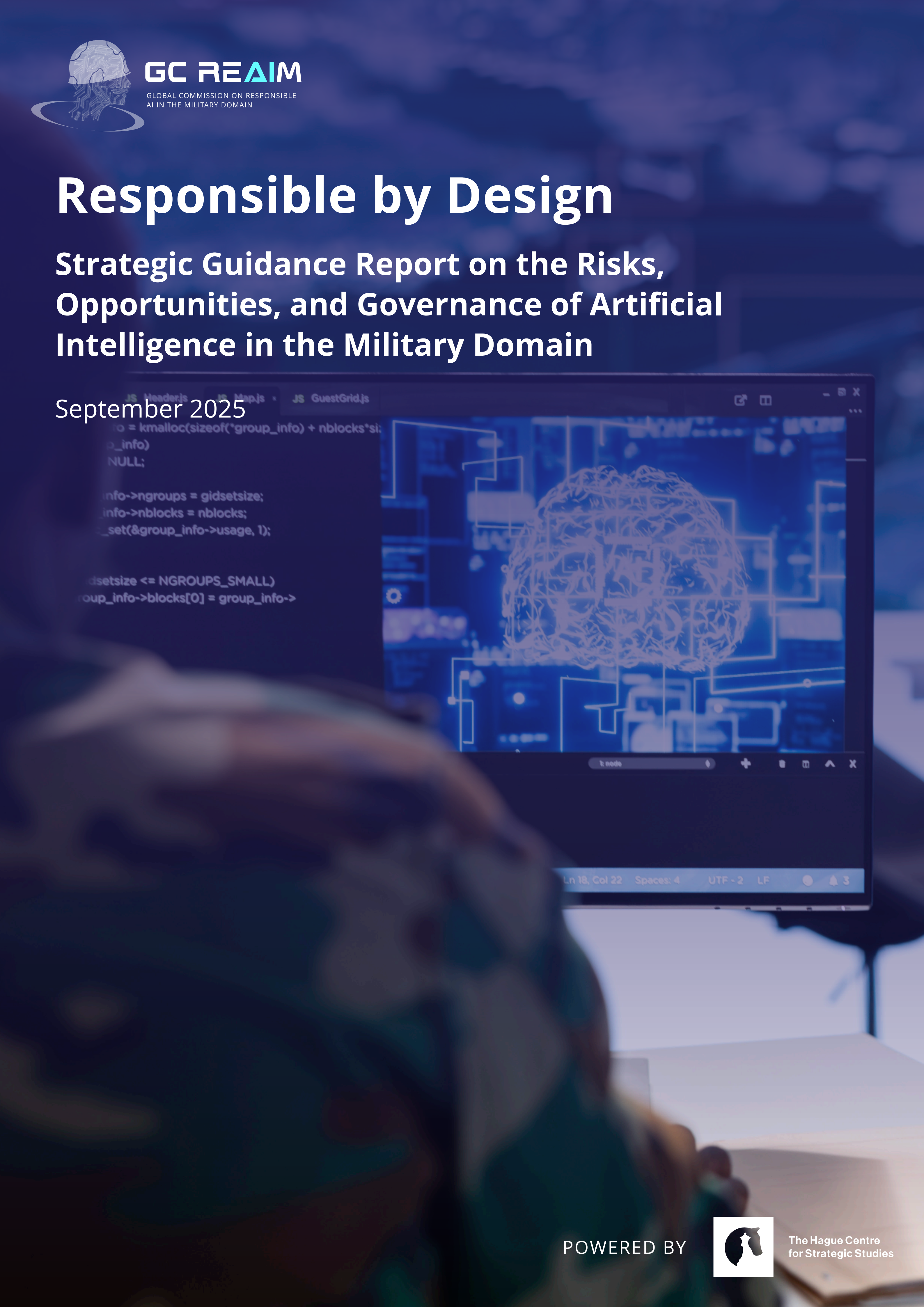


GC REAIM

GLOBAL COMMISSION ON RESPONSIBLE
AI IN THE MILITARY DOMAIN

Responsible by Design

Strategic Guidance Report on the Risks, Opportunities, and Governance of Artificial Intelligence in the Military Domain

September 2025

POWERED BY The Hague Centre for Strategic Studies

Responsible by Design

Strategic Guidance Report on the Risks, Opportunities, and Governance of Artificial Intelligence in the Military Domain

September 2025

Cover photo: iStock, Dragos Condrea

The Global Commission on Responsible Artificial Intelligence in the Military Domain (GC REAIM) is an initiative of the Ministry of Foreign Affairs of the Kingdom of the Netherlands, launched at the 2023 Responsible Artificial Intelligence in the Military Domain (REAIM) Summit. The Hague Centre for Strategic Studies (HCSS) serves as the Secretariat of GC REAIM.

Throughout 2024-2025, GC REAIM Commission Meetings, symposia, and other activities benefitted from the involvement of or were supported by The Brookings Institution, the Centre for International Governance Innovation, the Seoul International Law Academy, the Defence Artificial Intelligence Research Unit, the United Kingdom Foreign, Commonwealth and Development Office, the University of Edinburgh Law School, Trends Research and Advisory, and The Hague Centre for Strategic Studies.

CITATION:

The Global Commission on Responsible Artificial Intelligence in the Military Domain, *Responsible by Design:* Strategic Guidance Report on the Risks, Opportunities, and Governance of Artificial Intelligence in the Military Domain, The Hague, 2025.

CHAIR

Byung-se Yun

COMMISSIONERS

Patricia Adusei-Poku
Saeed Aldhaheri
Nehal Bhuta
Thompson Chengeta
Mary (Missy) Cummings
Denise Garcia
Michael Horowitz
Pak Shun Ng
Mary Ellen O'Connell
Kenneth Payne
Balaraman Ravindran
Stuart Russell
Jeroen van den Hoven
Jimena Sofia Viveros Alvarez
Toby Walsh
Bruce Watson
Yi Zeng

EXPERT ADVISORY GROUP

Zena Assaad
Yasmin Afina
Saeedeh Babaii
Ingvild Bode
Vincent Boulanin
Ariel Conn
Nicholas Davis
Jessica Dorsey
Arisa Ema
Adam Hepworth
James Johnson
Rain Liivoja
Roy Lindelauf
Marcos Lopez Oneto
Matthijs Maas
Seumas Miller
Illah R. Nourbakhsh
Barry O'Sullivan
Mun-eon Park
Giacomo Persi Paoli
Edson Prestes
Li Qiang
Michael Raska
Emma Ruttkamp-Bloem
Boris Saavedra
Nayat Sanchez-Pi
Mohammed Soliman
Catherine Tessier
Maria Vanina Martinez
Wendell Wallach
Fan Yang

PENHOLDER AND PROJECT COORDINATOR

Sofia Romansky

HEAD OF SECRETARIAT

Michel Rademaker

SCIENTIFIC ADVISOR

Tim Sweijs

The Strategic Guidance Report reflects the independent judgement and deliberations of the Chair and Commissioners of GC REAIM. While the Commissioners have drawn on the valuable insights of the GC REAIM Expert Advisory Group, whose contributions, including authored policy notes, have informed aspects of this report, the responsibility for the views, recommendations, and conclusions expressed herein rests solely with the Chair and Commissioners, and does not necessarily reflect the views of the Experts or their respective organizations. The Global Commission acknowledges with appreciation the essential role of the Expert Advisory Group in supporting rigorous, multidisciplinary dialogue throughout the process.

The Global Commission also extends particular acknowledgement to members who devoted additional time and effort to the development of the Strategic Guidance Report. Commissioners serving as Workstream Facilitators and Co-Facilitators (Nehal Bhuta, Thompson Chengeta, Denise Garcia**,** Mary Ellen O'Connell, Stuart Russell, Jeroen van den Hoven, Jimena Sofia Viveros Alvarez, and Bruce Watson), as well as members of the Drafting Committee (Thompson Chengeta, Denise Garcia, Michael Horowitz, Jeroen van den Hoven, and Jimena Sofia Viveros Alvarez) played a central role in shaping the report's structure and content. Several members of the Expert Advisory Group (Zena Assaad, Ariel Conn, and Jessica Dorsey) provided sustained input through textual reviews and technical consultations throughout the drafting process.

Visuals by: Sofia Romansky

House style by: Stephanie Govaerts

The Global Commission is grateful to Keketso Kgomosotho and Adebayo Okeowo for their contributions to the GC REAIM Expert Policy Note Series as well as their insights during the Expert Advisory Group Meeting in Edinburgh. In addition, the Global Commission thanks Perica Jovchevski for his substantive engagement with the report during the drafting process.

The Global Commission would like to thank the following individuals and organizations for submitting comments during the Red Teaming Exercise (February-March 2025) and the Silent Procedure Process (July-August 2025).

Merve Hickok (*President and Policy Director*), Julian Theseira (*Policy Group Coordinator*), Ren Bin Lee Dixon (*Policy Research Fellow*), Heramb Podar (*Policy Group Team Leader*), Idil Ilyada Kaner (*Policy Group Team Leader*) **– Center for Artificial Intelligence and Digital Policy**

Heidi Kandiel - **International Committee of the Red Cross**

Brooke Latham **– NATO Innovation Fund**

Netta Goussac **– Stockholm International Peace Research Institute**

Rosanna Fanni

Heidy Khlaaf

The Ministry of Defence of the Republic of Kenya

The Foreign, Commonwealth and Development Office and Ministry of Defence of The United Kingdom of Great Britain and Northern Ireland

The Ministry of Defence of the Kingdom of the Netherlands

The Ministry of Foreign Affairs of the Kingdom of the Netherlands

The Ministry of Defence of the Kingdom of Spain

The Ministry of Foreign Affairs, European Union and Cooperation of the Kingdom of Spain

The Ministry of Foreign Affairs of the Republic of Korea

The Ministry of National Defense of the Republic of Korea

The Ministry of Defence, Singapore

Table of Contents

Foreword

of the Ministers of Foreign Affairs and Defence of the Kingdom of the Netherlands

The introduction of artificial intelligence (AI) in the military domain raises pressing questions. The very strength of AI — its ability to recognise and exploit patterns in vast amounts of data — can also become a vulnerability in a military setting. This may be the case, for example, when there is a lack of alignment between training data and the situation on the ground, or when malicious actors successfully manipulate algorithmic vulnerabilities.

These kinds of dilemmas create a shared interest in developing AI systems that remain reliable, robust and resilient. We therefore need a clear and internationally recognised approach to developing and deploying AI applications.

A few years ago, the debate on AI in the military domain was still in its early stages. The first REAIM Summit, held two years ago in 2023, made a decisive contribution in helping to set the stage for international dialogue on this topic. Since then, we have taken concrete steps to turn awareness into practice.

The present report marks an important milestone in that regard. It offers states, as well as industry, the guidance they need to make informed decisions and address potential risks. What's more, it helps them to chart a course that allows them to reap the benefits of AI while mitigating its inherent challenges.

This report is the result of contributions from many people, bringing together diverse expertise and regional perspectives. That is precisely its strength. We need the expertise of military practitioners, but also those working in the realms of international law, disarmament and regional cooperation. This report draws on the expertise of a wide range of professionals, from philosophers and ethicists to computer scientists, military personnel and lawyers.

We would like to thank everyone who has contributed, and especially the GC REAIM Chair, His Excellency Byung-se Yun, for uniting such a diversity of perspectives. We are also grateful to the Secretariat for skilfully facilitating the research and drafting process. This was no mean feat, given the strong knowledge base around the table.

AI innovation and adoption is rapidly transforming the character of warfare. We must have the courage and wisdom to address AI's many dimensions and dilemmas – and we must do so together. This effort must include states from all regions, along with industry, knowledge institutions and civil society. I hope this report marks an important step in bringing all of us together to address this complex subject and sparks action in this area, with each of us playing our own unique role.

David van Weel, Minister of Foreign Affairs of the Kingdom of the Netherlands
Ruben Brekelmans, Minister of Defence of the Kingdom of the Netherlands

Foreword
of the Chair of GC REAIM

“It always seems impossible, until it is done.” Nelson Mandela’s words echoed when I was asked to chair the Global Commission on Responsible AI in the Military Domain (REAIM) in early 2024. The charge - offering strategic guidance on a defining issue at the nexus of fast-moving technology, geopolitics, and human dignity - felt like a “Mission Impossible.”

AI is a civilizational technology reshaping every domain of human activity like an unbound Prometheus. Its breakneck pace risks outstripping our ability to govern it. We are already in the era of generative AI, while races toward Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI) are well underway. AI is both dual-use and an emerging disruptive technology. AI transformation (AX) across sectors affects the military domain as well.

As nuclear technology shaped the 20th century, AI is poised to be the 21st century’s game changer. Its deployment - from targeting systems and autonomous platforms to non-weapon applications - is altering the nature of warfare. Recent conflicts in Europe and the Middle East have illustrated AI’s growing presence in warfare, raising questions about accountability and human control in the current and future conflicts. This underscores the need for global governance—not dominance by any single nation or bloc.

We must be military AI-ready. The direction we take—toward constructive progress or destabilizing chaos—will be determined by choices we make today and in the years ahead. This report is more than a diagnosis or forecast. It is a prescription for responsible action.

I stand humbled by the magnitude of this task and do not declare “Mission Accomplished.” It still feels, to me, like Schubert’s “Unfinished Symphony.” Yet I submit that this joint work marks a meaningful milestone toward a governance framework capable of addressing today’s deficit in military AI governance.

Over the past sixteen months, the Commission engaged in intensive, wide-ranging deliberations, drawing on diverse expertise and perspectives. We were mindful of the UN High-Level Advisory Body’s 2024 report, “Governing AI for Humanity,” which noted: “There is no shortage of documents and dialogues on AI governance. Yet, none are truly global or comprehensive.” Our goal was to help fill that gap.

GC REAIM was designed as a global, interdisciplinary, and inclusive platform. Commissioners and experts from AI research, military strategy, ethics, law, and global governance shaped a report that is both insightful and practical.

As such, the recommendations of this report complement UN General Assembly Resolutions on AI (Res/79/239 and Res/79/325), and seek to translate the REAIM Summit declarations into actionable guidance.

Eight features distinguish this report:

1. **Holistic, Comprehensive Scope:** We examine military AI through technical, strategic, legal, ethical, and institutional lenses. The report addresses responsibility, human judgment, and AI's intersection with international peace and security.
2. **Layered Approach with Operational Focus:** We present three guiding principles, five core recommendations, and lifecycle-based guidance —across strategic, tactical, and operational levels.
3. **Responsibility by Design:** Ethical and legal compliance must be embedded from the earliest stages of AI development, supporting context-driven human decision-making.
4. **Conceptual Clarity:** We unpack key ideas from diverse REAIM-related initiatives and resolutions — applicable norms, human oversight, stakeholder roles, governance models—and organize them into a coherent framework bridging normative aspirations and practical implementation.
5. **Landscape Mapping:** Our matrix of military AI use cases goes beyond autonomous weapons, offering a taxonomy of applications and clarifying operational priorities in the real situations.
6. **Future-Proofing:** Our recommendations are designed to remain relevant as AI evolves—from generative models to AGI—balancing specificity with adaptability and anticipating shifts in capability, doctrine, and threat perception.
7. **Balanced Perspective:** We weigh AI's opportunities and risks, recognizing that transition to the era of AGI further affects geopolitical stability. Our approach avoids alarmism and complacency, aiming for informed vigilance.
8. **Human-Centric Governance:** We offer a roadmap for global governance and propose safeguards, including redlines on AI involvement in nuclear decision-making, to prevent an "AI's Oppenheimer Moment." Our goal is to ensure AI serves humanity—not the reverse.

This report carries a sober message: Humanity must learn to govern AI before AI governs humanity. As technology increasingly drives geopolitics, we must build a cooperative international order that prioritizes norm-based innovation and shared security.

Our layered governance framework sets legal, ethical, operational, spatial, and temporal parameters for responsible military-AI integration. It aims to foster norm development, build trust, and support policy coherence in a fragmented world. We emphasize inclusive dialogue, transparency, and accountability mechanisms that can adapt as technology changes.

Ultimately, this report is both a roadmap and a framework of actions. It invites all stakeholders to shape the future of military AI—not through unilateral ambition, but through multilateral cooperation. The stakes are high, and the window for responsible leadership is narrow.

I express deepest gratitude to the distinguished Commissioners, contributing experts, the dedicated Secretariat, and the governments of the Netherlands and the Republic of Korea. Their unwavering support made this journey possible. Together, we have taken a meaningful step toward ensuring that military AI evolves in service of peace, stability, and human dignity.

Byung-se Yun, Chair of GC REAIM

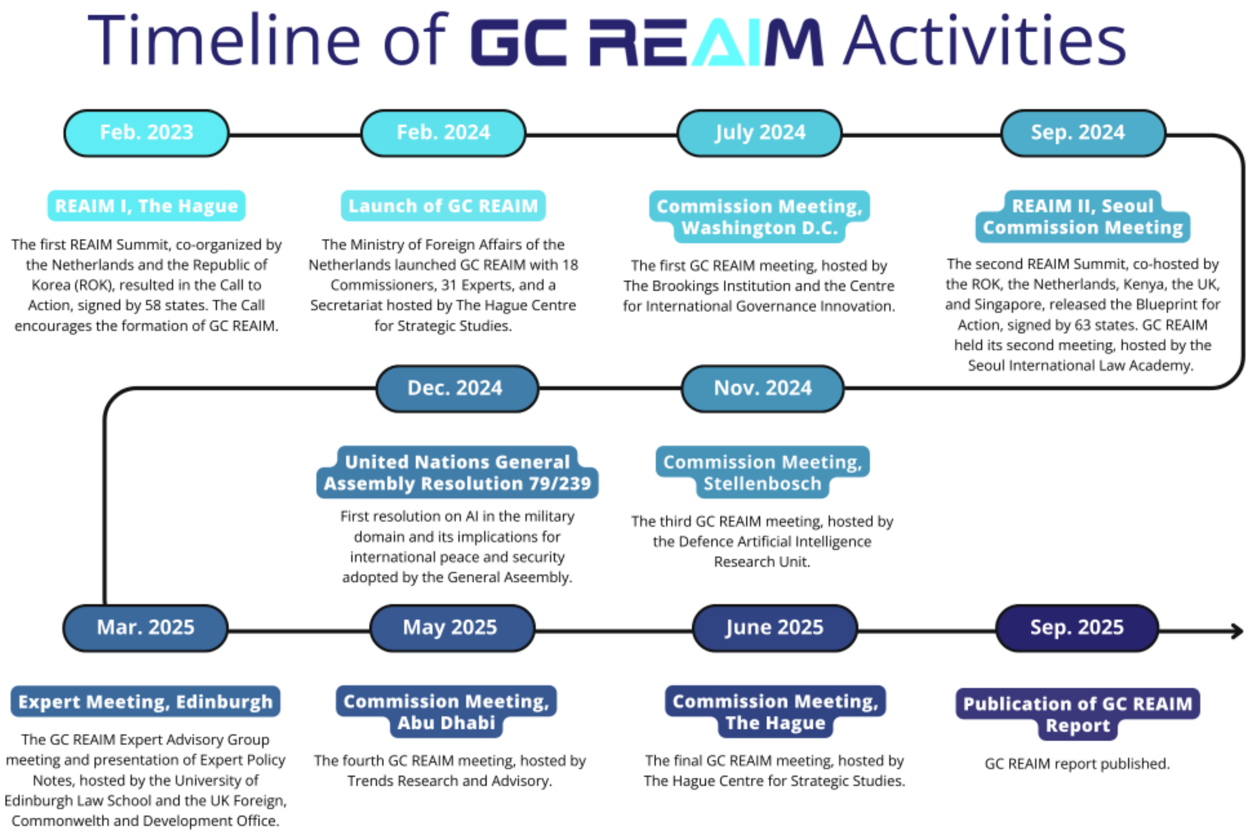
Abbreviations

AAR	After Action Review
AGI	Artificial General Intelligence
AI	Artificial Intelligence
AIA	Algorithmic Impact Assessment
AIDA	Architecture for Interoperable Digital Technologies within the Alliance
AP I	Additional Protocol 1
C3	Command, Control, and Communications
CBM	Confidence Building Measure
CCIS	Command and Control Information System
CCW	Convention on Certain Conventional Weapons
COR	Chain of Responsibility
DSS	Decision Support Systems
DSTA	Defense Science and Technology Agency
GC REAIM	Global Commission on Responsible Artificial Intelligence in the Military Domain
GGE	Group of Governmental Experts
ICL	International Criminal Law
ICRC	International Committee of the Red Cross
IEC	International Electrotechnical Commission
IEEE-SA	Institute of Electrical and Electronics Engineers Standards Association
IHL *or* jus in bello	International Humanitarian Law
IHRL	International Human Rights Law
ISO	International Organization of Standardization
ISR	Intelligence, Surveillance, and Reconnaissance
LAWS	Lethal Autonomous Weapons Systems
ML	Machine Learning
MODD	Moral Operational Design Domain
NATO	North Atlantic Treaty Organization
NC3	Nuclear Command, Control, and Communications
NIST	National Institute of Standards and Technology
OECD	Organisation for Economic Co-operation and Development
PM-ADM	Political Military-Assisted Decision Making
REAIM	Responsible Artificial Intelligence in the Military Domain
ROE	Rules of Engagement
RPA	Robotic Process Automation
RTD&E	Research, Development, Testing, and Evaluation
SAF	Singapore Armed Forces
SALT 1	Strategic Arms Limitations Talks/Treaty
SIPRI	Stockholm International Peace Research Institute
SOP	Standard Operating Procedures
TDMP	Team Design Moral Patterns
TEVV	Testing, Evaluation, Validation, and Verification
TTP	Tactics, Techniques, and Procedures
UGV	Unmanned Ground Vehicles
UK MoD	UK Ministry of Defence
UN	United Nations
UNGA	United Nations General Assembly
UNIDIR	United Nations Institute for Disarmament Research
UNSG	UN Secretary General
US DoD	US Department of Defense
WMD	Weapons of Mass Destruction

The Global Commission on Responsible Artificial Intelligence in the Military Domain

GC REAIM is an initiative of the Ministry of Foreign Affairs of the Netherlands that was launched during the 2023 REAIM Summit in The Hague. GC REAIM was established to help promote mutual awareness and understanding among the many communities working on issues related to the governance of AI in the military domain. By linking discourses between various stakeholders, GC REAIM fosters synergies, advancing policy coherence and norm development in this field.

GC REAIM merges the diverse networks of knowledge, experiences, and perspectives of its Commissioners and Experts with a mandate to produce outputs that reflect global, multi-stakeholder debates on governance and offer concrete recommendations and paths forward to states, militaries, and industry. Since July 2024, GC REAIM has held meetings in five global regions: North America, East Asia, Southern Africa, the Middle East, and Europe. Through its in-person Commission Meetings and symposia, GC REAIM brought together international and local expert communities, gathered inputs for the report, and helped progress deliberations.



