in a military context hinders formulation of protocols and rules, which are a pre-requisite for building confidence. 'Black Box' nature of deep neural networks-based systems make transparency technically challenging.

- **Trade-off between Principles**. Despite of adoption of ReAIM principles by various nations, there is a technical limitation of 'Ability to Code' these theoretical constructs and principles into algorithms. There is also a challenge of these being divergent to each other, leading to a requirement of trade-off between them, based on the applications or functions that these systems are designed to perform.

- **National Security and IP**. Considerable reluctance in disclosing military AI capabilities due to confidentiality requirements and tendency of vendors to strictly protect their IP to maintain their lead in competition reduces confidence in the systems.

'Black Box' nature of deep neural networks-based systems make transparency technically challenging.

- **Verification Problem**. Unlike physical platform-centric weapon systems, verification of software-based AI systems is very difficult and complex to accomplish.

- **Performance Transparency**. This refer to the problem where the true performance of the AI systems is not correctly reflected and their capabilities are often exaggerated, which erodes confidence.

Conclusion

As the advancements in niche technology like AI and its rapid proliferation in modern warfare are expected to progress and neither strict governance or control nor a blanket ban on its utilisation is feasible, multilateral CBMs appear to be the only pragmatic solution to the responsible development and deployment of AI military systems. It is through these CBMs that effective guardrails are implemented, to ensure that this niche technology is fielded in a responsible and reliable manner, under human control. There is a requirement to increase the commitment of nations towards these CBMs and arrive at mutually acceptable code of conduct with respect to the use of this powerful technology in the best possible interest of humanity for a safer and prosperous tomorrow.

Unlike physical platform-centric weapon systems, verification of software-based artificial intelligence systems is very difficult and complex to accomplish.

Endnotes

[1] Yuval Abraham, "'Lavender': The AI Machine Directing Israel's Bombing Spree in Gaza", *+972 Magazine*, 03 Apr 2024, accessed 02 Apr 2026, https://www.972mag.com/lavender-ai-israeli-army-gaza/

[2] Harry Booth, "What the Numbers Show About AI's Harms", *TIME*, 19 Jan 2026, accessed 03 Apr 2026, https://time.com/7346091/ai-harm-risk/

[3] IEEE, "Background, Mission, and Activities of the IEEE Global Initiative', *The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems*, accessed 03 Apr 2026, https://standards.ieee.org/wp-content/uploads/import/documents/other/ec_about_us.pdf

[4] T Yigit, U Kose, and N Sengoz, "Robotics Rights and Ethics Rules", *Journal of Multidisciplinary Developments* 3, no. 1 (2018): 30–37, accessed 03 Apr 2026, http://www.jomude.com/index.php/jomude/article/view/58

[5] ICRC, "What Is International Humanitarian Law?", *Geneva: International Committee of the Red Cross (ICRC)*, accessed 03 Apr 2026, https://www.icrc.org/sites/default/files/document/filelist/what_is_ihl.pdf

[6] David J Gunkel, "The Other Question: Can and Should Robots Have Rights?", *Ethics and Information Technology* 20, no. 2 (2018): 87–99, accessed 03 Apr 2026, https://link.springer.com/article/10.1007/s10676-017-9442-4

[7] Ronald C Arkin, "Ethical Robots in Warfare", *IEEE Technology and Society Magazine* 28, no. 1 (2009): 30–33, accessed 03 Apr 2026, https://ieeexplore.ieee.org/document/4799405

[8] United States Department of Warfare, "Data, Analytics, and Artificial Intelligence Adoption Strategy", *Department of Defense*, Nov 2023, accessed 04 Apr 2026, https://media.defense.gov/2023/nov/02/2003333300/-1/-1/1/dod_data_analytics_ai_adoption_strategy.pdf

[9] United Kingdom Ministry of Defence, "Defence Artificial Intelligence Strategy", *Ministry of Defence*, June 2022, accessed 04 Apr 2026,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1082416/Defence_Artificial_Intelligence_Strategy.pdf

[10] Ioana Puscas, "Confidence-Building Measures for Artificial Intelligence: A Framing Paper", *United Nations Institute for Disarmament Research*, 2023, accessed 04 Apr 2026, https://unidir.org/publication/confidence-building-measures-for-artificial-intelligence-a-framing-paper/

[11] Ibid.

[12] Junghyun Yoon, "Key Implications of the 2nd REAIM Summit in Seoul", *Institute for National Security Strategy Issue Brief* 116, no. 13, 18 Oct 2024, accessed 03 Apr 2026, https://www.inss.re.kr/upload/bbs/BBSA05/202410/F20241018153457508.pdf

[13] Manan Suri, "Confidence-Building Measures in Responsible Artificial Intelligence", paper presented at the *Round Table Discussion on Confidence-Building Measures (CBMs) in Responsible AI in the Military Domain ()*, New Delhi.