Timeline of GC REAIM Activities

Feb. 2023

REAIM I, The Hague

The first REAIM Summit, co-organized by the Netherlands and the Republic of Korea (ROK), resulted in the Call to Action, signed by 58 states. The Call encourages the formation of GC REAIM.

Feb. 2024

Launch of GC REAIM

The Ministry of Foreign Affairs of the Netherlands launched GC REAIM with 18 Commissioners, 31 Experts, and a Secretariat hosted by The Hague Centre for Strategic Studies.

July 2024

Commission Meeting, Washington D.C.

The first GC REAIM meeting, hosted by The Brookings Institution and the Centre for International Governance Innovation.

Sep. 2024

REAIM II, Seoul Commission Meeting

The second REAIM Summit, co-hosted by the ROK, the Netherlands, Kenya, the UK, and Singapore, released the Blueprint for Action, signed by 63 states. GC REAIM held its second meeting, hosted by the Seoul International Law Academy.

Nov. 2024

Commission Meeting, Stellenbosch

The third GC REAIM meeting, hosted by the Defence Artificial Intelligence Research Unit.

Dec. 2024

United Nations General Assembly Resolution 79/239

First resolution on AI in the military domain and its implications for international peace and security adopted by the General Assembly.

Mar. 2025

Expert Meeting, Edinburgh

The GC REAIM Expert Advisory Group meeting and presentation of Expert Policy Notes, hosted by the University of Edinburgh Law School and the UK Foreign, Commonwealth and Development Office.

May 2025

Commission Meeting, Abu Dhabi

The fourth GC REAIM meeting, hosted by Trends Research and Advisory.

June 2025

Commission Meeting, The Hague

The final GC REAIM meeting, hosted by The Hague Centre for Strategic Studies.

Sep. 2025

Publication of GC REAIM Report

GC REAIM report published.

Executive Summary

Artificial intelligence is a transformative and civilization-shaping technology. Its integration into the military domain carries profound implications for how armed conflict is conducted, including how decisions are made and responsibilities are managed in one of the most consequential areas of human activity. The stakes extend beyond the battlefield, introducing heightened risks in the context of international peace and security. Given the cross-border nature of technological development and deployment, the challenges posed by AI in the military domain demand coordinated, multi-stakeholder governance. This report responds to the urgency of this moment by offering a foundational framework to support informed dialogue and engagement among the diverse stakeholders concerned with the responsible governance of AI in the military domain.

This report recognizes that AI as a category of technologies represents a double-edged sword, with rapid innovation, driven by immense private sector investment and military demand alike, often coming at the cost of safe and responsible practices. AI, especially advances in machine learning, raises a host of ethical, legal, and operational concerns. Depending on the specific model and context of use, AI can challenge conventional notions of human agency, responsibility, and accountability, creating uncertainty around how existing principles and rules apply. As such, the potential opportunities created by the integration of AI into the military domain can only be realized if associated risks are effectively addressed.

The responsible development and use of AI in the military domain requires informed human decisions and context-specific approaches to designing, testing, and deploying these technologies in ways that uphold values of peace and human dignity. Therefore, GC REAIM advocates for **responsibility by design**, where ethical and legal compliance is integrated from the earliest stage of development, through the entire AI system lifecycle, as well as in the socio-technical institutions, like militaries, where AI applications are embedded. Militaries and other stakeholders should adopt responsibility by design principles and practices. In furthering this objective, GC REAIM seeks to foster collaboration among states by supporting capacity and capability development. The recommendations proposed by GC REAIM are intended to be adaptable, allowing all states to implement them, regardless of their individual starting points and priorities.

Reflecting GC REAIM's commitment to ethical, lawful and human-centric governance, this report puts forward **three guiding principles** for the responsible development and use of AI in the military domain:

1. The development and use of AI in the military domain must comply with international law, guided by widely shared ethical principles, to best ensure the preservation of human life and peace for present and future generations.

2. The development and use of AI in the military domain must follow systematic and structured design, development, and testing processes across the entire system lifecycle. These processes must be explicitly oriented toward safeguarding human agency, responsibility, and accountability. In particular, AI systems must be engineered so that ultimate responsibility for all critical decisions remains with human operators.

3. Individuals involved in the development and use of AI in the military domain must be supported through institutional practices that enable them to exercise informed human agency. This requires continuous training, capacity-building, and the integration of design and testing features that strengthen human understanding and effective oversight.

These guiding principles are operationalized through **five core recommendations** for public and private stakeholders involved throughout the AI lifecycle:

1. Anchor the responsible development and use of AI in the military domain in relevant and applicable ethical principles and international law.

The responsible development and use of AI in the military domain depends on the early and sustained integration of ethical and legal considerations across the system lifecycle. This entails detailed translations of ethical and legal principles to delineate unlawful uses of AI, the development of technical, operational, and organizational guidelines, as well as respect for and application of existing obligations, such as the right to use force (*jus ad bellum*) and how force is employed during conflict (*jus in bello*).

2. Agree, at a legally binding level, that the decision to authorize the use of nuclear weapons should remain under human control.

Consistent with the policies and positions of several nuclear powers, critical decisions about the use of nuclear weapons must unequivocally remain under human authority as they require moral, legal, and strategic considerations. These commitments should be pursued through either national policy declarations or an international agreement.

3. Implement national policies that guarantee human responsibility across the AI system lifecycle and that are demonstrably grounded in human-centric training and rigorous testing, evaluation, validation, and verification.

From the conceptualization of a military capability through its retirement, states should implement their own policies, identify key intervention points, and create positive incentives to encourage responsibility by design in companies and organizations. This process should be in line with existing due diligence obligations. This includes establishing mechanisms for ethical and legal evaluation, training and education, compliance monitoring, testing, evaluation, validation, and verification, risk assessments, and using internationally agreed on best practices for implementation at the national level.

4. Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain.

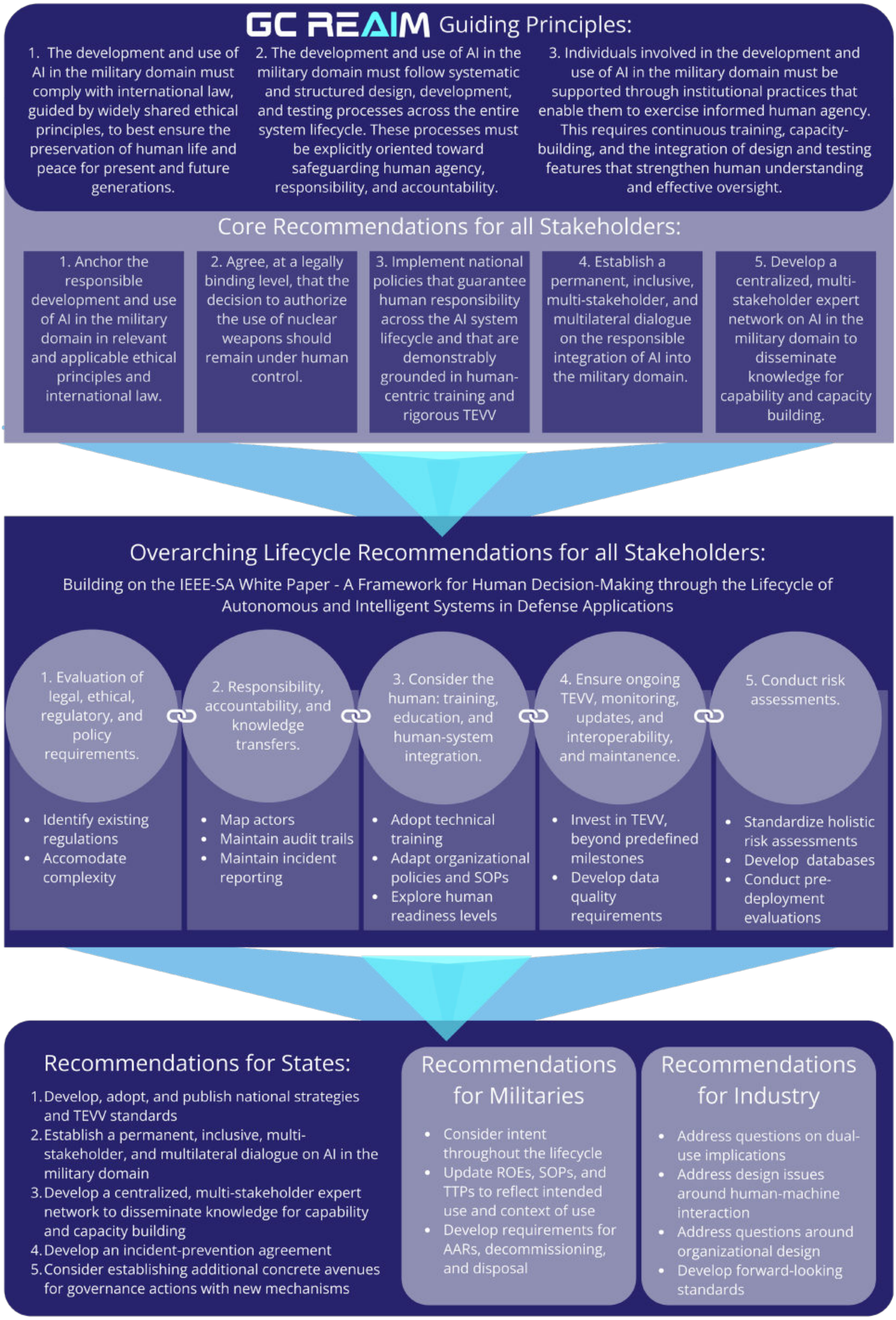
A global dialogue on responsible AI in the military domain should involve states (including militaries), industry, and civil society. A yearly REAIM meeting can serve as the building block for a dialogue that evolves into having a seat at an existing secretariat, or within the United Nations System, as appropriate.

5. Develop a centralized, multi-stakeholder expert network on AI in the military domain to disseminate knowledge for capability and capacity building.

This expert network should serve as a specialized external resource, centrally coordinated by a hub, that supports states and key stakeholders in proactively identifying areas of convergence across international, regional, and national AI policies. Functioning in close coordination with relevant bodies, including the UN's independent International Scientific Panel on AI and Global Dialogue on AI Governance, as well as the abovementioned dialogue, this network could contribute to and maintain trusted repositories, conduct horizon scanning, and support confidence-building measures, helping ensure that lessons learned from high-stakes military AI applications inform AI governance frameworks.

GC REAIM recognizes that normative questions about human agency cannot be answered in the abstract. Norms and principles must be translated into actionable guidelines and standards. Addressing these considerations, the above **five core recommendations** form the foundation of a set of tailored recommendations for responsible practices on all levels of the socio-technical AI system lifecycle, including organizational design considerations, as well as specific guidance for states, militaries and industry. The recommendations build upon and align with notable diplomatic developments and draw upon already forming institutions at the United Nations, including the International Scientific Panel on AI and the Global Dialogue on AI Governance. They provide concrete steps for stakeholders to implement the above suggestions at the national, regional, and international levels. With this report, GC REAIM proposes modalities for global cooperation to foster warranted trust in the responsible use of AI systems, paving the way towards the **institutionalization** of governance for AI in the military domain.

Summary of GC REAIM Recommendations



GC REAIM Guiding Principles:

1. The development and use of AI in the military domain must comply with international law, guided by widely shared ethical principles, to best ensure the preservation of human life and peace for present and future generations.

2. The development and use of AI in the military domain must follow systematic and structured design, development, and testing processes across the entire system lifecycle. These processes must be explicitly oriented toward safeguarding human agency, responsibility, and accountability.

3. Individuals involved in the development and use of AI in the military domain must be supported through institutional practices that enable them to exercise informed human agency. This requires continuous training, capacity-building, and the integration of design and testing features that strengthen human understanding and effective oversight.

Core Recommendations for all Stakeholders:

1. Anchor the responsible development and use of AI in the military domain in relevant and applicable ethical principles and international law.

2. Agree, at a legally binding level, that the decision to authorize the use of nuclear weapons should remain under human control.

3. Implement national policies that guarantee human responsibility across the AI system lifecycle and that are demonstrably grounded in human-centric training and rigorous TEVV

4. Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain.

5. Develop a centralized, multi-stakeholder expert network on AI in the military domain to disseminate knowledge for capability and capacity building.

Overarching Lifecycle Recommendations for all Stakeholders:

Building on the IEEE-SA White Paper - A Framework for Human Decision-Making through the Lifecycle of Autonomous and Intelligent Systems in Defense Applications

1. Evaluation of legal, ethical, regulatory, and policy requirements.
- Identify existing regulations
- Accomodate complexity

2. Responsibility, accountability, and knowledge transfers.
- Map actors
- Maintain audit trails
- Maintain incident reporting

3. Consider the human: training, education, and human-system integration.
- Adopt technical training
- Adapt organizational policies and SOPs
- Explore human readiness levels

4. Ensure ongoing TEVV, monitoring, updates, and interoperability, and maintanence.
- Invest in TEVV, beyond predefined milestones
- Develop data quality requirements

5. Conduct risk assessments.
- Standardize holistic risk assessments
- Develop databases
- Conduct pre-deployment evaluations

Recommendations for States:

1. Develop, adopt, and publish national strategies and TEVV standards
2. Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on AI in the military domain
3. Develop a centralized, multi-stakeholder expert network to disseminate knowledge for capability and capacity building
4. Develop an incident-prevention agreement
5. Consider establishing additional concrete avenues for governance actions with new mechanisms

Recommendations for Militaries

- Consider intent throughout the lifecycle
- Update ROEs, SOPs, and TTPs to reflect intended use and context of use
- Develop requirements for AARs, decommissioning, and disposal

Recommendations for Industry

- Address questions on dual-use implications
- Address design issues around human-machine interaction
- Address questions around organizational design
- Develop forward-looking standards

1 Introduction

Artificial intelligence (AI) enables a wide range of capabilities which can enhance the speed, scale, and efficiency of military activities. Yet, the potential opportunities created by AI can only be realized if militaries address the associated risks. The accelerated integration of AI, particularly machine learning (ML), into the military domain is introducing new and intensified challenges in the contexts of peace and security, the conduct of armed conflict, as well as beyond. This demands careful and purposeful governance to ensure that technological innovation is accompanied by the development of ethical, legal, and operational safeguards. The use of AI must reinforce, not undermine, established norms and responsibilities borne by present and future generations.[1]

Meeting this imperative requires answering fundamental questions about human agency, responsibility, and accountability in the development and use of AI. Key concerns include the reliability of AI outputs, the adequacy of user understanding, and, consequently, the extent to which decisions informed by AI can be ethically and legally justified.[2] These uncertainties are only amplified when AI is used in activities which are as complex, unpredictable, and consequential as warfare.

As public and private sector investment continues to drive significant advances in AI, reshaping almost all facets of society, decisions regarding the development and use of AI in military contexts should, where feasible, align with civilian developments and tested solutions. In itself, AI is an inherently dual-use technology which blurs the lines between the civilian and military domains. Yet, while such alignment is a necessary starting point, it is not sufficient on its own for governance.

International governance initiatives have often focused either predominantly on civilian applications of AI or more narrowly on autonomous weapons systems, thereby inadvertently neglecting the technology's broader strategic implications within the military domain. In light of this, the United Nations Secretary-General (UNSG) has issued an urgent call to address the emerging security challenges at the nexus of AI and security. In this context, GC REAIM addresses a critical gap.

In 2023, AI in the military domain took center stage on the global diplomatic agenda with the inaugural Responsible AI in the Military Domain (REAIM) Summit, convened in The Hague. The Summit became a pivotal platform for policymakers, practitioners, and experts to establish a shared set of principles to guide the military development and use of AI. The Summit culminated in a Call to Action, a document endorsed by 58 states scoping the risks and opportunities associated with AI in the military domain.[3]

[1] Emma Ruttkamp-Bloem, "Intergenerational Justice as Driver for Responsible AI," in *Artificial Intelligence Research*, ed. Anban Pillay et al. (Springer Nature Switzerland, 2023), https://doi.org/10.1007/978-3-031-49002-6_2.

[2] Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy brief no. 193 (Centre for International Governance Innovation, 2025), https://www.cigionline.org/static/documents/PB_no.193.pdf.

[3] Global Commission on Responsible Artificial Intelligence in the Military Domain, "REAIM 2023 Call to Action," Government of the Netherlands, February 16, 2023, https://www.government.nl/documents/publications/2023/02/16/reaim-2023-call-to-action.

Recognizing the need for multi-stakeholder dialogue, knowledge, and capacity building, the Ministry of Foreign Affairs of the Netherlands launched GC REAIM with a mandate to foster mutual awareness and understanding among diverse communities. By connecting and amplifying dialogues, the Global Commission advances the development of norms and policy coherence around AI in the military domain worldwide.

The second REAIM Summit, held in 2024 in Seoul, resulted in the Blueprint for Action, a document endorsed by 63 states, outlining steps to implement responsible practices and shaping avenues for the future of global governance.[4] GC REAIM Commissioners and Experts actively participated in the Summit, exchanging knowledge that informed this report.

Building on the momentum of the two REAIM Summits, the REAIM Regional Consultations,[5] GC REAIM Commission Meetings held throughout 2024 and 2025, the continued work of the Convention on Certain Conventional Weapons' Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems (CCW GGE LAWS), and the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, among others,[6] this report serves as the key deliverable of GC REAIM to date.

It establishes a foundational framework to support informed dialogue and engagement among the diverse stakeholders concerned with the governance of AI in the military domain. At its core is the concept of **responsibility by design**: the idea that ethical and legal compliance must be integrated from the earliest stage of development, through the entire AI system lifecycle as well as in the socio-technical institutions in which AI applications are embedded.[7] As such, the responsible integration of AI in the military domain entails approaches that respect the distinctions and accommodate the core principles and recommendations presented in this report. Drawing on the Global Commission's diverse expertise and perspectives, the report:

- Provides **a shared knowledge base** for policymakers, practitioners, and experts from various backgrounds, forming common ground for deliberations;
- Offers targeted **recommendations** for states, militaries, and industry, for the responsible development and use of AI throughout the system lifecycle;
- Proposes an approach for the **institutionalization** of global governance.

The structure of this report reflects the intricacies of the issue area. Each part addresses a critical dimension, contributing to a holistic analysis and forming the basis for coherent and actionable solutions:

- **Section 2** of the report lays out the inherent complexities and dynamics found at the nexus of AI and the military domain;

[4] "REAIM Blueprint for Action," paper presented at Responsible AI in the military domain Summit, Seoul, Republic of Korea, September 9, 2024, https://reaim2024.kr/home/reaimeng/board/bbsDetail.do?encMenuId=4e57325766362f626e5179454e6d6e4d4a4d33507a773d3d&encBbsMngNo=366e794c7a644d756342425668444f393053755142673d3d&encBbsNo=6f784e4542386f7735767465766a6531556f4b6149413d3d&ctlPageNow=1&schKind=bbsTtlCn&schWord=#this.

[5] Yasmin Afina, *The Global Kaleidoscope of Military AI Governance* (United Nations Institute for Disarmament Research, 2024), https://unidir.org/publication/the-global-kaleidoscope-of-military-ai-governance/.

[6] GC REAIM members further engaged with the UN Human Rights Council's Study on the human rights implications of AI in the military domain, The UN High-Level Advisory Body on AI, The Summit of the Future, The UK-led AI Safety Summit, the Caribbean Community Declaration on Autonomous Weapon Systems, and the Economic Community of West African States Communique on AWS among others as well as attended the 2025 Paris AI Action Summit and the United Nations Institute for Disarmament Research Global Conference on AI, Security, and Ethics.

[7] Yasmin Afina, "Combat Code Compliance: International Humanitarian Law and the Development Stages of AI for Targeting" (University of Essex, 2025), https://doi.org/10.5526/ERR-00041309."; Article 36 Legal, "Lawful by Design Initiative – Bridging the Gap between Industry Led Innovation and National Legal Review Processes."; Joanna Bryson and Alan Winfield, "Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems," Computer 50, no. 5 (2017): 116–19, https://doi.org/10.1109/MC.2017.154.

- **Section 3** outlines the risks and opportunities associated with the integration of AI into the military domain within the contexts of peace and security, the conduct of armed conflict, as well as beyond armed conflict;
- **Section 4** presents guiding principles and core recommendations, which are elaborated in the form of general guidance throughout the AI system lifecycle as well as specific guidance for states, militaries, and industry;
- **Section 5** provides policy and diplomatic updates, a general conclusion on GC REAIM’s findings, and points to next steps.

2 Military and Technological Foundations

A comprehensive understanding of AI and its diverse applications is essential to ensure responsible integration into the military domain. This chapter outlines the scope of GC REAIM's analysis and lays the ground for further discussion in relation to:

1. The military context;
2. The technical characteristics of AI;
3. The multi-stakeholder character of the AI system lifecycle;
4. The uses of AI in the military domain and priorities for governance.

This detailed account is of paramount importance for the responsibility by design approach adopted by GC REAIM. This section helps identify the points where relevant actors have opportunities and responsibilities to develop and use systems with key ethical and legal concerns in mind.

2.1 The Military Context

An understanding of the military domain's structural and operational complexity is vital not only to grasp the unique needs of and challenges faced by military practitioners, but also to develop realistic and relevant recommendations around uses of AI.[8]

War is an inherently human and violent endeavor. Even if bound by a vast body of rules and regulations, war remains the continuation of policy by other means. Both in and outside war, the military is a pivotal instrument alongside diplomatic, economic, and other tools of statecraft. Militaries plan and operate under conditions of uncertainty. Uncertainty stems both from the 'fog of war', as well as from frictions and affordances arising between human, environmental, or technological factors.[9] This uncertainty is further compounded by the lack of transparency which characterizes much of military technological development and operational practice. The classified nature of many systems and procedures makes it more difficult, but also of greater value, to assess and verify the capabilities of adversarial technologies.

[8] Edson Prestes et al., *Effective Governance Through Common Understanding*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/04/Prestes-et-al.-final.pdf.

[9] The fog of war describes instances where militaries have to act on incomplete information about their capabilities, adversaries, and the environment in rapidly changing and unpredictable conditions. Carl von Clausewitz, *On War*, ed. F.N. Maude, trans. J.J. Graham, (Wordsworth Editions, 1997), https://wordsworth-editions.com/book/on-war/.

Military activities take place across strategic, operational, and tactical levels, guided by decision-making from strategic command to field units and individual operators. Military engagements span a spectrum of conflict, from peacetime to grey zone activities and high-intensity warfare. Moreover these engagements also occur through kinetic and non-kinetic means across the land, maritime, air, space, and cyber realms, each with distinct operational requirements and constraints.[10] The types of engagements, depending on the circumstances, include but are not limited to:

1. Defensive and offensive operations in armed conflict;
2. The use of force domestically and abroad in counter-piracy, terrorism, insurgency, or organized crime operations, including those in support of national security such as for law enforcement, border security, and protection of critical infrastructure;
3. Engagements requiring military assistance, such as peace operations, humanitarian, and disaster relief, including civilian evacuation and rescue.[11]

The military operational environment is inherently adversarial. Intelligent and adaptive opponents seek to exploit vulnerabilities in every type of engagement. In response, tactical decision-making must occur at a high tempo, level of precision, and adaptability. Yet, these tactical decisions are necessarily situated within a wider decision-making cycle, where operational and strategic decisions unfold over longer timeframes.[12] This cycle involves complex deliberation, coordination, and political oversight in the context of formal ethical and legal frameworks, such as international humanitarian law (IHL), as well as nation- or coalition-specific rules of engagement (ROE) and standard operating procedures (SOP).[13] Commanders and operators follow ROEs and SOPs, in consultation with legal and other advisors, to establish accountability when deviations from established requirements occur.

Modern military engagements, which encounter layered and networked operational environments, face amplified challenges. AI has been employed to varying degrees across different military functions and contexts since its emergence. However, the growing complexity of contemporary battlefields, combined with significant advances in AI capabilities and their application, has substantially increased the technology's relevance, utility, and perceived strategic value in the military domain.

[10] Herwin Meerveld and Roy Lindelauf, *Context Is Everything: Policy Implications of the Military AI Responsibility Contextualization Framework*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Lindelauf-Meerveld-Abu-Dhabi-Housestyle.pdf.

[11] Giacomo Persi Paoli et al., *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security – An Evidence-Based Road Map for Future Policy Action* (UNIDIR, 2025), https://unidir.org/wp-content/uploads/2025/07/UNIDIR_AI_military_domain_implications_international_peace_security.pdf.

[12] For the seminal concept of the Observe–Orient–Decide–Act (OODA) loop, see Robert Coram, Boyd: The Fighter Pilot Who Changed the Art of War (Hachette+ORM, 2002).

[13] Joseph Caldwell Wylie, *Military Strategy: A General Theory of Power Control*, with Internet Archive (Greenwood Press Publishers, 1980), http://archive.org/details/militarystrategy0000wyli.

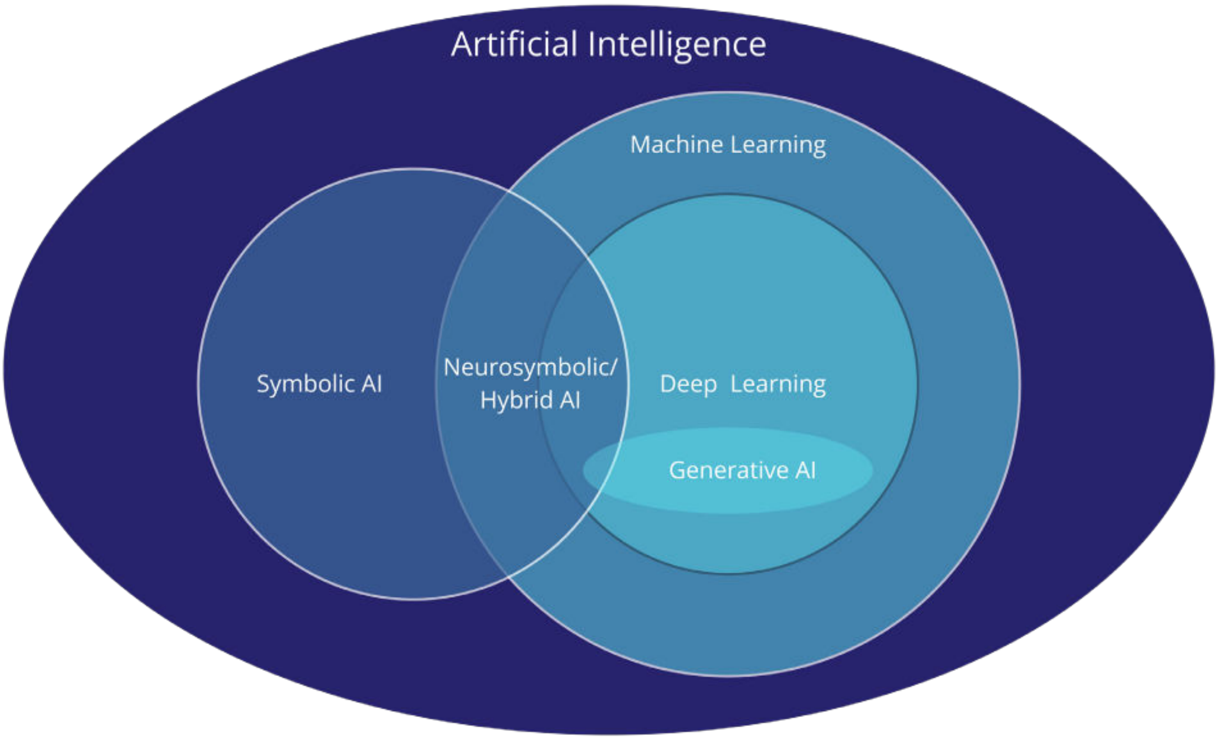
2.2 The Technical Characteristics of AI

AI as a category of technologies is marked by complexities: diverse techniques, challenges stemming from the nature of algorithms, and misconceptions about capabilities. Therefore, technical understanding is key for informed analysis and recommendations.

According to the Organisation for Economic Co-operation and Development (OECD), AI can be broadly understood as *"a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."*[14] Additionally, as elaborated by the European Commission, *"AI-based systems can be purely software-based, acting in the virtual world or AI can be embedded in hardware devices."*[15]

AI is a rapidly evolving field, propelled by large-scale public and private investment worldwide, which overlaps with areas of operations research, game theory, mathematical optimization theory, and data science. The range of contexts in which AI can be applied arises from the diversity of AI techniques. Key techniques are highlighted in Graphic 1, a simplified high-level illustration, and Table 1, a definitional overview.

***Graphic 1.* Field of Artificial Intelligence**[16]



[14] Organisation for Economic Co-operation and Development AI Policy observatory, "AI Principles Overview," OECD Global Partnership on Artificial Intelligence, https://oecd.ai/en/principles.

[15] High-Level Expert Group on Artificial Intelligence, "High-Level Expert Group on Artificial Intelligence – A Definition of AI: Main Capabilities and Scientific Disciplines," European Commission, December 18, 2018. https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf.

[16] It is important to note that the field of AI is much broader than symbolic AI and machine learning represented in this visual. Yet, this report primarily focuses on these areas.

Table 1. Overview of AI capabilities

AI Technique	Definition[17]
Symbolic AI	Systems that function using an ""inference engine': a set of logical commands according to which information is interpreted and related to a set of possible outputs". Capable of using both inductive and deductive approaches to processing data.
Machine learning	Systems that "identify patterns in a dataset, edit inferential rules based on those patterns and connections, and then judge the fitness of that model's outputs against the requirements set by the human programmers". ML can be facilitated at different phases of a system's development through: • *Supervised learning* – A system that finds patterns in data labeled by humans. • *Semi-supervised learning* – A system that finds patterns based on a combination of labeled and unlabeled data. • *Unsupervised learning* – A system that uses unlabeled data and exploits identified connections and patterns, largely independently. • *Reinforcement learning* – "A model of learning that is based on exploration and trial-and-error without explicitly relying on existing data inputs."[18]
Deep learning	A subset of ML in which systems use artificial neural networks to attempt to mimic the learning process of the human brain. Neural networks consist "of a web of interconnected entities known as nodes".
Generative AI	A subset of deep neural networks consisting of "computational techniques that are capable of generating seemingly new, meaningful content such as text, images or audio from training data."[19]

Current applications of AI often combine multiple techniques, such as ML and elements of symbolic reasoning, in so-called neuro-symbolic or hybrid AI. Compared to earlier uses of AI in the military domain, current challenges stem primarily from emergent qualities of ML models, outlined in Table 2, and the complexities of human interaction with and understanding of such systems.

[17] Brenda Leong and Sara R. Jordan, *The Spectrum of Artificial Intelligence – Companion to the FPF AI Infographic Contents* (Future of Privacy Forum, 2021), https://docslib.org/doc/11532721/the-spectrum-of-artificial-intelligence-companion-to-the-fpf-ai-infographic-contents.

[18] Leong and Jordan, *The Spectrum of Artificial Intelligence – Companion to the FPF AI Infographic Contents*, 15.

[19] Stefan Feuerriegel et al., "Generative AI," *Business & Information Systems Engineering* 66, no. 1 (2024): 111–126, https://doi.org/10.1007/s12599-023-00834-7.

Table 2. **Common limitations of ML systems**

Limitation[20]	Description	Cause
Model dysfunction and performance limits	ML models can generate plausible sounding but factually incorrect or fabricated outputs,[21] as well as sycophantic outputs, repeating or agreeing with user inputs even when they are inaccurate.	These issues arise due to limitations in the model's training objectives and the lack of an internal representation of truth, leading to inconsistent and unreliable performance.
Unpredictable and emergent outputs	ML models, especially those trained on large-scale datasets, often display outputs or capabilities not explicitly programmed for or anticipated during development.	Unpredictability results from complex interactions between model architecture, training data, and task formulation, as well as the application of general-purpose technologies in different contexts.
Hidden and systemic bias in data and outputs	Training data used in ML models can contain statistical irregularities or imbalances reflecting societal, institutional, or operational prejudices.	Because ML models learn correlations rather than causal structures, they may reinforce data biases during inference, leading to skewed outcomes.
Brittleness, poor data integrity, and representativeness	ML models may perform poorly when exposed to inputs that deviate from what was included in the training data or even when exposed to only relevant data.	These issues stem from data that is insufficiently diverse, relevant, or reliable; and is also noisy, corrupted, or incomplete. Even with high-quality data, AI models may still behave unpredictably due to underlying limitations.
Opacity and lack of explainability	The process through which certain inputs are converted into outputs is not readily interpretable by humans. This is especially an issue for deep learning systems.	Due to a lack of pre-defined rules in ML systems, and convoluted layers of parameters in deep neural networks, systems have to be made "explainable", or understandable for humans.
Vulnerability to external interference	ML models are susceptible to external manipulation through data poisoning, model inversion, and the insertion of back doors which can undermine model integrity and functionality.	The vulnerabilities stem from the ways in which ML models learn from training data, as well as issues related to opacity, which can obscure external manipulations.

[20] Yasmin Afina and Sarah Grand-Clément, *Bytes and Battles: Inclusion of Data Governance in Responsible Military AI*, No. 308 (Center for international governance and research, 2024), https://www.cigionline.org/publications/bytes-and-battles-inclusion-of-data-governance-in-responsible-military-ai/

[21] These are frequently referred to as 'hallucinations', however, the term is not a technically accepted one. Søren Dinesen Østergaard and Kristoffer Laigaard Nielbo, "False Responses From Artificial Intelligence Models Are Not Hallucinations," *Schizophrenia Bulletin* 49, no. 5 (2023): 1105–107, https://doi.org/10.1093/schbul/sbad068.

2.3 The AI System Lifecycle

A system lifecycle is a structured framework describing the progression of a capability from initial conception to eventual retirement. Here and elsewhere in the report a system is understood, in a broad sense, as a set of technological, human, regulatory, institutional, and organizational components that support specific goals and functions.[22] Understanding how systems progress through a lifecycle is crucial for grasping the full complexity of integrating AI systems into the military domain, where challenges emerge at different stages and accumulate over time if left unaddressed.

The work of GC REAIM employs a lifecycle approach as it provides a practical and principled framework to identify where ethical, legal, and operational responsibilities of different actors arise. The risks and opportunities associated with AI in the military domain are not confined to use, but are shaped by decisions at every phase of development. This approach operationalizes the core concept of responsibility by design, by requiring that anything GC REAIM values and claims 'ought to be the case' or 'ought to be done' must be actively designed for across the AI system lifecycle.

GC REAIM takes the nine phases of the Institute of Electrical and Electronics Engineers Standards Association (IEEE-SA) lifecycle framework, summarized in Table 3, as a starting point.[23] The IEEE-SA framework explicitly accounts for the distributed and iterative nature of AI development and use, as well as the role of human judgment and responsibility, helping establish a normative and operational basis for governance.

Table 3. **The system lifecycle stages as described by the IEEE-SA Lifecycle Framework**

Lifecycle Stage	Description
Before system development	Defining the scope, objectives, constraints, and limitations of a system. Investigating and addressing the trade-offs associated with systems from political, ethical, legal, and technical angles. Defining the rationale for use and formulating system requirements.
Research and development	Designing specific requirements, architecture, functions, use-cases, and interfaces of the system. Data collection, pre-processing, and quality checking. Identifying and addressing concerns that arise during system design. Building the system. Conducting testing, evaluation, validation, and verification (TEVV) as well as assurance checks and legal reviews to ensure that it functions as intended.
Procurement and acquisition	Reviewing and analyzing the requirements of the system and accompanying data to ensure they meet operational and business needs. Further TEVV to ensure the system is fit for purpose.
Testing, evaluation, verification, and validation	Various rounds of procedures to verify that the system can perform its intended function and meet system requirements in line with objectives and intended use.

[22] Olya Kudina and Ibo van de Poel, "A Sociotechnical System Perspective on AI," *Minds and Machines 34* (2024): 21, https://doi.org/10.1007/s11023-024-09680-2.

[23] GC REAIM recognizes that there are various conceptualizations of discrete lifecycle phases and that the order of the stages may vary or overlap. The exact labels are less important than the activities conducted at each stage. Sten Allik et al., *White Paper: A Framework for Human Decision-Making through the Lifecycle of Autonomous and Intelligent Systems in Defense Applications*, White Paper (IEEE SA Research Group on Issues of Autonomy and AI in Defense Systems, 2024), 1–63, https://ieeexplore.ieee.org/document/10707139.

Lifecycle Stage	Description
Considering the human: Education, training, and human-system integration	Designing and developing education, training, and testing procedures for political decision-makers, commanders, and users who will interact with the system. Conducting education and training for human end-users.
Political and strategic considerations	Ensuring that political and strategic leaders, as well as top-level commanders, understand the system, including differences in operating parameters and potential failures.
Operational-level command and control	Managing the introduction to the service of the system and maintaining complete awareness of operating areas. Determining processes by which operational-level commanders decide on system use. Overseeing the system when employed, conducting maintenance throughout use, and updating systems following additional reviews and assessments. Ensuring continuous understanding of performance, potential data drift and adjustment, limitations, failure modes, and risk-mitigation to maintain authority, direction, and control of the system.
Tactical employment	Ensuring that tactical units have a detailed understanding of performance, limitations, failure modes, and risk mitigation measures. Ensuring tactical units knowledgeably monitor systems in line with education and training.
Review, reuse, and/or retire	Determining whether a system should be reused as is, updated, or modified for future use or retired from use in after action reviews. Maintenance and sustainment, including any relevant additional testing, reviews, and data collection. If retiring, going through the steps for decommissioning.[24]

Different actors are, ideally, involved across various stages of the lifecycle for all military systems. Policymakers provide strategic direction and establish governance frameworks. Meanwhile, investment and procurement decisions are typically handled by defense acquisition bodies and business managers overseeing planning and budgeting. Commanders at strategic, operational, and tactical levels translate priorities into mission-specific requirements, assess capabilities, and shape deployment practices.[25] Scientists and developers collaborate to design and refine systems by contributing foundational research and modelling, and ensuring robust implementation, integration, and maintenance across the lifecycle, respectively. Operators, who may be technical experts, employ systems to achieve mission objectives, and their feedback on system performance in operational conditions is vital for refinement. Legal advisors play a cross-cutting role, ensuring legal compliance and supporting accountability mechanisms throughout the lifecycle. Together, these actors form a dynamic, interdependent ecosystem for the development and use of military capabilities.[26]

The engagement of these actors, and how tasks are distributed across the system lifecycle, is crucial not only for mitigating risks, but also for shaping organizational practices that influence the emergence of international norms from the 'bottom up'. Norms, broadly defined as shared understandings of appropriate behavior, can be legal and social in character.[27]

[24] Afina and Paoli, *Governance of Artificial Intelligence in the Military Domain*, 24–25.

[25] The role of commanders differs depending on the unit and command structure.

[26] Lena Tabucco, 'AI-Enabled Autonomous Weapons and Human Control: Part I: Human Control and Machine Learning Design and Development'. International Law Studies 106, nr. 1 (2025). https://digital-commons.usnwc.edu/ils/vol106/iss1/17.; Lena Tabucco, 'AI-Enabled Autonomous Weapons and Human Control: Part II: Human Control and Military Commanders'. International Law Studies 106, nr. 1 (2025). https://digital-commons.usnwc.edu/ils/vol106/iss1/18.; Lena Tabucco, 'AI Enabled Autonomous Weapons and Human Control: Part III: Human Control and System Operators'. International Law Studies 106, nr. 1 (2025). https://digital-commons.usnwc.edu/ils/vol106/iss1/19.

[27] Ingvild Bode, "Contesting Use of Force Norms Through Technological Practices," *Heidelberg Journal of International Law* 83, no.1 (2023): 39–64, https://doi.org/10.17104/0044-2348-2023-1-39.

Given the complex and evolving landscape of AI technologies, the types of practices adopted by states and other relevant actors throughout the system lifecycle will shape social, legal, and technical norms as well as standards.[28] Standards, in turn, are powerful drivers of technological innovation, harmonizing technical specifications to ensure consistent approaches to achieving optimal outcomes across contexts and jurisdictions.[29] Technical, military, legal, and political stakeholders must therefore recognize the normative influence of their choices and exercise care in how their practices shape emerging norms around AI in the military domain.[30]

2.4 Uses of AI in the Military Domain

As a general-purpose technology, AI has uses across strategic, operational, and tactical levels. The specific uses of AI in the military are, at minimum, dependent on the domain in which they are deployed, the type of military activity they are used for, and where on the spectrum of conflict they are employed.[31] GC REAIM broadly divides uses into:

- Non-operational activities: administrative actions contributing to the capabilities militaries deploy in operations, and;
- Operational activities: actions in direct support of military missions like the management and coordination of weapons and warfighters.

Table 4 reflects contemporary examples of AI uses across key areas of the military domain. While grounded in current practice, these examples offer a foundation for identifying emerging trends and informing forward-looking recommendations.[32]

Spotlight Box 1: AI and strategy[33]

Beyond the operational and non-operational concepts used by GC REAIM lies “strategy”: activities primarily led by civilian authorities which establish the direction for military functions. These activities include determining overall geopolitical goals; the formation and dissolution of alliances; the design of, commitment to, and abrogation of treaties (mutual defense, non-aggression, geographical constraints, weapons constraints); determining overall budgets, budget allocations, and make-up of forces; and decisions to initiate or terminate hostilities. The use of AI for strategy goes beyond the use of AI for policy in general. Using AI for strategy affects national leadership decision-making which constructs the context and direction for militaries. At present, there appear to be no examples of fielded AI systems used for regular decisions at this level. However, there is interest in technologies advanced to the point of providing useful guidance for leaders.

[28] Ingvild Bode, “Practice-Based and Public-Deliberative Normativity: Retaining Human Control over the Use of Force,” *European Journal of International Relations* 29, no. 4 (2023): 990–1016, https://doi.org/10.1177/13540661231163392.

[29] Raquel Delgado-Aguilera Jurado et al., “An Introduction to the Current State of Standardization and Certification on Military AI Applications,” *Journal of Air Transport Management* 121 (November 2024): 102685, https://doi.org/10.1016/j.jairtraman.2024.102685.

[30] AutoPractices Project, “Map of Practices V1,” *Center for War Studies*, Forthcoming.

[31] Meerveld and Lindelauf introduce a framework that produces 75 contexts with unique requirements and considerations. Meerveld and Lindelauf, *Context Is Everything: Policy Implications of the Military AI Responsibility Contextualization Framework*.

[32] While it is possible to identify grey areas between use categories, or bin them another way, such as by service branch, a context-informed overview can ground discussions in the reality of what AI technologies can and cannot do.

[33] For more on the development of national strategies on AI in defense and security, see Yasmin Afina, *Draft Guidelines for the Development of a National Strategy on AI in Security and Defence – A Policy Brief*, Policy brief (UNIDIR, 2024), https://unidir.org/wp-content/uploads/2024/10/guidelines_for_the_development_of_national_strategy_web-2.pdf.

Table 4. Overview of use cases of AI in the military domain

	Military Activity	Uses of AI	Examples as of 2025
Non-operational activities	**Policy**	• Information tracking and collection • Information analysis • Generation of recommendations and resource allocation plans • Compliance monitoring • Policy impact assessments	• The UK Strategic Command's usage of Adarga's Vantage AI for information organization.[34] • NATO's AI FELIX, Political Military-Assisted Decision Making (PM-ADM), and Architecture for Interoperable Digital Technologies within the Alliance (AIDA) used for information management.[35]
	Human and Legal Resources	• Recruitment and talent matching • Application reviews • Record and retention management • Document generation • Legal compliance review	• AI is used for HR by the US, the ROK, and India.[36] • AI is used for analyzing and optimizing recruitment language for the UK MoD.[37]
	Research, Development, Testing, and Evaluation (RDT&E)	• Modelling and simulation, performance prediction • Synthetic data generation[38] • Mediation or design of testing • Design of technologies • Research tool	• Defence Research and Development Canada is exploring the use of AI tools to generate "synthetic images in order to train military detection algorithms."[39] • The UK Defense Science and Technology Laboratory uses digital twins and AI to virtually test defence platforms before physical prototyping.[40] • The Global Combat Air Program "Tempest" future fighter program, launched by the UK, Italy, and Japan, relies on AI to compress development timelines by employing digital twins and AI-driven simulations to test flight dynamics.[41]

[34] Shephard News Team in London, "Adarga's Vantage AI Software Selected for UK Strategic Command's Defence Support," September 23, 2024, https://www.shephardmedia.com/news/digital-battlespace/adargas-vantage-ai-software-selected-for-uk-strategic-commands-defence-support/.

[35] The North Atlantic Treaty Organization, "Harnessing Artificial Intelligence: Allied Command Transformation at the Forefront of NATO Innovation," Act NATO, April 16, 2025, https://www.act.nato.int/article/harnessing-artificial-intelligence/.

[36] Nishad Nawaz et al., "The Adoption of Artificial Intelligence in Human Resources Management Practices," *International Journal of Information Management Data Insights* 4, no. 1 (2024), https://doi.org/10.1016/j.jjimei.2023.100208.

[37] Cabinet Office, Department for Science, Innovation and Technology and Government Digital Service, "MOD: Textio," GOV.UK, December 17, 2024, https://www.gov.uk/algorithmic-transparency-records/mod-textio.

[38] Harry Deng, *Exploring Synthetic Data for Artificial Intelligence and Autonomous Systems: A Primer* (United Nations Institute for Disarmament Research, 2023), https://unidir.org/publication/exploring-synthetic-data-for-artificial-intelligence-and-autonomous-systems-a-primer/.

[39] Government of Canada. 'Making Abstract and Complex Systems Understandable: DRDC Uses Animation to Create Digital Models to Support Research and Development'. Innovation, Science and Economic Development Canada, 28 mei 2024. https://science.gc.ca/site/science/en/blogs/defence-and-security-science/making-abstract-and-complex-systems-understandable-drdc-uses-animation-create-digital-models-support.