[14] Desta Haileselassie Hagos and Danda B Rawat, "Neuro-Symbolic AI for Military Applications", *IEEE Transactions on Artificial Intelligence* 5, no. 12 (2024): 6012–6026, accessed 03 Apr 2026, https://doi.org/10.1109/TAI.2024.3444746

[15] US DoD, "Department of Defense Directive 3000.09: Autonomy in Weapon Systems", *United States Department of Defense*, 25 Jan 2023, accessed 06 Apr 2026, https://media.defense.gov/2023/Jan/25/2003149928/-1/-1/0/DOD-DIRECTIVE-3000.09-AUTONOMY-IN-WEAPON-SYSTEMS.PDF

[16] GPAI, "Algorithmic Transparency in the Public Sector: A State-of-the-Art Report of Algorithmic Transparency Instruments", *Global Partnership on Artificial Intelligence*, May 2024, accessed 07 Apr 2026, https://gpai.ai/projects/responsible-ai/algorithmic-transparency-in-the-public-sector.pdf

[17] Puscas, "Confidence-Building Measures".

[18] National Institute of Standards and Technology (NIST), "Artificial Intelligence Risk Management Framework", *US Department of Commerce*, Jan 2023, accessed 07 Apr 2026, https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

[19] Geetha Aradhyula, "Ethical and Responsible AI Frameworks", *IRE Journals* 9, no. 5, Nov 2025, accessed 03 Apr 2026, https://www.irejournals.com/paper-details/1711694

[20] Press Information Bureau, "India AI Governance Guidelines—Enabling Safe and Trusted AI Innovation", *Ministry of Electronics and Information Technology*, Nov 2025, accessed 04 Apr 2026, https://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc2025115685601.pdf

[21] Ioana Puscas, "AI and International Security: Understanding the Risks and Paving the Path for Confidence-Building Measures", *United Nations Institute for Disarmament Research (UNIDIR)*, 12 Oct 2023, accessed 08 Apr 2026, https://unidir.org/publication/ai-and-international-security-understanding-the-risks-and-paving-the-path-for-confidence-building-measures/

[22] Organisation for Economic Co-operation and Development (OECD), "AI Principles", *OECD Policy Sub-Issue*, accessed 08 Apr 2026, https://www.oecd.org/en/topics/sub-issues/ai-principles.html

About the Author

Major General Pawan Anand, AVSM, PhD (Retd) is a distinguished officer of the Indian Army's Corps of Engineers, whose career spanned nearly four decades and exemplifies excellence in leadership, operational acumen, and strategic vision. A recipient of President's Gold Medal at the National Defence Academy (NDA) and Sword of Honour at the Indian Military Academy (IMA), the officer has held diverse command and staff appointments across operational and strategic domains. He saw active combat during the Kargil War and subsequently led operations in counter-insurgency environments in Jammu and Kashmir. As Chief Engineer, South Western Command, he spearheaded the development of critical military infrastructure projects of national importance. The officer is presently serving as Director, Centre for Emerging Technologies and Atma Nirbhar Bharat at the USI of India.

Colonel Ashish Nagaich, SM, is a 1999 batch Officer, commissioned into Corps of Signals. An alumnus of NDA (95th Course) and IMA (105 Regular Course), the officer has attended the Defence Services Staff College-66 & Higher Command-51 courses and has done his MTech in Computer Science, Communications and Information Technology from Devi Ahilya Vishwavidyalya, Indore. The officer has had varied regimental experiences, which include providing active field communications in Plains, Mountain, and Deserts along Western and Northern borders and Command of an Counter Insurgency Force Signal Regiment in Northern Command. The officer holds a Master's degree in Science and Technology and is perusing his research fellowship on 'Integration & Deployment of Artificial Intelligence in Enterprise Networks'.

About the United Service Institution of India (USI

The USI was founded in 1870 by a soldier scholar, Colonel (later Major General) Sir Charles MacGregor 'For the furtherance of interest and knowledge in the Art, Science and Literature of National Security in general and Defence Services, in particular'. It commenced publishing its Journal in 1871. The USI also publishes reports of its members and research scholars as books, monographs, and occasional papers (pertaining to security matters). The present Director General is Vice Admiral Sanjay J Singh, SYSM, PVSM, AVSM, NM, PhD (Retd).

(Estd. 1870)

United Service Institution of India (USI)

Rao Tula Ram Marg, Opposite Signals Enclave, New Delhi-110057

Tele: 2086 2316/ Fax: 2086 2315, E-mail: dde@usiofindia.org

₹250