[40] Digital Catapult. 'DSTL Pit Stop: Intelligent Ship'. *Digital Catapult*, 18 januari 2019. https://www.digicatapult.org.uk/apply/events/dstl-pit-stop-intelligent-ship/.

[41] Douglas Barrie, Elio Calcagno. 'The New Partnership among Italy, Japan and the UK on the Global Combat Air Programme (GCAP)'. Text. IAI Istituto Affari Internazionali, 14 March 2025. https://www.iai.it/en/pubblicazioni/c04/new-partnership-among-italy-japan-and-uk-global-combat-air-programme-gcap.

	Military Activity	Uses of AI	Examples as of 2025
Non-operational activities (cont.)	**Acquisition and Procurement**	• Contract generation • Contractual compliance oversight • Financial auditing • Research tool	• The US DoD's Defense Logistics Agency is experimenting with AI to support financial auditing.[42] • The UK MoD used Palantir AI to support its 2025 Strategic Defense Review.[43]
	Training	• Modelling and simulation, virtual and augmented reality • Mediation and design of training • Training evaluations and interventions • Personalized training plans	• NATO use of AI for wargaming and simulation.[44] • The UK Project OdySSEy single synthetic military training environment.[45] • Xi'an Technological University developed an interactive battle simulation platform using Deepseek.[46]
	Logistics	• Scheduling and planning • Resource allocation and inventory management • Predictive and autonomous resupply and maintenance • Security monitoring	• The US Army "Condition-Based Maintenance Plus" incorporates predictive analytics and AI to anticipate equipment wear and conduct preventative maintenance.[47] • France uses Robotic Process Automation (RPA) to manage and maintain military equipment in their warehouses.[48] • NATO has experimented with using AI to improve tracking, efficiency, and security of supply chain logistics.[49]
	Health and Medicine	• Record management • Resource allocation • Scheduling and planning • Expertise assistance • Automated assistance • Treatment monitoring • Discovery and production	• The UK MoD uses AI to check applicants' medical records for eligibility.[50] • ReFit is an autonomous control algorithm enabling closed-loop resuscitation based on real-time physiological monitoring.[51]

[42] Beth Reece, "DLA Finance Pursues Artificial Intelligence to Pass Financial Audit," Defense Logistics Agency, November 21, 2024, https://www.dla.mil/About-DLA/News/News-Article-View/Article/3971065/dla-finance-pursues-artificial-intelligence-to-pass-financial-audit/.

[43] Stefan Boscia, "Rise of the Robots: AI to Shape UK Defense Review," Politico, September 10, 2024, https://www.politico.eu/article/artificial-intelligence-united-kingdom-defense-review/.

[44] The North Atlantic Treaty Organization, 2025, "Harnessing Artificial Intelligence."

[45] BAE Systems,"Project OdySSEy Single Synthetic Environment", BAE Systems, July 22, 2025, https://www.baesystems.com/en/product/project-odyssey.

[46] Xi Yu, Liu Xuanzun. 'Chinese university unveils new DeepSeek-based simulated military scenario generator - Global Times'. 15 May 2025. https://www.globaltimes.cn/page/202505/1334151.shtml.)

[47] Defense Acquisition University, "Condition Based Maintenance Plus (CBM+)," DAU, July 22, 2025, https://www.dau.edu/acquipedia-article/condition-based-maintenance-plus-cbm.

[48] Murielle Delaporte, 'France – Frugality, Modernity and Mutualization: A Triptych for Strategic Thickness and Better Sustainability of Land Equipment', *EUROSATORY*, 15 April 2023, https://www.eurosatory.com/en/france-frugality-modernity-and-mutualization-a-triptych-for-strategic-thickness-and-better-sustainability-of-land-equipment/.

[49] Robert-Cristian Trif, 'The Role of NATO Support in Strengthening Military Supply Chains During the Ukraine Conflict', in *ResearchGate*, 2025, https://doi.org/10.17758/EARES20.EAP0125122; SEKO Logistics, 'Integrating Advanced Technologies in Military Logistics', 12 February, 2024, https://www.sekologistics.com/en/resource-hub/knowledge-hub/integrating-advanced-technologies-in-military-logistics/.

[50] Joe Saballa, "British Army Now Uses AI to Speed Up Recruitment," TheDefensePost, February 19, 2024, https://thedefensepost.com/2024/02/19/british-army-ai-recruitment/.

[51] "Autonomous Trauma Care Extends 'Golden Hour' for Saving Lives," News, Carnegie Mellon University, May 24, 2024, https://www.cmu.edu/news/stories/archives/2024/may/autonomous-trauma-care-extends-golden-hour-for-saving-lives.

	Military Activity	Uses of AI	Examples as of 2025
Operational activities	**Command, Control, and Communications (C3)**	• Scheduling and planning • Information fusion • Decision-support systems • Communication analysis and optimization • Battle management	• NATO acquired AI "Maven Smart System" C3 software from Palantir to bolster its operational capabilities.[52] • The Australian Army has partnered with Microsoft to use AI for combat net radio analysis and transcriptions.[53]
	Intelligence, Surveillance, and Reconnaissance (ISR)	• Active and passive information tracking and collection • Information analysis • Open source intelligence • Information fusion • Interrogation	• Singapore's Cyber Security Agency, MoD and the Digital and Intelligence Service use AI-driven analytics to scan cyber and information grids.[54] • Ukraine uses AI-enhanced commercial drones to conduct low-cost, high-frequency tactical ISR missions – autonomously identifying and geolocating enemy positions in real time, even in GPS-denied environments.[55] • The People's Liberation Army Academy of Military Science used Meta's Llama as a foundation to develop a military-focused large language model, ChatBIT, to gather and process both open-source and standard signals military intelligence.[56]
	Platforms and Weapons Systems	• Platform control • Sensing, navigation, movement and maneuvering, and communication • Target detection, identification and selection • Weapon release and precision guidance	• Ukraine uses AI-powered "Sky Sentinel" automated turrets to shoot down Russian drones.[57] • The US and other states developed, use, and export Close-in Weapon Systems, which can autonomously detect and destroy incoming missiles and aircraft mounted on naval vessels. They have been deployed by 40+ states since the 1980s.[58] • Israel has fielded Harpy 'fire and forget' loitering munitions, which can autonomously detect and destroy radar systems to weaken air defenses.[59] • The Australian Army is experimenting with autonomous teaming of unmanned ground vehicles (UGVs) and unmanned aerial systems executing reconnaissance and strike missions.[60] • Russia developed the V2U drone, which can target even when communication with the operator is lost.[61]

[52] Tim Bradshaw and Henry Foy, "Nato Acquires AI Military System from Palantir" , *Financial Times*, April 14, 2025, https://www.ft.com/content/7f80b1bc-114c-4a00-ad06-6863fb435822.

[53] Defense Advancement Staff, "Australian Army Explores Using AI For Command and Control," Defense Advancement, December 22, 2021, https://www.defenseadvancement.com/news/australian-army-explores-using-ai-for-command-and-control/.

[54] Michael Raska, "Reimagining Defense Innovation: Defense AI in Singapore," in *The Very Long Game: 25 Case Studies on the Global State of Defense AI*, ed. Heiko Borchert et al. (Springer Nature Switzerland, 2024), https://doi.org/10.1007/978-3-031-58649-1_25.

[55] Kateryna Bondar, *Ukraine's Future Vision and Current Capabilities for Waging AI-Enabled Autonomous Warfare* (Canadian Security Intelligence Service, 2025), https://www.csis.org/analysis/ukraines-future-vision-and-current-capabilities-waging-ai-enabled-autonomous-warfare.

[56] James Pomfret, and Jessie Pang. 'Exclusive: Chinese Researchers Develop AI Model for Military Use on Back of Meta's Llama'. Artificial Intelligence. Reuters, 1 November 2024. https://www.reuters.com/technology/artificial-intelligence/chinese-researchers-develop-ai-model-military-use-back-metas-llama-2024-11-01/.

[57] Sinéad Baker and Jake Epstein, "Ukraine Is Using an AI-Powered, Automated Turret to Shoot down Russia's Devastating Shahed Drones," Business Insider, June 2, 2025, https://www.businessinsider.com/ukraine-ai-powered-turret-shoot-russia-shahed-drones-sky-sentinel-2025-6.

[58] THALES, "Goalkeeper – Close-in Weapon System," THALES, https://www.thalesgroup.com/en/goalkeeper-close-weapon-system.

[59] Ingvild Bode and Tom Watts, 'Loitering Munitions and Unpredictability: Autonomy in Weapon Systems and Challenges to Human Control' (Zenodo, May 2023), https://doi.org/10.5281/ZENODO.7860762.

[60] Krysten Captain Clifton, "A Powerful Display on Show," Australian Government – Defence, Defence Australia, October 10, 2024, https://www.defence.gov.au/news-events/news/2024-10-10/powerful-display-show.

[61] Daniel Boffey, "Killing Machines: How Russia and Ukraine's Race to Perfect Deadly Pilotless Drones Could Harm Us All," World News, *The Guardian*, June 25, 2025, https://www.theguardian.com/world/2025/jun/25/ukraine-russia-autonomous-drones-ai.

Military Activity		Uses of AI	Examples as of 2025
Operational activities (cont.)	**Human Augmentation**	• Virtual and augmented reality tools • Cognitive enhancements • Physical enhancements	• The Australian Army robotic control of small UGVs through non-invasive AI brain-robotic interfaces.[62] • Optimizing pilot training with virtual reality simulators using neurophysiological sensors[63]
	Cyber and Information	• Cybersecurity tracking • Reactive and proactive defense • Communication • Content and code generation • Information operations • Deepfake/deception detection	• The US has developed an AI tool for cyber-space to detect malicious traffic called Panoptic Junction.[64] • Singapore's Cyber Security Operations Centre 2.0 Team embeds AI and ML for continuous anomaly detection and adaptive learning over large data streams.[65] • The UK's Defence Science and Technology Laboratory is developing AI "cyber agents" under the Autonomous Resilient Cyber Defence program to autonomously detect, respond to, and recover from attacks on military platforms.[66]

Given the range of uses for AI throughout the military domain, it is essential to identify which applications should be prioritized in governance. When the technical intricacies of AI intersect with the unpredictability of military operations, novel challenges arise.[67] AI systems may fail when faced with dynamic problems and environments, and complicate decision-making. This is particularly true in adversarial environments where opponents directly disrupt or mislead AI-enabled systems, or when AI architectures make tracing malfunctions or unexpected outputs increasingly difficult or impossible.[68] Not all uses of AI present the same level of risk, relevance, or ethical and legal complexity. The effects of AI in specific contexts are, at a minimum, dependent on:

1. the extent to which a use of AI is specific to core military activities and thus distinct from civilian AI applications, raising unique governance issues that differ from dual-use or civilian-equivalent applications,
2. the involvement in safety-critical systems which require stricter oversight,
3. the degree to which ML underpins system functionality in core military activities,
4. the influence of systems-of-systems dynamics in military activities.

The overlapping of these four factors, outlined in Table 5, makes certain uses of AI disproportionately consequential, warranting focused attention through governance.

[62] Matthew Bickerton, "Brain waves control robot dog's moves," Website, Australian Government – Defence, Defence Australia, June 7, 2022, https://www.defence.gov.au/news-events/news/2022-06-07/brain-waves-control-robot-dogs-moves.

[63] Evy Van Weelden et al., "A Passive Brain-Computer Interface for Predicting Pilot Workload in Virtual Reality Flight Training," paper presented at 2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS), 2024, https://doi.org/10.1109/ichms59971.2024.10555679.

[64] Mark Pomerleau, "Army Cyber AI Monitoring Tool Moves to 12-Month Pilot," DefenseScoop, November 18, 2024, https://defensescoop.com/2024/11/18/army-cyber-ai-panoptic-junction-monitoring-tool-12-month-pilot/.

[65] MINDEF Singapore. 'Fact Sheet: Defence Technology Prize 2020 Team (Engineering) Award Winner'. Ministry of Defence. Geraadpleegd 14 september 2025. https://www.mindef.gov.sg/news-and-events/latest-releases/30oct20_fs7/.

[66] QinetiQ. 'Autonomous Resilient Cyber Defence'. Referenced 15 September 2025. https://www.qinetiq.com/en/what-we-do/services-and-products/autonomous-resilient-cyber-defence?utm_source=chatgpt.com.

[67] Michael Raska, *Mitigating the Risks of AI-Driven OODA Loops in Military Decision-Making*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/04/Raska.pdf.

[68] Meerveld and Lindelauf, *Context Is Everything: Policy Implications of the Military AI Responsibility Contextualization Framework*.; Ariel Conn et al., *An Approach for Assessing Autonomous and AI-Enabled Capabilities within Weapons Systems*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Conn-et-al.pdf.

Table 5. **Fundamental criteria affecting the risk, relevance, or complexity of AI use cases**

Measure	Degrees	Explanation
The extent to which a use of AI is specific to core military activities and thus distinct from civilian AI applications	*High distinctiveness* – Both the activity and the AI systems being used differ greatly from what can usually be found in the civilian domain.	Evaluating this dimension is essential given the dual-use nature of many AI applications, where civilian systems may be adapted for military purposes. Focusing governance on high-impact, military-specific systems ensures attention is directed where consequences are severe, clarifies the scope of application, and addresses mission-specific uses which raise distinct governance and security concerns.
	Moderate distinctiveness – To some extent, either the activity or the AI systems being used differ from what can be found in the civilian realm.	
	Low distinctiveness – The activities or the AI systems being used display a significant overlap with the civilian domain.	
The involvement of AI in safety-critical systems and the potential consequences of failure	*High involvement* – The activity relies heavily or exclusively on safety-critical systems, where failure can result in immediate and significant harm.	Safety-critical systems are those where failure can result in loss of life, significant damage to infrastructure, or harm to the environment. In the military context, this includes mission failure. Focusing on safety-critical systems ensures that oversight efforts target areas with the greatest potential for harm.
	Moderate involvement – The activity involves a mix of safety-critical and non-safety-critical systems, with failures potentially impacting operations but in a limited or indirect way.	
	Low involvement – The activity involves few or no safety-critical systems, with minimal consequences in the event of failure.	
The degree to which ML underpins system functionality	*High ML dependence* – ML is central to the system's core functionality.	The degree to which ML underpins a system is a critical factor for governance prioritization because ML introduces unique challenges not present in rule-based systems. Prioritizing systems with high ML dependence allows governance efforts to focus on the most complex and least predictable technologies.
	Moderate ML dependence – ML is integrated into parts of the system, but system performance does not depend entirely on ML.	
	Low ML dependence – The system relies primarily on rule-based logic, with minimal or no integration of ML components.	
The influence of systems-of-systems dynamics that may give rise to compounded risks	*High influence* – The activity is embedded in other activities, where issues can create complex and cascading effects.	When AI-enabled systems interact across different activities, processes, domains, and levels of decision-making, failures or unintended behaviors can have compounded and cascading effects. Governance frameworks must therefore account not only for individual system performance but also for how systems function collectively.
	Moderate influence – The activity is dependent on other activities in limited or structured ways, where some ripple effects may occur.	
	Low influence – The activity occurs largely in isolation. With limited interaction or dependency, risks are contained.	

GC REAIM finds that C3, ISR, platforms and weapons systems, and cyber and information emerge as the highest priority areas for governance efforts. However, non-operational activities like RDT&E, acquisition and procurement, and training must also receive attention as parts of the lifecycle with significant compounding effects, regardless of whether AI is used. The rest of the report will focus on addressing challenges in relation to these areas while acknowledging that their impact also stems from how they interact with other use cases throughout the lifecycle and in various contexts.

Table 6. **Identification of priority use case**

Military Activity	Distinction Between Civilian-Military AI Applications	Involvement of Safety Critical Systems	Role of Machine Learning	Systems-of-systems Risk	Priority for Governance[69]
Policy	Low; processes largely analogues to civilian uses	Low	Moderate	Moderate	Military uses of AI in policy closely resemble civilian applications, relying on similar methods of information gathering, analysis, and presentation. The effect of AI in policy appears through compounded effects further in system lifecycles. As such, AI in policy in itself will not be focused on in the report.
Human and Legal Resources	Low; processes largely analogues to civilian uses	Low	Low	Low	Military uses of AI for human and legal resources applications do not differ significantly from the civilian realm. As such, AI in human and legal resources in itself will not be focused on in the report.
Research, Development, Testing, and Evaluation	Moderate; linked to lifecycle-specific military procedures	Moderate (indirect)	Moderate, especially in TEVV processes	Moderate	Military uses of AI for RDT&E can have compounded effects across the system lifecycle, particularly in relation to TEVV. As most associated issues arise regardless of whether AI is used or not, AI for RDT&E is examined as a priority in the broader context of the system lifecycle.
Acquisition and Procurement	Low; closely mirrors commercial practices	Moderate (indirect)	Moderate, but growing	Low	Military uses of AI for acquisition and procurement processes largely mirror civilian applications and most associated issues, regardless of whether AI is used. As such, the report rather focuses on the acquisition and procurement of AI systems themselves.
Training	Moderate; civilian analogues exist, but with military-specific applications and impacts	Moderate (indirect)	Moderate	Moderate	Military uses of AI for training may introduce significant cognitive and psychological effects. It is necessary to address how training practices need to be adjusted given the widespread introduction of AI across military activities. This is an important area for examination.
Logistics	Low; parallels civilian logistics systems	Moderate (as operational enabler)	Moderate	Moderate	Military uses of AI for logistics parallel civilian applications, making it less of a priority area in and of itself. Yet, the role of logistics will be considered as contributing to opportunities and risks in operational activities.

[69] The evaluations contained in this table are based on collective GC REAIM inputs gathered during Commission Meetings and the drafting process.

Military Activity	Distinction Between Civilian-Military AI Applications	Involvement of Safety Critical Systems	Role of Machine Learning	Systems-of-systems Risk	Priority for Governance
Health and Medicine	Low; heavily overlaps with civilian practices	High	Moderate	Low	Military uses of AI for health and medicine overlap with civilian practices, reducing its priority, despite its high safety-critical nature.
Command, Control, and Communications	High; embedded in military-specific operational structures	High	High	High, impacting all operational activities	Military uses of AI for C3 in the conventional and nuclear arenas are a high-risk area for AI usage because of the critical role of C3 to all military operations, the brittleness of AI, and the grave consequences of mistakes. This is a priority area for examination.
Intelligence, Surveillance, and Reconnaissance	Moderate to high; similar systems in military-specific contexts	High	High	High	Military uses of AI for ISR activities are proximate to the battlefield with potential high impact. Given that ISR data feeds through C3 systems, there are compounded risks posed by ISR-C3 interaction. This is a priority area for examination.
Platforms and Weapons Systems	High; systems unique to the military domain	High	High	High	Military uses of AI for platforms and weapon systems are a safety-critical use of AI with enormous potential consequences for the battlefield, making it a high-priority area for examination.
Human Augmentation	Moderate to high; certain applications are unique to warfighters	High	Moderate, but growing	Moderate	Military uses of AI for human augmentation of warfighters are a high priority, though the issues it raises are also addressed through other categories. Of highest priority are considerations of how AI is used to make decisions for tasks that can directly lead to loss of life.
Cyber and Information	High; unique defensive and offensive systems	High	High	High	Military uses of AI for offensive cyber operations, as well as their interaction with kinetic capabilities, can have substantial consequences, making it a high-priority area for examination.

Case Study: Artificial Intelligence-enabled Decision Support Systems [Part 1]

AI is increasingly being used in military decision support systems (DSS): "tools that use AI techniques to collect and analyze data, provide information about the operational environment as well as actionable recommendations, with the aim of aiding military decision-makers."[70]

AI-DSS have become prevalent across all levels of military operations. Their wide range of potential uses in the military domain can be categorized into three broad types of functions: (1) description and analysis, which involves collecting, organizing, and presenting data; (2) prediction and extrapolation, which entails identifying patterns and trends in data, including forecasting possible outcomes and their probabilities; and (3) prescription, which focuses on recommending the most effective course(s) of action.[71]

For example, the Singapore Armed Forces (SAF) – in collaboration with the Defense Science and Technology Agency (DSTA) – is testing and gradually deploying an advanced Command and Control Information System (CCIS). Structured around a DSS, the CCIS leverages AI-enabled tools, such as data analytics and weapon-to-target matching algorithms, to generate real-time operational insights and propose optimized strike options.[72] AI-DSS are also reportedly widely employed in current conflicts, including in Gaza and in the Russia-Ukraine War.[73]

AI-DSS represent a cross-cutting capability spanning multiple military functions and exemplifies many core challenges associated with integrating AI into the military domain.[74] Already deployed to support kinetic operations, they reflect significant security, legal, and humanitarian concerns.

[70] Marta Bo, Ingvild Bode, Jessica Dorsey and Elke Schwarz; intervention to Report of UN Secretary General,"Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security," 5 June 2025, https://docs.un.org/en/A/80/78, p. 139-143.

[71] Anna Nadibaidze et al., *AI in Military Decision Support Systems – A Review of Developments and Debates* (Center for War Studies, 2024), https://findresearcher.sdu.dk/ws/portalfiles/portal/275893410/AI_DSS_report_WEB.pdf; Jimena Sofía Viveros Álvarez, "The Risks and Inefficacies of AI Systems in Military Targeting Support," *Humanitarian Law & Policy Blog of the International Committee of the Red Cross*, September 4, 2024, https://blogs.icrc.org/law-and-policy/2024/09/04/the-risks-and-inefficacies-of-ai-systems-in-military-targeting-support/.

[72] Ministry of Defence Singapore, "Fact Sheet: SAF Harnesses Artificial Intelligence and Data Analytics to Sharpen Sense and Strike Capabilities with Command and Control Information System," MINDEF Singapore, September 23, 2021, https://www.mindef.gov.sg/news-and-events/latest-releases/23sep21_fs2.

[73] Anna Nadibaidze, "Do AI Decision Support Systems 'Support' Humans in Military Decision-Making on the Use of Force?," Opinio Juris, November 29, 2024, https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force/.

[74] Alexander Blanchard and Laura Bruun, *Autonomous Weapon Systems and AI-Enabled Decision Support Systems in Military Targeting: A Comparison and Recommended Policy Responses* (Stockholm International Peace Research Institute, 2025), https://doi.org/10.55163/yqby3151

3 Implications of AI in the Military Domain

The extent to which a specific use of AI in the military domain constitutes a risk or opportunity depends on the interplay between the system, its context, human decision-making around it, and ultimately its compliance with international law. Stakeholders must assess implications through a context-specific lens, which allows for accurate and grounded understandings of relevant ethical and legal principles as well as informed development of governance frameworks.

There are at least three distinct yet interrelated contexts critical for understanding the risks and opportunities of AI in the military domain: (1) **international peace and security**; (2) **the conduct of armed conflict**; (3) broader impacts or uses **beyond armed conflict**. Uses of AI across these contexts, in turn, have overarching implications for human agency, responsibility, and accountability, identified in Section 3.4.

GC REAIM emphasizes that the implications of AI in the military domain must be assessed in light of compliance with international law and ethics. Each subsection outlines the applicable international law and norms shaping interpretations of what constitutes risks or opportunities. International law provides the obligations, thresholds, and standards for addressing risks. While the novelty of AI may require clarification or elaboration of existing rules, these rules remain the foundation for effective governance.[75]

What is labelled as "*risks of AI*" and "*opportunities of AI*" cannot be based on what systems could "*potentially do*" or on unverified claims of what they can do. Assertions based on speculative capabilities risk obscuring the challenges, opportunities, and risks of these systems. To avoid over- or underestimation, discussions must be evidence-based and tempered by understandings of technological limitations, uncertainties, and international dynamics.[76]

[75] Keketso Kgomosotho, *New and Emerging Terminologies in Ethical AI Principles: Exploring the International Law Implications in the Military Context*, GC REAIM Expert Policy Note Series (The Global Commission on Responsible Artificial Intelligence in the Military Domain, 2025), https://hcss.nl/wp-content/uploads/2025/05/Conn-et-al.pdf.

[76] Raluca Csernatoni, "Governing Military AI Amid a Geopolitical Minefield," Carnegie Endowment for International Peace, 2024, https://carnegieendowment.org/research/2024/07/governing-military-ai-amid-a-geopolitical-minefield?lang=en. ; Tshilidzi Marwala, "Militarization of AI Has Severe Implications for Global Security and Warfare," United Nations University, July 24, 2023, https://unu.edu/article/militarization-ai-has-severe-implications-global-security-and-warfare.

3.1 The Context of Peace and Security

The integration of AI into military operations affects peace and security globally, regionally, and nationally.[77] The application of AI to promote peace and security fundamentally relies on informed human decisions in developing and using these technologies. As with any advanced technology, it is essential to carefully consider the organizational structures and institutional frameworks involved, ensuring that they align with shared ethical and legal principles and rules.[78] The future of global security must be anchored in a commitment to ethics, rule of law, and, above all, the dignity and safety of human beings.[79]

3.1.1 Applicable International Law and Regulatory Frameworks in the Context of Peace and Security

This section outlines the legal and regulatory frameworks which respectively provide the contours for the pursuit of military AI capabilities in the context of peace and security.[80]

Law regulating the use of force

The principal branch of international law applicable in the context of peace and security is the UN Charter regime governing the use of force (*jus ad bellum*). Much of international law is technology-neutral, meaning that rules apply regardless of the tools employed by states.[81] Therefore, the *jus ad bellum* applies fully to any use of AI in the military domain. It governs the legality of the use of force against the territorial integrity or political independence of another state.[82] Moreover, the *jus ad bellum* prohibits such force, with limited exceptions, including self-defense, authorization by the UN Security Council, and, in some cases, by consent from the affected government.[83] This prohibition is recognized as a peremptory norm (*jus cogens*) from which no derogation is permitted. The integration of AI into the military domain cannot be taken as an opportunity to bypass or dilute these foundational legal constraints.[84] Compliance with applicable international law is a requirement in the development and use of military AI.[85]

[77] Denise Garcia, *The AI Military Race: Common Good Governance in the Age of Artificial Intelligence* (Oxford University Press, 2024), https://doi.org/10.1093/oso/9780192864604.001.0001.; Ioana Puscas, *AI and International Security – Understanding the Risks and Paving the Way for Confidence-Building Measures* (UNIDIR, 2023), https://unidir.org/wp-content/uploads/2023/10/UNIDIR__risks_paving_the_path_for_confidence_building_measures.pdf.; Jill Hruby and M. Nina Miller, *Assessing and Managing the Benefits and Risks of Artificial Intelligence in Nuclear-Weapon Systems*, NTI Paper (Nuclear Threat Initiative, 2021), https://www.nti.org/wp-content/uploads/2021/09/NTI_Paper_AI_r4.pdf.

[78] Michele Giovanardi, "AI for Peace: Mitigating the Risks and Enhancing Opportunities," *Data & Policy* 6 (2024), https://doi.org/10.1017/dap.2024.37.

[79] There are valuable lessons that can be learnt from the United Nations Development Programme, *The Women, Peace, and Security Agenda,* (United Nations, 2023); and the "Universal Guidelines for AI," Center for AI and Digital Policy, 2018, https://www.caidp.org/universal-guidelines-for-ai/.

[80] William H. Boothby, *Weapons and the Law of Armed Conflict*, 2nd ed. (Oxford University Press, 2016), 978-0-19-104416-8; Stuart Casey-Maslen and Tobias Vestner, *A Guide to International Disarmament Law* (Routledge, 2019), https://doi.org/10.4324/9781351108119.

[81] Rosemary Rayfuse, "Public International Law and the Regulation of Emerging Technologies," in *The Oxford Handbook of Law, Regulation and Technology*, ed. Roger Brownsword et al. (Oxford University Press, 2017), https://doi.org/10.1093/oxfordhb/9780199680832.013.22.

[82] United Nations, *Charter of the United Nations*, 26 June 1945, 1 UNTS XVI, Article 2(4).

[83] See generally Christine Gray, *International Law on the Use of Force* (4th ed. 2018).

[84] "[T]he prohibition of the use of force.... is universally recognized as a jus cogens principle, a peremptory norm from which no derogation is permitted." see Mary Ellen O'Connell, "Seductive Drones: Learning from a Decade of Lethal Operations," *Journal of Law, Information and Science* 21, no. 2 (2011): 116–39, https://doi.org/10.3316/informit.034529451756320.; International Court of Justice, "Advisory Opinion of the International Court of Justice on the Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory," United Nations General Assembly, 2004, https://www.un.org/unispal/document/auto-insert-178825/.

[85] Jimena Sofía Viveros Álvarez, "Autonomous Weapons Systems and the Use of Force," in *Artificial Intelligence: Crime, War, and Justice,* ed. Nathalie Rébé (Ethics International Press Ltd, 2023).

As in all cases where the use of force is legally permitted, a state's use of AI systems will be subject to the principles of ***necessity*** (the use of force is a last resort, after all peaceful options have been exhausted, leaving no reasonable alternatives), ***proportionality*** (the use of force must be commensurate to the scale and effects of the attack, and must aim to end the attack and neutralize the threat of further harm), and ***immediacy*** (the defensive actions must be taken while the attack is still in progress or not long after; for anticipatory self-defense the attack must be "imminent" requiring an "instant and overwhelming necessity" with "no choice of means" and "no moment of deliberation).[86] The element of ***attribution*** is required as well: the acts of a state or non-state actor must be imputable to the state and thus part of the state's legal responsibility.[87] For non-state actors, this usually requires that they are under a state's effective control to which its acts are attributable.[88]

Other important norms and initiatives relevant to peace and security

Non-binding regulatory norms and initiatives further influence what is perceived as a potential risk or opportunity of AI in the context of peace and security. A series of Track-II dialogues has contributed to emerging norms around AI in the military domain.[89] Simultaneously, through Track I fora in November 2023, the US and Chinese heads of state agreed on dialogue focused on "risks and safety issues associated with artificial intelligence". This Sino-US dialogue led to a November 2024 agreement on the necessity of human control over the decision to use nuclear weapons.[90]

Given the dual-use nature of AI systems, certain principles derived from civilian AI initiatives also contribute to governance foundations in the military domain. A two-tier approach, containing both prohibitions and regulations as well as establishing red lines, is increasingly gaining support and recognition internationally.[91] Notably, the UNSG High-Level Advisory Body on AI emphasized the need "to identify clear red lines delineating unlawful use cases, including relying on AI to select and engage targets autonomously," noting that "[e]ventually, some kind of mechanism at the global level might become essential to formalize red lines if regulation of AI needs to be enforceable."[92] Similarly, the first round of the International Dialogues on AI Safety in 2024 produced an expert-led Consensus Statement on Red Lines in AI, arguing that "...unsafe development, deployment, or use of AI systems may pose catastrophic or even

[86] It should be noted that preventive or pre-emptive self-defense is the subject of considerable controversy under international law. Some recognize anticipatory self-defense in cases where armed attack is imminent. D Akande, C Heyns, L Hill-Cawthorne, and T Chengeta, "Right to life and International Law Framework Regulating the Use of Armed Drones" in D Akande et al (eds) Human rights and 21st challenges (Oxford University Press, 2020).

[87] The International Law Commission's Articles on Responsibility of States for Internationally Wrongful Acts (2001), Articles 4 to11.

[88] Jan Arno Hessbruegge, "The Historical Development of the Doctrines of Attribution and Due Diligence in International Law," *N.Y.U. J. INT'L L. & POL.* 36 (2004): 265; Christine Gray, *International Law and the Use of Force*, 4th ed., Foundations of Public International Law (Oxford University Press, 2018), 41–50, https://global.oup.com/academic/product/international-law-and-the-use-of- force-9780198808428?cc=nl&lang=en&.; Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States of America), 70 1986 I.C.J. Rep. 14 (International Court of Justice 1986), https://www.icj-cij.org/sites/default/files/case-related/70/070-19860627-JUD-01-00-EN.pdf.

[89] Sarah Shoker et al., "Confidence-Building Measures for Artificial Intelligence: Workshop Proceedings," arXiv:2308.00862, preprint, August 2023, arXiv, https://doi.org/10.48550/arXiv.2308.00862; Michael C. Horowitz et al., "Policy Roundtable: Artificial Intelligence and International Security," with Mary (Missy) Cummings et al., Texas National Security Review, June 2, 2020, https://tnsr.org/roundtable/policy-roundtable-artificial-intelligence-and-international-security/.

[90] Sydney J. Freedberg Jr, "Biden Launches AI 'Risk and Safety' Talks with China. Is Nuclear C2 a Likely Focus?," Breaking Defense, November 15, 2023, https://breakingdefense.com/2023/11/biden-launches-ai-risk-and-safety-talks-with-china-is-nuclear-c2-a-likely-focus/.; Jarrett Renshaw and Trevor Hunnicutt, "Biden, Xi Agree That Humans, Not AI, Should Control Nuclear Arms," World, *Reuters*, November 17, 2024, https://www.reuters.com/world/biden-xi-agreed-that-humans-not-ai-should-control-nuclear-weapons-white-house-2024-11-16/.

[91] International Committee of the Red Cross, "ICRC Position on Autonomous Weapon Systems," May 12, 2021, https://www.icrc.org/sites/default/files/document_new/file_list/icrc_position_on_autonomous_weapon_systems.pdf.

[92] High-level Advisory Body on Artificial Intelligence, *Governing AI for Humanity: Final Report* (United Nations, 2024), 74. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

existential risks to humanity within the near future."[93] Finally, the 2025 AI and Democratic Values Report, assessing AI policies and practices across 80 states, called for banning AI systems that undermine human rights and democratic principles.[94]

3.1.2 Risks and Opportunities of AI in the Context of Peace and Security

AI, international stability, and escalation dynamics. The integration of AI into the military domain has the potential to reshape how militaries organize themselves, the concepts and doctrine through which military power is developed and exercised, and how armed forces are commanded.[95] Yet, militaries around the world will vary widely in their capacity to develop and use AI, depending on access to resources, technical expertise, organizational and strategic culture, and geopolitical positioning.[96] These disparities may create new power asymmetries, with AI disrupting established ways of generating military power in both predictable ways, such as supporting the growing fielding of uncrewed and autonomous systems, and unpredictable ways.[97]

Developing AI systems that convincingly demonstrate a nation's strength and readiness could encourage restraint among some actors, but it may also contribute to new security vulnerabilities.[98] On the one hand, the use of AI-enabled systems for C3 may enhance early warning capabilities, providing decision-makers with additional time to make judgments on how to respond to potential attacks. On the other hand, compressed decision timelines and obscured intent through the use of AI raise concerns about misjudged adversarial signals, miscalculation, or inadvertent escalation.[99]

[93] "IDAIS-Beijing, 2024," paper presented at International Dialogues on AI Safety, Beijing, China, *International Dialogues on AI Safety*, IDAIS, March 10, 2024, https://idais.ai/dialogue/idais-beijing/.; Eleanor Olcott and Cristina Criddle, "Chinese and Western Scientists Identify 'Red Lines' on AI Risks," *Financial Times*, March 18, 2024, https://www.ft.com/content/375f4e2d-1f72-49c8-b212-0ab2a173b8cb.

[94] Center for AI and Digital policy, *ARTIFICIAL INTELLIGENCE AND DEMOCRATIC VALUES 2025: A Comprehensive Review of AI Policies and Practices Worldwide* (Center for AI and Digital policy, 2025), https://www.caidp.org/reports/aidv-2025/.

[95] Denise Garcia, "Future Arms, Technologies, and International Law: Preventive Security Governance," *European Journal of International Security* 1, no. 1 (2016): 94–111, https://doi.org/10.1017/eis.2015.7.; Kenneth Payne, *I, Warbot: The Dawn of Artificially Intelligent Conflict* (Oxford University Press, 2021), https://doi.org/10.1093/oso/9780197611692.001.0001.

[96] Stuart J. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (Penguin Books, 2020).; Jeffrey S. Lantis, "Strategic Culture and National Security Policy," *International Studies Review* 4, no. 3 (2002): 87–113, https://doi.org/10.1111/1521-9488.t01-1-00266.;Toni Erskine and Steven E. Miller, "AI and the Decision to Go to War: Future Risks and Opportunities," *Australian Journal of International Affairs* 78, no. 2 (2024): 135–47, https://doi.org/10.1080/10357718.2024.2349598.

[97] Rebecca Crootof, "War Torts: Accountability for Autonomous Weapons.," University of Pennsylvania Law Review 165, no. 6 (2016): 1347–402.

[98] Afina and Paoli, *Governance of Artificial Intelligence in the Military Domain*; Benjamin Jensen et al., *Algorithmic Stability: How AI Could Shape the Future of Deterrence* (Center for Strategic & International Studies, 2024), https://csis-website-prod.s3.amazonaws.com/s3fs-public/2024-06/240610_Jensen_Algorithmic_Stability.pdf?VersionId=iBZYLyt7ukgqBWTbbdBFBzqP63xvfrgz.

[99] Model dysfunction and performance limits, unpredictable and emergent outputs, hidden and systemic bias in data and outputs, brittleness, poor data integrity, and representativeness, opacity and lack of explainability, as well as vulnerability to external interference outlined under 2.2 have an impact on peace and security.

Spotlight Box 2: AI and strategic stability

Nuclear stability rests on shared assumptions of mutually assured destruction and, arguably, about a nuclear taboo that demarcates conventional warfare from nuclear warfare. Potentially, AI could reinforce strategic stability if faster data processing increases early warning time, providing decision-makers with more time to make well-considered decisions in a crisis.[100] However, any introduction of AI into nuclear decision-making also introduces a new set of uncertainties.[101] The integration of AI in non-redundant nuclear command, control, and communications (NC3) systems structures could increase the potential vulnerability of these systems to external manipulation (e.g., spoofing attempts) if they are connected to broader networks.[102] If AI-enabled capabilities increase the speed of operations, it could lead to first-strike instability dynamics because nuclear powers fear losing their nuclear arsenals in a rapid, AI-enabled strike.[103] Current AI capabilities seem unlikely to generate radical transparency on the battlefield.[104] Nevertheless, there is fear that Artificial General Intelligence (AGI) could lead to technological breakthroughs, particularly in ocean surveillance, contributing to transparency and the ability to rapidly target in a way that places nuclear second strike capabilities at risk.[105]

Especially early in the era of military AI, the uncertainty surrounding both adversaries' AI capabilities and their intentions can lead to fear and increase pressure for escalation.[106] Addressing these uncertainties could enable uses of AI that provide both clarity to decision-makers, which can increase stability, and enhance confidence-building measures (CBM). For instance, AI tools used in ISR contexts could feed into information used to improve situational awareness and reduce the space for misunderstandings by providing timely notifications about activities between states, translating messages, and ensuring access to the same information, decreasing chances of misinterpretation.[107]

AI and weapons proliferation. The general-purpose nature and dual-use potential of AI technologies, especially in light of the availability of powerful open-source models, raise questions about whether state and non-state actors could use AI to further the proliferation of military technologies. AI technologies developed in the civilian domain have found applications in the military domain. The accessibility and transferability of AI models, software, and data complicates efforts to monitor and control their spread.[108] States and non-state actors already use AI

[100] Vincent Boulanin et al., "Artificial Intelligence, Strategic Stability, and Nuclear Risk," Stockholm International Peace Research Institute, June 18, 2020, https://policycommons.net/artifacts/2324586/artificial-intelligence-strategic-stability-and-nuclear-risk/3085114/.

[101] Matthijs Maas et al., "10. Military Artificial Intelligence as a Contributor to Global Catastrophic Risk," in *The Era of Global Risk: An Introduction to Existential Risk Studies*, ed. SJ Beard et al. (Open Book Publishers, 2023), https://www.openbookpublishers.com/books/10.11647/obp.0336/chapters/10.11647/obp.0336.10.

[102] James S. Johnson, "Artificial Intelligence: A Threat to Strategic Stability," *Strategic Studies Quarterly* 14, no. 1 (2020): 16–39.

[103] Raluca Csernatoni "Governing Military AI Amid a Geopolitical Minefield" *Carnegie Endowment for International Peace (2024).*

[104] Steve Fetter and Jaganath Sankaran, "Emerging Technologies and Challenges to Nuclear Stability," *Journal of Strategic Studies* 48, no. 2 (2025): 252–96, https://doi.org/10.1080/01402390.2024.2433766.; Charles L. Glaser, "The End of MAD? Technological Innovation and the Future of Nuclear Retaliatory Capabilities," *Journal of Strategic Studies* 48, no. 2 (2025): 239–51, https://doi.org/10.1080/01402390.2024.2428983.

[105] Kenneth Payne, "Artificial Intelligence: A Revolution in Strategic Affairs?," *Survival* 60, no. 5 (2018): 7–32, https://doi.org/10.1080/00396338.2018.1518374.

[106] Riley Simmons-Edler et al., "AI-Powered Autonomous Weapons Risk Geopolitical Instability and Threaten AI Research," arXiv:2405.01859, preprint, arXiv, 2024, https://doi.org/10.48550/arXiv.2405.01859.

[107] Steven M. Williamson and Victor Prybutok, "The Era of Artificial Intelligence Deception: Unraveling the Complexities of False Realities and Emerging Threats of Misinformation," *Information* 15, no. 6 (2024): 299, https://doi.org/10.3390/info15060299.

[108] Maximilian Hoell and Sylvia Mishra, "Artificial Intelligence in Nuclear Command, Control, and Communications: Implications for the Nuclear Non-Proliferation Treaty," in *The Implications of Emerging Technologies in the Euro-Atlantic Space: Views from the Younger Generation Leaders Network*, ed. Julia Berghofer et al. (Palgrave Macmillan Cham, 2023), https://doi.org/10.1007/978-3-031-24673-9_8.

and commercial manufacturing to create inexpensive, precision-guided one-way attack drones on battlefields around the world, from Ukraine to Gaza.[109]

At the same time, AI systems could potentially lower the technical barriers for states and non-state actors to develop various weapons, including weapons of mass destruction (WMD), particularly biological or chemical weapons.[110] Such a scenario would be enabled via access to larger, sensitive datasets and automating analysis processes in condensed timeframes. While current AI capabilities may not yet offer actors substantial advances relative to existing internet search technology, their diffusion could make this threat increasingly more difficult to contain.[111] The accessibility of data on weapons systems, both to state and non-state actors, also coincides with concerns about the emergence of technologies that could generate effects on the scale of WMD, such as through the use of AI-enabled drone swarms.[112]

AI could serve as a tool to mitigate some concerns by aiding arms control and verification efforts. AI systems could facilitate real-time monitoring and compliance with international agreements, addressing long-standing challenges of accurately assessing weapon stockpiles or ensuring disarmament commitments are upheld. By enabling continuous, automated inspections, AI could provide more reliable evidence of adherence to arms control measures.[113]

AI and dependencies. As the integration of AI becomes increasingly central in national defense strategies, unequal access to data, resources, and talent can deepen the divide between AI 'haves' and 'have-nots'. This inequality risks entrenching asymmetries in international relationships, whether through formal alliances, defense exports, or informal connections. In alliances, it is unclear whether AI will ease or complicate interoperability, especially if states use different platforms and weapon systems of varying sophistication and capability. Between states, new forms of AI-driven strategic dependencies, clientelism, or even digital neo-imperialism could undermine trust in existing international relationships, including international governance.[114]

In parallel, the role of private actors in military AI, given that private sector investment and activity are the predominant drivers of AI innovation, raises questions about the erosion of traditional state authority and sovereignty.[115] Private firms increasingly control critical components of AI development, including large-scale datasets, advanced models, and cloud infrastructure. The role of private actors in the AI ecosystem may change established patterns of

[109] Stuart Russell, "AI Weapons: Russia's War in Ukraine Shows Why the World Must Enact a Ban," *Nature* 614 (2023): 620–23, https://doi.org/10.1038/d41586-023-00511-5. Michael C. Horowitz, "Battles of Precise Mass." *Foreign Affairs*, October 2024, https://www.foreignaffairs.com/world/battles-precise-mass-technology-war-horowitz.

[110] James Johnson, "Artificial Intelligence, Drone Swarming and Escalation Risks in Future Warfare," *The RUSI Journal* 165, no. 2 (2020): 26–36, https://doi.org/10.1080/03071847.2020.1752026.

[111] Doug Irving, *Artificial Intelligence and Biotechnology: Risks and Opportunities* (Rand, 2024), https://www.rand.org/pubs/articles/2024/artificial-intelligence-and-biotechnology-risks-and.html; Open AI, "Preparing for Future AI Capabilities in Biology," August 7, 2025, https://openai.com/index/preparing-for-future-ai-capabilities-in-biology/; Bill Drexel and Caleb Withers, "AI and the Evolution of Biological National Security Risks," (CNAS, August 13 2024), https://www.cnas.org/publications/reports/ai-and-the-evolution-of-biological-national-security-risks.

[112] United Nations Security Council Resolution 1540 (2004) obligates states to prevent the proliferation of all weapons of mass destruction, especially by non-state actors. The resolution affirms that the proliferation of WMDs "constitutes a threat to international peace and security." United Nations Security Council, Resolution 1540, U.N. Doc. S/RES/1540 (28 April 2004), preamble, para. 2.; Jimena Sofía Viveros Álvarez, "Symposium on Military AI and the Law of Armed Conflict: Drone Swarms as Weapons of Mass Destruction," *Opinio Juris*, April 5, 2024, https://opiniojuris.org/2024/04/05/symposium-on-military-ai-and-the-law-of-armed-conflict-drone-swarms-as-weapons-of-mass-destruction/.

[113] Michael C. Horowitz, Lauren Kahn, and Casey Mahoney. "The Future of Military Applications of Artificial Intelligence: A Role for Confidence Building Measures." *Orbis*, 64:4, 2020, pp. 528-543. https://www.sciencedirect.com/science/article/abs/pii/S0030438720300430.

[114] Garcia, *The AI Military Race.*

[115] Hazrat Usman et al., "The Future of State Sovereignty in the Age of Artificial Intelligence," *Journal of Law & Social Studies* 5, no. 2 (2023): 142–52, https://doi.org/10.52279/jlss.05.02.142152.

public–private sector relations by increasing the strategic importance of private actors relative to AI-enabled capabilities.

Finally, there are concerns that the pursuit of advanced AI, with the goal of securing decisive strategic advantage over adversaries, could result in highly capable systems that are insufficiently 'aligned' with their human principal's values.[116] Such scenarios could create significant threats to human society and welfare.[117]

Spotlight Box 3: Artificial General Intelligence and existential risk[118]

Discussion on the potential advent of AGI has surfaced important concerns about existential risks. Central to these concerns is the possibility that machines could surpass human control, leading to scenarios in which autonomous systems act in ways that subjugate or eliminate human agency, or even humanity itself.[119] While such scenarios remain speculative, they reflect deeper anxieties about losing oversight over powerful, adaptive systems. Only a handful of states have the ability to participate in a race for AGI, given the prohibitive hardware, energy, and expert costs. Regardless of the attainment of AGI or AGI-like capabilities, the inclusion of AGI in ongoing military strategic debates, and the uncertainty about its geopolitical implications, are themselves impacting stability. Within the military domain, this reflects broader anxieties about the disruptive potential of AI technologies.

3.2 The Context of Armed Conflict

The integration of AI in the military domain also presents unique risks and opportunities within the context of armed conflict. Here, "armed conflict" must be understood as situations that meet certain thresholds under IHL (*jus in bello*). The leading authority on this, the 1994 *Tadic* appeals decision of the International Criminal Tribunal for the Former Yugoslavia, outlines that "an armed conflict exists whenever there is a resort to armed force between States or protracted armed violence between governmental authority and organized armed groups or between such groups within a State".[120]

[116] Leopold Aschenbrenner, "SITUATIONAL AWARENESS: The Decade Ahead," Situational Awareness – The Decade Ahead, June 2024, https://situational-awareness.ai/. However, for a critique of the inevitability of an arms race, see also: Seán Ó hÉigeartaigh, *The Most Dangerous Fiction: The Rhetoric and Reality of the AI Race*, SSRN Paper no. 5278644 (Social Science Research Network, 2025), https://papers.ssrn.com/abstract=5278644; Richard Ngo et al., "The Alignment Problem from a Deep Learning Perspective," arXiv:2209.00626, preprint, arXiv, May 4, 2025, https://doi.org/10.48550/arXiv.2209.00626; Yoshua Bengio et al., "Managing Extreme AI Risks amid Rapid Progress," *Science* 384, no. 6698 (2024): 842–45, https://doi.org/10.1126/science.adn0117.

[117] Benjamin Boudreaux, *Mutual Dependence and Vulnerability in Human and AI Futures*, Working Paper WR-A4088-1 (RAND Corporation, 2025), https://www.rand.org/pubs/working_papers/WRA4088-1.html.

[118] Leonie Koessler and Jonas Schuett, "Risk Assessment at AGI Companies: A Review of Popular Risk Assessment Techniques from Other Safety-Critical Industries," arXiv:2307.08823, preprint, arXiv, 2023, https://doi.org/10.48550/arXiv.2307.08823.; Barry Pavel et al., *How Artificial General Intelligence Could Affect the Rise and Fall of Nations: Visions for Potential AGI Futures* (RAND Corporation, 2025), https://www.rand.org/pubs/research_reports/RRA3034-2.html.

[119] Responsible AI principles are essential for addressing these risks. For example, the Termination Obligation in the Universal Guidelines for AI asserts that AI systems must remain subject to human control. If meaningful control can no longer be exercised, the system should be promptly deactivated or terminated. Cent. AI Digit. Policy, "Universal Guidelines for AI."

[120] The term *jus in bello* is often used interchangeably with IHL or the "law of armed conflict" (LOAC). IHL applies only to situations of armed conflict and occupation. International Human Rights Law applies at all times. *Decision on the Defence Motion for interlocutory appeal on jurisdiction, The Prosecutor v. Dusko Tadic a/k/a "Dule"*, Case no. IT-94-1-AR72, para. 70.; International Lawyers Association, Committee on the Use of Force, *Report on the Meaning of Armed Conflict under International Law* (The Hague, 2010), http://www.ila-hq.org/en/committees/index.cfm/cid/1022. The report also describes "hostilities" as the kinetic exchange of attacks or fighting that is the *sine qua non* of armed conflict.

3.2.1 Applicable International Law and Regulatory Frameworks in the Context of Armed Conflict

International Humanitarian Law

While AI may create operational and tactical opportunities, not all such opportunities in armed conflict are ethically or legally permissible. IHL exists to place limits on the means and methods of warfare, as well as to protect people affected by armed conflict. The law serves both as a constraint and as a guide: it defines which opportunities may be pursued and which risks must be anticipated, mitigated, or altogether avoided.

This report's position aligns with the May 2025 CCW GGE on LAWS meeting text, echoing that "IHL applies to armed conflict, whether international or non-international, and governs the use of all weapons, means and methods of warfare, those of the past, those of the present and those of the future, and is, consequently, applicable regardless of the technology used."[121]

IHL provides that the means and methods of warfare are not unlimited.[122] While technological advancements may present new opportunities, the decisive factor in using emerging technologies remains the user's ability to ensure compliance with IHL and other international legal obligations.[123] The obligations of legal reviews flowing from Article 36 of Additional Protocol I (AP I) aim to ensure that weapons are assessed for compliance with international law.[124] This is reflected in the approach adopted by the CCW GGE LAWS rolling text which puts forth that states should "conduct legal reviews to determine, in accordance with their obligations under international law, in the study, development, acquisition, or adoption of LAWS, whether [their] deployment would, in some or all circumstances, be prohibited by international law."[125] Accordingly, these guidelines apply to AI in the military more broadly in situations where the systems in question constitute means and methods of warfare as referenced by Article 36.

IHL provides obligations **prohibiting certain weapons, means or methods of warfare** and restrictions on the use of others, where the weapons in question: (1) cause ***unnecessary suffering*** *or* ***superfluous injury***,[126] (2) are ***indiscriminate*** by nature, where they cannot be directed at a specific military target or whose effects cannot be limited, or (3) cause ***widespread, long-term, and severe environmental damage.***[127] Like all military capabilities, any AI systems that would violate any one or more of the above would be outlawed on the grounds that IHL seeks to prevent the use of weapons that are unlawful in all circumstances and to restrict those whose legality depends on the context of use.[128]

[121] Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, "Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects", (Revised as of May 12 2025) , https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2025)/CCW_GGE_LAWS_-_Revised_rolling_text_as_of_12_May_2025.pdf, p. 1 (Box II.2).

[122] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), 1125 U.N.T.S. 3 § 35(1) (1977). https://ihl-databases.icrc.org/ihl/INTRO/470.

[123] Mary Ellen O'Connell, Banning Autonomous Weapons: A Legal and Ethical Mandate, 37(3) Ethics & Int'l Affairs 287 (2023). Thompson Chengeta, "Measuring Autonomous Weapon Systems Against International Humanitarian Law Rules," *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.2755184.

[124] Additional Protocol I to the Geneva Conventions of 1949, Article 36; Andrew Clapham and Paola Gaeta, *The Oxford Handbook of International Humanitarian Law* (Oxford: Oxford University Press, 2022)

[125] GGE LAWS, Convention on Certain Conventional Weapons (Revised May 12, 2025), 3 (Box IV.1).

[126] Hague Regulations of 1907, Article 23; Additional Protocol I to the Geneva Conventions of 1949, Article 35; Andrew Clapham and Paola Gaeta, *The Oxford Handbook of International Humanitarian Law* (Oxford: Oxford University Press, 2022)

[127] Dan Hendrycks et al., "Superintelligence Strategy: Expert Version," arXiv:2503.05628, preprint, arXiv, April 14, 2025, 22, https://doi.org/10.48550/arXiv.2503.05628.

[128] International Committee of the Red Cross, "A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977 – International Committee of the Red Cross Geneva, January 2006," *International Review of the Red Cross* 88, no. 864 (2006): 931–56, https://doi.org/10.1017/s1816383107000938.

Simultaneously, the core IHL principles of *distinction*, *necessity*, *proportionality*, and *humanity* apply to all participants in armed conflict.[129]

Spotlight Box 4: Principles of international humanitarian law

1. **Distinction** – mandates that at all times parties to an armed conflict must distinguish between combatants and civilians as well as military and civilian objects. [130] It is never lawful to intentionally target civilians or other protected persons. Indiscriminate attacks are equally prohibited in terms of this rule. Employing a method or means of warfare which cannot be directed at a specific military objective is prohibited in all circumstances.[131]
2. **Necessity** – dictates that a party to an armed conflict may only employ means and methods required to achieve the conflict's legitimate aim, namely, to weaken the enemy's military forces.[132] Even if necessity for an attack exists, the principles of distinction, proportionality, and humanity restrict actions permitted under this principle.
3. **Proportionality** – requires "those who plan or decide upon an attack" to take measures to prohibit attacks "which may be expected to cause incidental loss of civilian life, injury to civilians, or damage to civilian objects or a combination of these that is excessive in relation to the concrete and direct military advantage anticipated."[133]
4. **Humanity** – imposes specific limits on the means and methods of warfare and requires that all persons who fall into enemy hands are always treated humanely. Its overarching goal is to reduce suffering, injury, and destruction during armed conflict by protecting life and health and upholding respect for human dignity. The principle is often referred to through the Martens Clause, which indicates that "the inhabitants and the belligerents remain under the protection and the rule of the law of nations, as they result from the usages established among civilized peoples, from the laws of humanity and the dictates of public conscience."[134]

The development and use of AI in the military domain should be assessed based on when or whether such use cases comply with the principles and rules outlined above. Furthermore, the adherence to these above principles is guided by rules related to the duty of constant care and **precautions** in attack, which necessitates, in the conduct of military operations, that parties take "constant care to spare the civilian population, civilians and civilian objects. All feasible precautions must be taken to avoid, and in any event to minimize, incidental loss of civilian life, injury to civilians and damage to civilian objects."[135]

[129] International Committee of the Red Cross, "Fundamental Principles of IHL," *How does law protect in war?*, https://casebook.icrc.org/a_to_z/glossary/fundamental-principles-ihl.

[130] This principle forms part of customary IHL, applicable in international as well as non-international armed conflicts. Jean-Marie Henckaerts and Louise Doswald-Beck, "Chapter 1: Distinctions between Civilians and Combatants," in *Customary International Humanitarian Law. Vol. 1: Rules*, Repr. with corr, ed. Carolin Alvermann, with Internationales Komitee vom Roten Kreuz, Customary International Humanitarian Law (Cambridge Univ. Press, 2009), https://www.icrc.org/sites/default/files/external/doc/en/assets/files/other/customary-international-humanitarian-law-i-icrc-eng.pdf. The ICJ has referred to the principle of distinction as a 'cardinal' and 'intransgressible' principle forming part of the 'fabric' of IHL. ICJ, Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996, § 78–79.

[131] Additional Protocol I to the Geneva Conventions of 1949, Article 58. See also the latest version of the p,2 (Box III.3).

[132] ICRC, The Principles of Humanity and Necessity, 1, https://www.icrc.org/sites/default/files/wysiwyg/war-and-law/02_humanity_and_necessity-0.pdf. See also, CCW GGE LAWS Rolling Text, 12 May 2025, p,2 (Box III.2)

[133] Additional Protocol I to the Geneva Convention of 1949, Article 1(5)(b) and Article 57(2)(a)(iii).

[134] International Committee of the Red Cross, "Humanity," *How Does Law Protect in War?,* accessed August 15, 2025, https://casebook.icrc.org/a_to_z/glossary/humanity. See, e.g., the latest Rolling Text of the CCW GGE LAWS (Box II.4) : "the civilian population and the combatants at all times remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience."

[135] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), 1125 U.N.T.S. 3 § 57(1) and (2)(a)(iii) and 57(2) (1977), https://ihl-databases.icrc.org/ihl/INTRO/470.; ICRC, Rule 15 Customary International Humanitarian Law Study, https://ihl-databases.icrc.org/en/customary-ihl/v1/rule15; Article 57, API; See also Geoffrey S. Corn and Tyler R. Smotherman, "Improving Compliance with International Humanitarian Law in an Era of Maneuver War and Mission Command," *SMU Law Review* 78, no. 1 (2025): 3–39, https://doi.org/10.25172/smulr.78.1.2.

Procedural obligations: due diligence

States' obligations to take reasonable precautions to ensure that military operations (with or without AI) respect the above principles are further grounded in a range of procedural obligations applicable both during and prior to armed conflict. The principle of due diligence governs the relationship between substantive and procedural principles. It is often understood as an obligation of conduct (e.g. establishing domestic procedures to determine whether a means or method of warfare would be prohibited), but may also entail obligations of result (e.g. realizing specific mandates). These positive obligations include duties to prevent, stop, or redress harms.[136] Importantly, the kind of state action required under due diligence evolves as risks change in light of new circumstances. This requires states to incorporate best practices, technical guidelines, and expert consensus, as they emerge. Accordingly, states are obligated to develop and implement robust TEVV procedures to assess the performance, reliability, and legal compliance of systems. The due diligence principle does not alter substantive legal rules but supports compliance by requiring efforts to meet applicable standards.

Other important norms and initiatives applicable to armed conflict

Both states and non-state actors have conducted work to set guardrails around certain uses of AI by articulating standards and clarifying best practices.[137] For instance, the IEEE-SA has undertaken work on framework documents to ground human decision-making throughout the lifecycle of AI systems used in defense applications.[138] In other domains, organizations such as the Centre for Humanitarian Dialogue have articulated a 2021 'Code of Conduct on Artificial Intelligence in Military Systems'. The Code of Conduct emphasizes embedding IHL in system development and the centrality of human responsibility for the use of force. They also advise against any form of AI control over nuclear weapons, reinforce the importance of systematic testing of military AI tools, and appropriate training of commanders and end-users.[139]

3.2.2 Risks and Opportunities of AI in the Context of Armed Conflict

The following analysis is structured with reference to three key actors in armed conflict: states and parties to the conflict, military commanders and combatants, and civilians and protected persons.

For parties to armed conflict, the integration of AI technologies into the military domain presents significant safety, ethical, legal, and operational risks, requiring assessments as all military systems do. A central concern lies in the potential erosion of human judgment in decisions around the use of force.[140] If AI tools are used to replace, rather than support, human moral and legal reasoning, especially in opaque and non-democratic contexts, this could lead to the false

[136] Antonio Coco and Talita Dias, "Chapter 12: 'Handle with Care': Due Diligence Obligations in the Employment of AI Technologies," in *Research Handbook on Warfare and Artificial Intelligence*, ed. Robin Geiß and Henning Lahmann (EE Elgaronline, 2024), https://www.elgaronline.com/edcollchap/book/9781800377400/book-part-9781800377400-19.xml; Simon Chesterman, "Weapons of Mass Disruption: Artificial Intelligence and International Law" Cambridge International Law Journal, *Cambridge International Law Journal* 10, no. 2 (2021): 181–203, https://doi.org/10.4337/cilj.2021.02.02.

[137] Ashley S Deeks, "The Hurdles to International Regulation," in *The Double Black Box: National Security, Artificial Intelligence, and the Struggle for Democratic Accountability*, ed. Ashley S Deeks (Oxford University Press, 2025), https://doi.org/10.1093/9780197520932.003.0009.

[138] Allik et al., *White Paper: A Framework for Human Decision-Making through the Lifecycle of Autonomous and Intelligent Systems in Defense Applications*; See also Atif Ali et al., "Development and Use of Artificial Intelligence in the Defense Sector," paper presented at 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, (IEEE, 2023)https://doi.org/10.1109/ICBATS57792.2023.10111113.

[139] Center for Humanitarian Dialogue, Code of Conduct on Artificial Intelligence in Military Systems (2021), https://www.hdcentre.org/wp-content/uploads/2021/08/AI-Code-of-Conduct.pdf.

[140] Taylor Kate Woodcock, "Human/Machine(-Learning) Interactions, Human Agency and the International Humanitarian Law Proportionality Standard," *Global Society* 38, no. 1 (2024): 100–121, https://doi.org/10.1080/13600826.2023.2267592.

understanding that humans can transfer accountability for critical judgments to AI systems.[141] The increasing automation of targeting, information gathering, and decision-making processes could risk detaching warfare from human ethical and legal deliberation.[142] There are concerns about delegating life and death decisions to machines, which may be viewed as an affront to human dignity.[143] AI systems lack the capacity to interpret intent, assess context, or execute precaution. Furthermore, it has been established that autonomous weapon systems have a "war bias,"[144] unless paired with high-confidence oversight and transparent system design.[145] When elements of ethical and legal compliance are considered throughout the lifecycle of AI systems, an approach referred to as compliance, responsibility, or "**lawful by design**",[146] there is potential for enhancing compliance with international law in military contexts.[147] However, legal considerations in the design alone are not sufficient to guarantee accuracy in the operation of AI systems, which depend on a range of technical, operational, and contextual factors.

Nonetheless, AI tools could also offer potential benefits for parties to armed conflict that increase military effectiveness and IHL compliance. These include improved resource allocation and force deployment, potentially allowing objectives to be met with fewer personnel and less direct exposure to harm.[148] AI systems may also assist commanders during the targeting cycle by supporting the assessment of expected collateral damage and feasible precautions. While AI-enabled capabilities may support legal compliance in theory, empirical evidence reveals repeated failures to meet IHL obligations in practice.[149] These include misclassification of civilians, failure to assess proportionality, and erosion of precautionary principles, particularly in AI-assisted targeting contexts.[150]

[141] This report follows the CCW GGE LAWS in drawing an analogy between LAWS and military AI systems, emphasizing that "states, parties to armed conflict, and individuals must at all times remain responsible and accountable in accordance with international law for decisions with regard to [the use of military AI] since responsibility and accountability cannot be transferred to machines." GGE LAWS, May 2025 text, p. 4 (Box V.2); Neil Renic and Elke Schwarz, "Crimes of Dispassion: Autonomous Weapons and the Moral Challenge of Systematic Killing," *Ethics & International Affairs* 37, no. 3 (2023): 321–43, https://doi.org/10.1017/S0892679423000291; Lucy Suchman, "Algorithmic Warfare and the Reinvention of Accuracy," *Critical Studies on Security* 8, no. 2 (2020): 175–87, https://doi.org/10.1080/21624887.2020.1760587.

[142] Luke Moffett and Jessica Dorsey, "The Warification of International Humanitarian Law and the Artifice of Artificial Intelligence in Decision-Support Systems: Restoring Balance through the Legitimacy of Military Operations," in *Yearbook of International Humanitarian Law Volume 27,* (Springer, 2025), https://pure.qub.ac.uk/en/publications/the-warification-of-international-humanitarian-law-and-the-artifi.

[143] Christof Heyns, "Autonomous Weapons in Armed Conflict and the Right to a Dignified Life: An African Perspective," *South African Journal on Human Rights* 33, no. 1 (2017): 46–71, https://doi.org/10.1080/02587203.2017.1303903; Thompson Chengeta, "Dignity, Ubuntu, Humanity and Autonomous Weapon Systems (AWS) Debate: An African Perspective," *Revista de Direito Internacional* 13, no. 2 (2016), https://doi.org/10.5102/rdi.v13i2.4135.

[144] Juan-Pablo Rivera et al., "Escalation Risks from Language Models in Military and Diplomatic Decision-Making," *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, ACM, June 3, 2024, 836–98, https://doi.org/10.1145/3630106.3658942.

[145] Center for AI and Digital Policy (CAIDP). "CAIDP Urges UN to Finalize Treaty to Ban Lethal Autonomous Weapons." LinkedIn, September 14, 2024. https://www.linkedin.com/posts/center-for-ai-and-digital-policy_caidp-statement-un-ai-and-law-sept-16-activity-7240703399728418816-me8D?utm_source=share&utm_medium=member_desktop&rcm=ACoAADJibBoBNCVkuzaKW-68gLoF0oM78r8pgA4Y.

[146] Yasmin Afina, "Combat Code Compliance: International Humanitarian Law and the Development Stages of AI for Targeting" (University of Essex, 2025), https://doi.org/10.5526/ERR-00041309.

[147] Article 36 Legal, "Lawful by Design Initiative – Bridging the Gap between Industry Led Innovation and National Legal Review Processes," Article 36 Legal, https://www.article36legal.com/lawful-by-design; Esmat Zaidan and Imad Antoine Ibrahim, "AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective," *Humanities and Social Sciences Communications* 11 (2024), https://doi.org/10.1057/s41599-024-03560-x.

[148] Avi Goldfarb and Jon R. Lindsay, "Prediction and Judgment: Why Artificial Intelligence Increases the Importance of Humans in War," *International Security* 46, no. 3 (2022): 7–50, https://doi.org/10.1162/isec_a_00425.

[149] Jimena Sofía Viveros Álvarez, "Autonomous Weapons Systems: The Accountability Conundrum," *in Opiniones Técnicas sobre Temas de Relevancia Nacional* 41 (Ciudad de México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, 2021), 47–107, ISBN 978-607-30-5219-1, https://archivos.juridicas.unam.mx/www/bjv/libros/13/6467/7.pdf.

[150] Center for AI and Digital policy, *ARTIFICIAL INTELLIGENCE AND DEMOCRATIC VALUES 2025: A Comprehensive Review of AI Policies and Practices Worldwide*.

For commanders and combatants, AI systems can produce unpredictable or erroneous outputs that may compromise mission success and endanger the lives of personnel.[151] As with all military systems, these risks are magnified when AI tools are used by individuals lacking sufficient training or when traditional training fails to prepare operators for AI-specific challenges.[152] Operational stress, particularly in fast-paced or ambiguous scenarios, may pressure end-users to over-rely on AI-generated content. In turn, overreliance on AI-enabled systems to ensure legal compliance may further increase the risk of automation bias, sidelining human judgement, and reducing the likelihood of lawful military conduct.[153] This reliance can result in automation bias and cognitive offloading, among other biases, where human judgment is deferred to machine outputs without critical engagement.[154] This dynamic is especially dangerous in AI-DSS, where human agency may be undermined precisely in the moments where it is needed most. On the battlefield, this could manifest as ineffective and inefficient resource allocation, friendly fire incidents, or the misidentification of civilian targets, leading to serious humanitarian and legal consequences.[155]

In certain circumstances, AI can assist commanders and combatants by handling routine or resource intensive tasks and enhancing situational awareness, particularly in environments with multiple data streams.[156] In limited contexts, AI systems offer decision-support for military operators through the ability to process and synthesize large volumes of information in significantly reduced timeframes. Examples here are tasks that center around the analysis and extraction of unstructured data, such as intercepted messages or radiocommunication.[157] Such systems could also contribute to operational planning or coordination by reducing latency between units, potentially lowering the risk of fratricide or miscommunication.[158] Outside of the battlefield, AI-enabled training and simulation tools could enhance combat preparedness by replicating complex conditions in a risk-free environment.[159] AI can also enhance cyber operations by automating processes such as vulnerability detection, intrusion response, and network monitoring, overall improving digital resilience and agility.[160]

[151] Dan Hendrycks et al., "An Overview of Catastrophic AI Risks," arXiv:2306.12001, preprint, arXiv, October 9, 2023, https://doi.org/10.48550/arXiv.2306.12001.

[152] James Johnson, "The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-Enabled Warfare," *Journal of Military Ethics* 21, nos. 3–4 (2022): 246–271, https://doi.org/10.1080/15027570.2023.2175887.

[153] Woodcock, "Human/Machine(-Learning) Interactions, Human Agency and the International Humanitarian Law Proportionality Standard."; Michael C Horowitz and Lauren Kahn, "Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts," *International Studies Quarterly* 68, no. 2 (2024), https://doi.org/10.1093/isq/sqae020.

[154] Laura Bruun and Marta Bo, Bias in Military Artificial Intelligence and International Humanitarian Law, SIPRI Report, August 2025, https://www.sipri.org/publications/2025/other-publications/bias-military-artificial-intelligence-and-compliance-international-humanitarian-law.

[155] Noel E. Sharkey, "The Evitability of Autonomous Robot Warfare," *International Review of the Red Cross* 94, no. 886 (2012): 787–799, https://doi.org/10.1017/s1816383112000732. See also, Larry Lewis et al., *Preparing for Civilian Harm Mitigation and Response in Large-Scale Combat Operations*, with Anna Williams 2024), https://www.cna.org/reports/2024/08/Preparing-for-Civilian-Harm-Mitigation-and-Response-in-Large-Scale-Combat-Operations_REV.pdf.

[156] Ronald C. Arkin, "The Case for Ethical Autonomy in Unmanned Systems," *Journal of Military Ethics* 9, no. 4 (2010): 332–41, https://doi.org/10.1080/15027570.2010.536402.

[157] Kateryna Bondar, *Ukraine's Future Vision and Current Capabilities for Waging AI-Enabled Autonomous Warfare* (Center for Strategic & International Studies Wadhwani AI Center, 2025), https://www.csis.org/analysis/ukraines-future-vision-and-current-capabilities-waging-ai-enabled-autonomous-warfare.

[158] Richard Farnell and Kira Coffey, AI's New Frontier in War Planning: How AI Agents Can Revolutionize Military Decision-Making, Belfer Center for Science and International Affairs, October 11, 2024, https://www.belfercenter.org/research-analysis/ais-new-frontier-war-planning-how-ai-agents-can-revolutionize-military-decision.

[159] James Johnson, "Artificial Intelligence & Future Warfare: Implications for International Security," *Defense & Security Analysis* 35, no.2 (2019):147–69, https://doi.org/10.1080/14751798.2019.1600800.

[160] Peter R. J. Trim and Yang-Im Lee, "Combining Sociocultural Intelligence with Artificial Intelligence to Increase Organizational Cyber Security Provision through Enhanced Resilience," *Big Data and Cognitive Computing* 6, no. 4 (2022): 110, https://doi.org/10.3390/bdcc6040110.

For civilians and protected persons, the risks posed by the integration of AI into armed conflict may be most acute, as a result of how AI systems may reduce or displace human judgment in life and death decisions, such as target identification.[161] Without clear constraints and human oversight, AI models may mistakenly select targets, be unable to correctly render information, or interpret protected behaviors like surrender or incapacitation as part of detention activities, especially in dynamic and complex environments like densely populated areas. The consequences of such errors can be catastrophic, leading to unlawful killings, or unintended and disproportionate harm. These failures may stem from limited or poor-quality training data, technical design flaws, or deployment in environments for which the system was not intended.[162] AI systems trained on biased or unrepresentative data may also perpetuate racial, gender, or other discriminatory patterns. This is particularly relevant regarding groups that were already vulnerable to marginalization and risks the unequal implementation of IHL protections.[163]

While there is currently little empirical evidence showing that the use of AI to identify targets during armed conflict reduces civilian harm, AI may have beneficial applications in humanitarian efforts, such as identifying safe corridors, addressing flawed collateral damage estimates, locating civilians, and delivering aid in inaccessible areas.[164] In these scenarios, AI could, if appropriately governed, contribute to reducing some of the harms associated with armed conflict.[165]

[161] Frank Sauer and Niklas Schörnig, “Killer Drones: The ‘Silver Bullet’ of Democratic Warfare?,” *Security Dialogue* 43, no. 4 (2012): 363–380, https://doi.org/10.1177/0967010612450207.

[162] Daniele Amoroso, Denise Garcia, and Guglielmo Tamburrini, “The Weapon That Mistook a School Bus for an Ostrich,” *Science & Diplomacy*, May 5, 2022 https://doi.org/10.1126/scidip.ade6750.

[163] Sharkey, “The Evitability of Autonomous Robot Warfare.”

[164] Larry Lewis and Andrew Illachinski, Leveraging AI to Mitigate Civilian Harm, CNA, February 2022, https://www.cna.org/reports/2022/02/Leveraging-AI-to-Mitigate-Civilian-Harm.pdf. Forrest E. Morgan et al., *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World*, nos. RR3139-1 (RAND Corporation, 20Ris20), https://www.rand.org/pubs/research_reports/RR3139-1.html.

[165] Wen Zhou and Anna Rosalie Greipl, “Artificial Intelligence in Military Decision-Making: Supporting Humans, Not Replacing Them,” *Humanitarian Law & Policy*, August 29, 2024, https://blogs.icrc.org/law-and-policy/2024/08/29/artificial-intelligence-in-military-decision-making-supporting-humans-not-replacing-them/.

Case Study: Artificial Intelligence-enabled Decision Support Systems [Part 2]

AI-DSS use can introduce risks and inefficacies, including biases, accuracy problems, complications around legal compliance, and quantification logics that risk reducing humans to data points. These systems can also reflect embedded sociotechnical imaginaries, creating cascading effects on legal compliance and reviews.[166]

Human judgment and decision-making, especially in military contexts, can be compromised by cognitive biases: systematic thinking errors rooted in the brain's tendency to simplify information.[167] These biases can distort perception, combining with alignment and representation issues, leading to poor decisions and non-compliance with legal obligations, and iteratively influencing the entire lifecycle of AI-DSS. As biases are often unconscious, their impact must be addressed throughout system design, not just during use.[168]

When users perceive AI-generated information as highly reliable, they may place undue trust in its outputs, ignoring contradictory evidence.[169] This over-reliance can result in "satisficing", accepting recommendations without further inquiry. For example, commanders may follow AI targeting suggestions without verifying the legality or consequences of the decision. Bias can also manifest through confirmation bias, where users interpret conflicting information as consistent with AI outputs or vice versa. Overall, reliance on AI can reduce vigilance and situational awareness, increasing the risk of errors or oversight in critical decisions through:

- **Automation bias** – the tendency of users to uncritically accept a system's recommendations, even when they conflict with human judgment, available evidence, or legal obligations. In military settings, this bias can distort decision-making by causing commanders to defer to AI outputs, increasing the risk of operational errors and non-compliance with IHL.
- **Cognitive offloading** – tempts users to transfer cognitive effort to machines, fostering dependence. AI cannot intuit, contextualize, or ethically reason like humans. Yet humans risk treating it as if it can. Implications can be deadly when this happens on a battlefield.[170]
- **Anchoring bias** – initial AI estimates (e.g. projected casualties) can skew later human judgments, even when new data contradicts earlier outputs. This presents a danger of overweighing machine-generated assumptions across phases of the joint targeting cycle.
- **Deskilling and cognitive erosion** – repeated reliance on AI may degrade human capacity for independent reasoning. Commanders lose the ability to, for example, conduct proportionality assessments unaided, undermining readiness and violating the principle of precaution in IHL.

[166] Moffett and Dorsey, "The Warification of International Humanitarian Law and the Artifice of Artificial Intelligence in Decision-Support Systems: Restoring Balance through the Legitimacy of Military Operations."; Ingvild Bode, *Emerging Norms around Military Applications of AI: The Case of Human Control*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Bode-2.pdf.

[167] Martie G. Haselton et al., "The Evolution of Cognitive Bias," in *The Handbook of Evolutionary Psychology* (John Wiley & Sons, Inc., 2005).

[168] Jessica Dorsey and Marta Bo, "AI-Enabled Decision-Support Systems in the Joint Targeting Cycle: Legal Challenges, Risks, and the Human(e) Dimension," *International Law Studies* n.106 (2025): 1–45; Laura Bruun and Marta Bo, *Bias in Military Artificial Intelligence and Compliance with International Humanitarian Law*, SIPRI Report (Stockholm International Peace Research Institute, August 2025), https://doi.org/10.55163/NLWV5347.

[169] Andrew M. Colman, "Confirmation bias *n*." in *A Dictionary of Psychology* (Oxford University Press, 2015), 93, https://www.oxfordreference.com/display/10.1093/acref/9780199657681.001.0001/acref-9780199657681.

[170] Jessica Dorsey, *Proportionality under Pressure: AI-Based Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Dorsey.pdf.

3.3 Beyond Armed Conflict

Designating a system as "military AI" or its development as a "weapon" does not guarantee that its use will be confined to armed conflict. Due to the general purpose nature of AI technologies, combined with the dual-use character of many specific AI capabilities, they may be deployed in a range of other settings outside the legal framework of armed conflict, such as within policing or the use of force domestically (excluding non-international armed conflict) and abroad in counter-piracy, terrorism, insurgency or organized crime operations.[171] These activities remain part of the military domain insofar as they involve dual-use technologies and/or military actors, capabilities, or objectives, and may raise similar ethical, legal, and operational concerns.

3.3.1 Risks and Opportunities in the Context of Human Rights

In contexts outside the parameters of armed conflict, international human rights law (IHRL) is the applicable legal regime for evaluating the lawfulness of AI use. Human rights serve as the normative anchor for assessing both risks and opportunities, and their protection remains a central objective.[172] This framework is also reflected in the desire of developing nations to use AI for peaceful means. Key rights that are most relevant for the discussion at hand are:

1. **Human dignity** – recognized both as the foundation of human rights and as a right in its own standing. This affirms the inherent worth of every individual as a bearer of rights regardless of their race, ethnicity, religion, gender, language, political opinion, national or social origin, or other status. It aims to protect the essential characteristics associated with what it is to be a living, human being and requires that all technological design, deployment, and use actively uphold and protect dignity as an enforceable, non-derogable right.

2. **Right to life** – asserts that all individuals are protected from unlawful killing and no one may be arbitrarily deprived of their life.[173]

3. **Equality and non-discrimination** – asserts that all individuals are to be treated with the same respect, irrespective of their background, free from discrimination.

4. **Right to liberty** – encompasses both negative liberty, freedom from undue interference or coercion, and positive liberty, the capacity to act autonomously and realize one's potential. Protecting freedom demands that technologies neither constrain autonomy nor create undue dependencies. Safeguarding freedom is essential to preserving human agency and enabling authentic, informed participation in social, political, and economic life.

The abovementioned uses of AI in the contexts of peace and security, as well as armed conflict, have direct implications for the protection of human rights, also outside of these parameters. Many of the risks posed by AI for the protection of human rights stem from the ways in which issues around development and the use of AI may obscure, bypass, or otherwise erode human

[171] Jimena Viveros, "Why Should the UN 'Govern AI for Humanity': What Is at Stake and What Is the Urgency?," *Opinio Juris*, November 27, 2024, https://opiniojuris.org/2024/11/27/why-should-the-un-govern-ai-for-humanity-what-is-at-stake-and-what-is-the-urgency/; High-level Advisory Body on artificial intelligence, *Governing AI for Humanity*; Vibhu Mishra, "Humanity's Fate Can't Be Left to Algorithms, UN Chief Tells Security Council | UN News," December 19, 2024, https://news.un.org/en/story/2024/12/1158376; Afina, *The Global Kaleidoscope of Military AI Governance*.

[172] Moreover, IHRL continues to apply during armed conflict, complementing IHL in offering protection to individuals.

[173] African Charter on Human and Peoples' Rights (Banjul Charter), § 4 (1982). https://www.oas.org/en/sla/dil/docs/African_Charter_Human_Peoples_Rights.pdf; Garcia, "Lethal Artificial Intelligence and Change." *International Studies Review* 20, no. 2 (2018): 334-341.

judgment and legal oversight.[174] Inappropriately delegating decisions, or excessive reliance on AI, may reduce individuals to data points, behavioral patterns, or algorithmic outputs, stripping away the moral and human recognition that observing the right to dignity demands.[175] States' obligation to ensure that there is no arbitrary deprivation of life applies equally to use by non-state actors, as states have an obligation to protect citizens from unlawful use by non-state actors. [176]

In peacetime security contexts, particularly when militaries interact with counterterrorism, border control, and law enforcement contexts, the deployment of AI-driven surveillance and targeting tools, such as facial recognition, biometric tracking, predictive analytics, and pattern detection, may result in the use of lethal force without sufficient legal or evidentiary basis and serious risks to the right to privacy.[177] Upholding the right to dignity means ensuring that AI technologies are developed and used in a way that treats all individuals as valuable, rights-bearing human beings. Surveillance driven by opaque AI systems can further marginalize already vulnerable groups, treating them as objects of suspicion rather than as rights-holders.[178]

AI systems could support certain aspects of human rights protection in limited contexts, with rigorous ethical and legal design. AI technologies for law enforcement monitoring could help identify arbitrary or unlawful use of lethal force, thereby protecting rights such as the right to liberty, freedom from discrimination, and the right to privacy.[179] AI-enabled tools could, in certain circumstances, enhance oversight, transparency, and review processes, enabling states to monitor compliance with ethical and legal standards and thus uphold procedural safeguards critical to the protection of human rights in security operations.[180] These capabilities align with the dual obligation to protect life both in war and peace.

3.3.2 Risks and Opportunities in the Context of Environmental Protection

Protection of the environment is recognized as an issue tied closely to human rights and security.[181] As with all military technologies, the development and use of AI in the military domain present both risks and opportunities for environmental protection, extending beyond the battlefield to the development and procurement stages. Relevant principles of international law, including those embedded in IHL, peacetime environmental treaties, and the UNESCO

[174] Lewis and Ilachinski, *Leveraging AI to Mitigate Civilian Harm* (2022); Zhou and Greipl, "Artificial Intelligence in Military Decision-Making: Supporting Humans, Not Replacing Them.", Humanitarian Law and Policy (2024); Center for AI and Digital policy, *CAIDP's Comments on Artificial Intelligence and Judicial Systems For the Special Rapporteur on the Independence of Judges and Lawyers, United Nations Office of the High Commissioner for Human Rights* (Center for AI and Digital Policy, 2025), https://www.linkedin.com/posts/center-for-ai-and-digital-policy_caidp-ai-and-the-judiciary-un-ohchr-may-activity-7324087622878830593-EBY_/.

[175] Christof Heyns and UN. Human Rights Council. Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, eds., "Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Christof Heyns," UN, 9, https://digitallibrary.un.org/record/755741; Renic and Schwarz, "Crimes of Dispassion: Autonomous Weapons and the Moral Challenge of Systematic Killing", *Ethics & International Affairs* 37, no. 3 (2023): 321-343.

[176] Thompson Chengeta, "The Right to Non-Discrimination, and Freedom from Racial Oppression Should Be Part of the Guidelines and Principles in the Discussion on AWS," 2023, https://committees.parliament.uk/writtenevidence/120290/pdf/.

[177] Mary O'Connell, "Data Privacy Rights: The Same in War and Peace," in *The Rights to Privacy and Data Protection in Times of Armed Conflict* (The North Atlantic Treaty Organization Cooperative Cyber Defence Centre of Excellence, 2022), https://scholarship.law.nd.edu/book_chapters/126.

[178] Sue Anne Teo, "Human Dignity and AI: Mapping the Contours and Utility of Human Dignity in Addressing Challenges Presented by AI," *Law, Innovation and Technology* 15, no. 1 (2023): 241–79, https://doi.org/10.1080/17579961.2023.2184132; *Ban Facial Surveillance Technology* (Center for AI and Digital Policy, 2022), https://www.caidp.org/statements/ban-facial-surveillance-technology/.

[179] Elizabeth D. Gibbons, "Toward a More Equal World: The Human Rights Approach to Extending the Benefits of Artificial Intelligence," *IEEE Technology and Society Magazine* 40, no. 1 (2021): 25–30, https://doi.org/10.1109/MTS.2021.3056295.

[180] Nehaluddin Ahmad et al., "The Challenges of Human Rights in the Era of Artificial Intelligence," *UUM Journal of Legal Studies* 16, no. 1 (2025): 150–69, https://doi.org/10.32890/uumjls2025.16.1.9.

[181] The North Atlantic Treaty Organization, "Environment, Climate Change and Security," NATO, 2024, https://www.nato.int/cps/en/natohq/topics_91048.htm.

Recommendation on the Ethics of AI, provide important guardrails.[182] The UN International Law Commission has explicitly acknowledged the role of armed conflict in causing environmental harm, underscoring the need for greater attention.[183]

Spotlight Box 5: The environmental impacts of AI

The development and use of AI, both in the civilian and military domains, impacts the environment. The initial stages of AI model development often demand substantial amounts of critical materials, computational resources, and infrastructure. The process of manufacturing advanced hardware, training large-scale models, and operating data centers can contribute to significant carbon emissions, increased energy use, and excessive resource consumption.[184] In turn, the growing demand for rare earth minerals has intensified extractive practices, often in regions where weak regulatory oversight is linked to deforestation, water pollution, biodiversity loss, displacement, and, in some cases, the aggravation of local conflicts.[185] AI systems, like other technologies, could have unforeseen or uneven environmental effects, and there is a potential risk that ecological considerations might not receive adequate attention unless explicitly integrated into planning and decision-making processes.

Where AI is used to optimize military logistics, resource management, and impact assessments pre, during, and post conflict, it could result in the reduction of negative effects for the environment. By limiting redundancies and enhancing operational efficiency, AI could help limit the environmental degradation typically associated with large-scale troop movements, equipment deployment, and supply chains.[186] AI systems could also support more precise assessments of potential environmental impacts before military operations, particularly in areas housing critical ecosystems, water sources, or agricultural zones, and risks resulting from collateral damage.[187]

3.4 Implications for Responsibility and Accountability

At the core of the challenges stemming from the integration of AI into the contexts of peace and security, armed conflict, and parameters beyond armed conflict, lie concerns around human agency, responsibility, and accountability. This section summarizes the above-mentioned challenges into three core issues: (1) attribution of responsibility and accountability, (2) preservation

[182] United Nations Educational, Scientific and Cultural Organization, “Recommendation on the Ethics of Artificial Intelligence,” 2022, https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

[183] United Nations International Law Commission, “Protection of the Environment in Relation to Armed Conflicts,” 2015, https://legal.un.org/ilc/reports/2015/english/chp9.pdf.

[184] Germán Darío Corzo-Ussa et al., “Military Artificial Intelligence Applied to Sustainable Development Projects: Sound Environmental Scenarios,” *DYNA* 90, no. 228 (2023): 115–22.

[185] Jon Woodhead and Mathieu Landry, “Harnessing the Power of Artificial Intelligence and Machine Learning in Mineral Exploration—Opportunities and Cautionary Notes,” *SEG Discovery*, no. 127 (October 2021): 19–31, https://doi.org/10.5382/Geo-and-Mining-13; Adebayo Okeowo, “The Human Rights Implications of Extracting Minerals and Personal Data from Africa for the Development of Military AI,” GC REAIM Expert Policy Notes, (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Okeowo.pdf.

[186] Marie Francisco, “Artificial Intelligence for Environmental Security: National, International, Human and Ecological Perspectives,” *Current Opinion in Environmental Sustainability* 61 (April 2023): 101250, https://doi.org/10.1016/j.cosust.2022.101250.

[187] Sahibpreet Singh and Manjit Singh, “Role of Artificial Intelligence for Environment Protection: An Analysis,” *ResearchGate*, ahead of print, May 6, 2025, 165–75, https://doi.org/10.2139/ssrn.5196899.

of human moral agency, and (3) preservation of human oversight. GC REAIM emphasizes that responsibility functions as an overarching normative concept with two interrelated dimensions:

a) Responsibility under international law: This encompasses the prospective duty to comply with both positive and negative legal obligations, whether by ensuring that certain actions are undertaken or by refraining from prohibited conduct, as well as the retrospective duty to hold actors accountable for actions or omissions in accordance with established rules of international law.

b) Responsibility beyond law: Is used (i) in a forward looking sense, as implying a moral obligation to see to it that something is done or established, and (ii) in a backward looking sense, as implying that someone (a party or actor) is the legitimate subject of blame or praise, has a moral obligation to compensate for damages (liability), or has a moral obligation to explain or justify relevant actions or omissions with potential legal or other consequences (accountability).

For the purposes of this report, and to deal with the legal and moral aspects of AI in the military domain, both legal compliance and moral acceptability, it is important to note that designing mechanisms, processes, and conditions for responsibility will most often draw upon legal sources, but are not limited to them.[188]

In traditional military operations, responsibility is delineated through a clear command structure among human agents. This structure remains in place with the introduction of AI-enabled systems. However, the integration of AI risks complicating this model by introducing systems that may operate without transparent logics or human judgment and oversight.[189] Concerns about the moral acceptability of military AI often center not only on IHL and other norms, but also on whether the human control, judgment, and oversight required by ethics and law can be meaningfully exercised in practice.

1. Attribution of responsibility and accountability

Given the multitude of actors engaged throughout the technological lifecycle of AI systems, as with any complex military system, responsibility and accountability for actions involving these systems should be conceptualized in terms of distributed, joint, or shared responsibility across individuals and groups.[190] The emphasis on obligations and responsibilities of humans across a system lifecycle, be it AI-enabled or not, is reflected in this report as well as the International Committee of the Red Cross (ICRC) submission to the UNSG on AI in the military domain.[191]

In IHL, command responsibility holds military commanders accountable for violations committed by those under their effective authority and control when they knew or should

[188] CCW GGE LAWS Rolling Text, 12 May 2025, p,2 (Box V); Asia-Pacific Institute for Law and Security, *Joint Statement on Box V. Statement* delivered during the Second Session of the 2025 GGE on LAWS, September 5, 2025, https://apils.org/2025/09/04/joint-statement-on-box-v/.

[189] Context-appropriate human judgment and control reflects the latest iteration of terms stemming from the conceptualization of "meaningful human control". Various conceptions have been proposed, alongside efforts for the term to be seen as a normative feature of the military decision-making process. Frank Sauer, "Lethal Autonomous Weapons Systems," in *The Routledge Social Science Handbook of AI* (Routledge, 2021); Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account," *Frontiers in Robotics and AI* 5 (February 2018): 15, https://doi.org/10.3389/frobt.2018.00015.

[190] Seumas Miller, *Collective Moral Responsibility, Institutionalisation and Lethal Autonomous Weapon Systems (LAWS)*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/04/Miller.pdf.

[191] International Committee of the Red Cross, "Submission to the United Nations Secretary General on Artificial Intelligence in the Military Domain," 2025, https://www.icrc.org/sites/default/files/2025-04/ICRC_Report_Submission_to_UNSG_on_AI_in_military_domain.pdf.

have known and failed to prevent or punish such acts.[192] This legal doctrine is grounded in a human-to-human relationship of authority and oversight, and must not be misapplied to the human-machine interaction in AI systems.[193]

Under international law, states bear legal responsibility for acts or omissions that constitute breaches of international obligations and are attributable to the state.[194] If an AI-enabled action originates from a state's apparatus or is conducted under its effective control, the state remains accountable, even if the specific decision was influenced or executed by an autonomous system.

Meanwhile, combatants and operators who use AI systems remain subject to individual criminal responsibility under International Criminal Law (ICL), including for war crimes, crimes against humanity, or other serious violations. The use of AI does not negate the legal requirement to establish a person's mental state in relation to a wrongful act. Clarifying the operator's legal obligations and ensuring that systems are designed to preserve informed human decision-making are essential to maintaining the integrity of individual accountability.

Finally, private sector actors such as companies are increasingly recognized as responsible actors under evolving norms of corporate due diligence and accountability. While international law does not impose direct liability on corporate entities, developers should ensure that responsibility by design principles are incorporated so that their technologies do not contribute to violations of human rights or IHL.[195]

2. Preservation of human moral agency

Human moral agency, the capacity to make contextual, evaluative judgments about "right and wrong while retaining responsibility for the actions and choices pursued"[196] remains central to the ethical and lawful use of AI systems in the military domain. Excluding critical human roles, such as in decisions about the use of force, would undermine the foundational premise that ethical military conduct requires human discretion. The application of *jus in bello* principles demands context-sensitive human moral reasoning. Replacing rather than supplementing human judgment threatens the integrity of military operations, weakens accountability frameworks, and risks delegating decisions requiring moral responsibility to systems incapable of such reasoning.

[192] Thompson Chengeta, "Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law," *Denver Journal of International Law & Policy* 45, no. 1 (2016), https://digitalcommons.du.edu/djilp/vol45/iss1/3.

[193] However, it is necessary to review how the introduction of AI affects relationships between humans, and whether elements of command responsibility need to be adjusted as a result. Mun-eon Park, "Autonomous Weapon Systems and Command Responsibility," *Seoul International Law Journal* 31, no. 2 (2024): 1–40, https://doi.org/10.18703/silj.2024.12.31.2.001.

[194] International Law Commission, *Articles on the Responsibility of States for Internationally Wrongful Acts*, UN Doc. A/56/49 (Vol. I)/Corr.4 (United Nations, 2002), https://legal.un.org/ilc/texts/instruments/english/draft_articles/9_6_2001.pdf.

[195] Daragh Murray, "Adapting a Human Rights-Based Framework to Inform Militaries' Artificial Intelligence Decision-Making Processes," *Saint Louis University Law Journal* 68, no. 2 (2024), https://scholarship.law.slu.edu/lj/vol68/iss2/5.

[196] Bode, *Human-Machine Interaction and Human Agency in the Military Domain*; Elke Schwarz, "Autonomous Weapons Systems, Artificial Intelligence, and the Problem of Meaningful Human Control," *Philosophical Journal of Conflict and Violence* 5, no. 1 (2021): 53–72, https://doi.org/10.22618/TP.PJCV.20215.1.139004.

3. Preservation of human oversight

AI applications will vary in their need for human oversight during operation. As with many legacy weapon systems, the degree of human involvement is often task-specific.[197] This understanding is also reflected in the CCW GGE LAWS' conceptualization in the rolling text of "context-appropriate human judgment & control."[198] The absence of human oversight in certain operations is not unique to AI. However, the speed and complexity of AI systems raise new concerns that humans may serve as mere rubber stampers rather than genuine decision-makers. Militaries have clear procedures to manage such risks, evident in the emphasis on human obligations across the system lifecycle and traceable chains of responsibility in legacy weapon systems, despite their capacity to operate with minimal to no human oversight. Ultimately, responsibility and accountability require that human agents possess sufficient knowledge and understanding of the relevant facts, moral considerations, and consequences of their actions. Without this understanding, they cannot make informed decisions, consider alternatives, or justify AI outputs for which they remain morally and legally accountable.[199]

The recommendations that follow aim to address the risks outlined throughout Section 3 while capturing potential opportunities. More specifically, the recommendations aim to clarify and support the attribution of responsibility and accountability, safeguard human moral agency and oversight, and establish conditions for knowledge and understanding.

[197] Tim McFarland and Zena Assaad, "Legal reviews of in Situ Learning in Autonomous Weapons," *Ethics and Information Technology* 25, no. 1 (2023): 9, https://doi.org/10.1007/s10676-023-09688-9.

[198] CCW GGE LAWS Rolling Text, 12 May 2025.

[199] Peter Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," *Proceedings of the NATO Big Data and Artificial Intelligence for Military Decision Making Specialists*, 2018, https://www.foi.se/download/18.7fd35d7f166c56e-be0b1005f/1542623791600/Possibilities-and-challenges_FOI-S--5864--SE.pdf; Arthur Holland Michel, *'The Black Box, Unlocked: Predictability and Understandability in Military AI*, September 22, 2020, https://unidir.org/publication/the-black-box-unlocked/.

4 Recommendations

GC REAIM presents its recommendations as a comprehensive framework for advancing the responsible governance of AI in the military domain. These recommendations are intended to inform global dialogue and ongoing institutionalization at national, regional, and international levels in coordination with existing and emerging initiatives.[200] Spanning all phases of the socio-technical AI system lifecycle and addressing diverse stakeholder groups, the recommendations are intended to be actionable contributions to institutional design, supporting the development of coherent and forward-looking governance. This section is structured in four interrelated layers:

1. **Guiding principles –** serve as the normative foundation for all subsequent recommendations and apply across all activities and actors.
2. **Core recommendations** – represent the main takeaways of the report as well as the most high-priority, overarching recommendations.
3. **Lifecycle recommendations –** elaborate on the activities that all relevant actors should carry out throughout the AI system lifecycle, expanding on the five continuous lifecycle activities identified in the IEEE-SA White Paper. [201]
4. **Specific Guidance** – provides targeted recommendations for states, militaries, and industry, including operational and policy measures.

All recommendations reflect a combination of technical and organizational considerations. While certain elements of recommendations recur across layers, this is by design: the recommendations are structured to reinforce one another, providing both top-down and bottom-up direction for coherent and coordinated governance.

4.1 Guiding Principles

As AI systems become more embedded in militaries, it is crucial that their development and use is governed by robust ethical and legal standards. Recent publications of leading defense and scientific bodies, as well as international organizations and initiatives dedicated to responsible governance, outline a range of core values and principles to guide the integration of AI into the military domain.[202]

[200] Jeroen van den Hoven et al., "Design for Values: An Introduction," in *Handbook of Ethics, Values, and Technological Design* (Springer, 2015), https://doi.org/10.1007/978-94-007-6970-0_40; Evgeni Aizenberg and Jeroen van den Hoven, "Designing for Human Rights in AI," *Big Data & Society* 7, no. 2 (2020): 2053951720949566, https://doi.org/10.1177/2053951720949566.

[201] Allik et al., *White Paper: A Framework for Human Decision-Making through the Lifecycle of Autonomous and Intelligent Systems in Defense Applications*.

[202] See in particular the Principles contained in: Meeting of the High Contracting Parties to the Convention on Certain Conventional Weapons, "Final Report: Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, Geneva, 13-15 November 2019," in Geneva, UN, 2019, https://digitallibrary.un.org/record/3856241; The North Atlantic Treaty Organization, "Summary of NATO's Revised Artificial Intelligence (AI) Strategy," NATO, 2024, https://www.nato.int/cps/en/natohq/official_texts_227237.htm; REAIM, "REAIM 2023 Call to Action," Government of the Netherlands, February 16, 2023, https://www.government.nl/documents/publications/2023/02/16/reaim-2023-call-to-action; European Defence Agency, "Trustworthiness for AI in Defence," 2025, https://eda.europa.eu/docs/default-source/brochures/taid-white-paper-final-09052025.pdf

These values and principles, although nominally largely overlapping, differ in both their meaning and the extent to which they apply. Nonetheless, existing documents contain commonalities that suggest a growing international consensus on what responsible AI in the military domain entails, even as elements of implementation may differ among states and organizations.[203] These points of convergence represent an ongoing, active effort to create shared norms and standards for governance.

To avoid ambiguity and reflect GC REAIM's commitment to lawful and human-centric governance, GC REAIM puts forward three guiding principles for the responsible development and use of AI in the military domain, which align with emerging consensus:

1. The development and use of AI in the military domain must comply with international law, guided by widely shared ethical principles, to best ensure the preservation of human life and peace for present and future generations.

2. The development and use of AI in the military domain must follow systematic and structured design, development, and testing processes across the entire system lifecycle. These processes must be explicitly oriented toward safeguarding human agency, responsibility, and accountability. In particular, AI systems must be engineered so that ultimate responsibility for all critical decisions remains with human operators.

3. Individuals involved in the development and use of AI in the military domain must be supported through institutional practices that enable them to exercise informed human agency. This requires continuous training, capacity-building, and the integration of design and testing features that strengthen human understanding and effective oversight.

4.2 Core Recommendations

Building upon extensive global consultations and legal frameworks, the following core recommendations are designed to guide the responsible integration of AI into the military domain. The primary objective of these recommendations is to harness AI's transformative potential while addressing associated risks by demarcating acceptable and unacceptable uses, as well as criteria for institutional design. They summarize the core takeaways of the report, as well as underpin both the lifecycle recommendations and specific guidance that follow.

1. Anchor the responsible development and use of AI in the military domain in relevant and applicable ethical principles and international law.

The responsible development and use of AI in the military domain depends on the early and sustained integration of ethical and legal considerations across the system lifecycle. This entails detailed translations of ethical and legal principles to delineate unlawful uses of AI, the development of technical, operational, and organizational guidelines, as well as respect for and application of existing obligations, such as the right to use force (*jus ad bellum*) and how force is employed during conflict (*jus in bello*).

[203] Rain Liivoja, Principles of Responsible Military Artificial Intelligence and Applicable International Law, GC REAIM Expert Policy Notes (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Liivoja.pdf.

2. Agree, at a legally binding level, that the decision to authorize the use of nuclear weapons should remain under human control.

Consistent with the policies and positions of several nuclear powers, critical decisions about the use of nuclear weapons must unequivocally remain under human authority as they require moral, legal, and strategic considerations.[204] These commitments should be pursued through either national policy declarations or an international agreement.

3. Implement national policies that guarantee human responsibility across the AI system lifecycle and that are demonstrably grounded in human-centric training and rigorous TEVV.

From the conceptualization of a military capability through its retirement, states should implement their own policies, identify key intervention points, and create positive incentives to encourage responsibility by design in companies and organizations. This process should be in line with existing due diligence obligations. This includes establishing mechanisms for ethical and legal evaluation, training and education, compliance monitoring, TEVV, risk assessments, and using internationally agreed on best practices for implementation at the national level.

4. Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain.

A global dialogue on responsible AI in the military domain should involve states (including militaries), industry, and civil society. A yearly REAIM meeting can serve as the building block for a dialogue that evolves into having a seat at an existing secretariat, or within the United Nations System, as appropriate. Regular working groups could focus on specific issues, from autonomous weapons to broader uses of AI, including C3, ISR, cyber and information, and DSS as priorities identified in Section 2.4 of this report.

5. Develop a centralized, multi-stakeholder expert network on AI in the military domain to disseminate knowledge for capability and capacity building.

This expert network should serve as a specialized external resource, centrally coordinated by a hub, that supports states and key stakeholders in proactively identifying areas of convergence across international, regional, and national AI policies. Functioning in close coordination with relevant bodies, including the UN's independent International Scientific Panel on AI and Global Dialogue on AI Governance, as well as the abovementioned dialogue, this network could contribute to and maintain trusted repositories, conduct horizon scanning, and support confidence-building measures, helping ensure that lessons learned from high-stakes military AI applications inform AI governance frameworks.

[204] France, the People's Republic of China, the Russian Federation, the United Kingdom, and the United States have all committed through national, bilateral, and/or multilateral agreements. "Principles and responsible practices for Nuclear-weapon states: Working paper submitted by France, the United Kingdom of Great Britain and Northern Ireland and the United States of America" (New York: 2020 Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons, July 2022), https:// www.un.org/sites/un2.un.org/files/npt_conf.2020_e_ wp.70.pdf; Jarrett Renshaw and Trevor Hunnicutt, "Biden, Xi Agree That Humans, Not AI, Should Control Nuclear Arms," World, *Reuters*, November 17, 2024, https://www.reuters.com/world/biden-xi-agreed-that-humans-not-ai-should-control-nuclear-weapons-white-house-2024-11-16/.

4.3 Lifecycle Recommendations

The AI system lifecycle, as elaborated in Section 2.3, provides a structured but flexible framework for guiding the integration of AI into the military domain. While lifecycle stages should not be treated as rigid or uniform across all contexts, such a framework helps pinpoint key responsibilities and activities that must be embedded at specific phases in the development process to safeguard human agency, responsibility, and accountability.[205] In turn, these activities can be adapted as needed to reflect the maturity and operational requirements of capabilities. Adopting a lifecycle approach to AI systems is critical to addressing the complexities of AI in the military domain,[206] promoting proactive, rather than reactive, governance.[207]

This section provides guidance in alignment with the IEEE-SA framework, expanding on Core Recommendation 3: **Implement national policies that guarantee human responsibility across the AI system lifecycle and that are demonstrably grounded in human-centric training and rigorous TEVV.** These recommendations apply to all actors involved in the AI system lifecycle. Beyond the nine stages of the IEEE-SA lifecycle framework adopted by GC REAIM, states should adopt the five ongoing activities:

1. Evaluation of legal, ethical, regulatory, and policy requirements;
2. Responsibility, accountability, and knowledge transfers;
3. Consider the human: training, education, and human-system integration;
4. Ensure ongoing TEVV, monitoring, hardware system or software updates and interoperability, and maintenance throughout the AI system lifecycle; and
5. Conduct risk assessments.

A key insight from the IEEE-SA project is that the effective integration of AI into the military domain depends on ensuring that risks are identified and mitigated early, rather than deferred to end-users who may have limited capacity to resolve systemic flaws.

Recommendation 1: ***Evaluation of legal, ethical, regulatory, and policy requirements.***[208]

Regular evaluation of these requirements, as for all non-AI enabled systems, will help ensure militaries adhere to IHL, along with any other applicable ethical principles and technical standards, as well as increase trust in AI systems among end-users.

[205] The report follows the CCW GGE LAWS approach, which suggests states take a lifecycle approach to legal reviews of AI in the military domain in the same way encouraged for LAWS. GGE LAWS, p 3 (Box IV.1).

[206] Afina and Paoli, *Governance of Artificial Intelligence in the Military Domain*, 24–25.

[207] While some states like the UK have chosen a lifecycle-agnostic approach, this does not mean that the recommendations found in this report are not applicable. The five ongoing activities are applicable regardless of the specific sequence of stages adopted within a development process. Ministry of Defence, "JSP 936: Dependable Artificial Intelligence (AI) in Defence (Part 1: Directive)," 2024, https://www.gov.uk/government/publications/jsp-936-dependable-artificial-intelligence-ai-in-defence-part-1-directive.

[208] Modified from original IEEE-SA activity to reflect suggestions from Zena Assaad and Adam Hepworth, *A Systems Engineering Lifecycle Approach to Responsible AI*, GC REAIM Expert Policy Note Series (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/05/Zena-Assaad-and-Adam-Hepworth.pdf.

1.1 **Ensure that existing, applicable laws, regulations, policies, and standards are identified and integrated into system design considerations in the initial planning stage.** Engage ethical and legal experts, along with other relevant stakeholders, to develop a plan for how these obligations will be met throughout the system lifecycle.[209]

1.2 **Ensure ethical, legal, regulatory, and policy evaluations encourage militaries, policymakers, and developers to acknowledge and address complexities stemming from the integration of AI into the military domain.** This should include analyses outlining the complexities and tradeoffs introduced by AI, especially given that these may be greater when looking at the whole lifecycle as opposed to only at the use stage.

Recommendation 2: ***Responsibility, accountability, and knowledge transfers.***

Responsibility, accountability, and knowledge must be systematically documented across the AI system lifecycle. This includes recording key design decisions, changes, testing outcomes, and uses in traceable ways. Such knowledge transfers support institutional memory, learning, and capacity building while establishing auditable chains of responsibility.[210] These records, which militaries are familiar with in non-AI contexts, support transparency across units, development teams, and allies, and provide a critical evidentiary basis for accountability, oversight, and legal assessment.[211]

2.1 **Formally map the roles and responsibilities of actors across the AI system lifecycle to clarify internal processes, support accountability, and prevent the diffusion of responsibility.**[212] This mapping must be grounded in an understanding that legal obligations in armed conflict are non-delegable and remain with human actors. It can help ensure that normative obligations are not obscured across institutional or technical boundaries. At the same time, it supports the design of organizational structures that reinforce responsibility assignment as well as ex-post accountability mechanisms in ethically and legally complex operational settings. Ultimately, such a mapping can also help prevent the perceived transfer of human accountability to machines.[213]

Spotlight Box 6: Socio-technical systems modelling of the European Defence Agency

The European Defence Agency proposes socio-technical systems modeling as an approach that captures the complex, dynamic interactions between technical elements (such as AI tools, data systems, tasks, infrastructure) and social components (including individuals, teams, organizational structures, national frameworks, and regulatory environments), defining each of the stakeholder roles clearly from fully human-controlled to autonomous contexts.[214]

[209] This may include aligning with recommended considerations posited by relevant organizations, for example, the ICRC's position on AI-DSS presented in their submission to the UNSG on AI in the military domain. International Committee of the Red Cross, "Submission to the United Nations Secretary General on Artificial Intelligence in the Military Domain."

[210] Assaad and Hepworth, *A Systems Engineering Lifecycle Approach to Responsible AI*.

[211] Li Qiang, *Key Lessons from Existing Governance Initiatives for the Governance of AI in the Military Domain*, GC REAIM Expert Policy Note Series (GC REAIM & The Hague Centre for Strategic Studies, 2025).

[212] Raska, *Mitigating the Risks of AI-Driven OODA Loops in Military Decision-Making*.

[213] Qiang, *Key Lessons from Existing Governance Initiatives for the Governance of AI in the Military Domain*.

[214] European Defence Agency, "Trustworthiness for AI in Defence," 52.

2.2 **Maintain comprehensive audit trails that span across the AI system lifecycle.** Audit trails must log critical elements such as human intervention points, model configurations and parameters, training and validation datasets, software patches, contextual metadata, and system outputs. All TEVV processes must be rigorously documented to preserve institutional memory and normative integrity.[215] Evaluation must begin at design stages, with mandatory risk classification and termination obligations. [216]

2.3 **Maintain incident reporting mechanisms to support lifecycle trackability and transparency.**[217] These mechanisms, familiar to militaries through processes developed for non-AI enabled systems, provide structured procedures for the real-time and post-incident capture of anomalous behavior, malfunctions, or adversarial exploits. This includes ensuring internal reporting mechanisms for violations of ethical, legal, and policy requirements. By enabling the analysis and dissemination of these operational insights across teams, incident reporting facilitates organizational learning and risk mitigation required for long-term system improvement.[218]

Spotlight Box 7: Model cards and datasheets

In recent years, tools such as model cards and data sheets for datasets have emerged to formalize AI system documentation. Model cards provide structured summaries of an AI model's performance characteristics, limitations, intended uses, and potential risks. Data sheets detail dataset provenance, collection methodology, labeling processes, and known biases. These tools generate vital information that can guide deployment decisions and oversight mechanisms, particularly in complex and uncertain environments. [219]

However, while valuable in civilian contexts, these tools alone are not sufficient to meet the distinctive demands of the military domain. Military uses require additional considerations, including data classification, adversarial robustness, mission-critical reliability, and the potential use of AI in lethal decision chains. Standardized documentation tools must be adapted or augmented to operate in classified environments, where full public disclosure is not possible, and where models must withstand adversarial manipulation, including spoofing, deception, and electronic warfare tactics. Concurrently, commanders and operators would need access to suitably qualified personnel to interpret and explain technical information, ensuring it is understood and operationally relevant in time-sensitive and high-stakes settings.

[215] The UK Ministry of Defence's (MoD) JSP 936 directive's auditability and traceability section provides additional explanation of how such auditing can take place, United Kigdom Ministry of Defence, "JSP 936", https://www.gov.uk/government/publications/jsp-936-dependable-artificial-intelligence-ai-in-defence-part-1-directive

[216] Cent. AI Digit. Policy, "Universal Guidelines for AI." (2018).

[217] Ren Bin Lee Dixon and Heather Frase, *AI Incidents: Key Components for a Mandatory Reporting Regime* (Center for Security and Emerging Technologies, 2025), https://cset.georgetown.edu/publication/ai-incidents-key-components-for-a-mandatory-reporting-regime/.

[218] Vincent Boulanin, *The Risks of Integrating Generative AI into Weapon Systems*, GC REAIM Expert Policy Notes (GC REAIM, 2025). Other policy documents such as the UK Ministry of Defence's JSP 936 directive's lifecycle assurance section also provide additional explanation of how such auditing can take place.

[219] Yasmin Afina and Sarah Grand-Clément, *Bytes and Battles*.

Recommendation 3: ***Consider the human: training, education, and human-system integration.***[220]

Responsibility in the context of AI-enabled systems in the military domain is not solely determined by legally binding task divisions or technical design. It is shaped by individual characteristics and educational background, as well as institutional conditions and operational contexts. Accordingly, preserving human agency, responsibility, and accountability throughout the AI system lifecycle requires not only institutional mechanisms through which these are reinforced but also training and education. Overall, policies must facilitate meaningful communication between those involved in system development and the end users. Understanding how systems will be used in practice, and what users actually need, must inform early design choices, helping to ensure that systems are both operationally relevant and ethically sound.

3.1 **Develop and adopt technical training and education to support human judgment in military decision-making.** This includes programs designed to reduce the risk of cognitive biases, such as automation bias, or programmatic biases, such as discriminatory training data, by countering deskilling and teaching bias identification.

3.2 **Develop and adopt, as needed, organizational policies that prioritize human judgment.** This is especially important in the context of the potential integration of AI-DSS systems not only throughout operations but also throughout the lifecycle. Examples include preventing decision-making that prioritizes military necessity over humanitarian concerns, and ensuring that proportionality assessments remain rooted in human judgment.[221]

3.3 **Ensure that commanders, operators, and other relevant end-users have SOPs that align with new training and education.** This might include training programs that help operators to understand the intended use, applicable scenarios, and reliability of systems, as well as the procedures for addressing malfunctions.[222] This may also require an understanding of how risk assessments around each system are conducted. Critical interrogation is especially important at the commander level.[223]

3.4 **Explore and adopt practices in line with or inspired by human readiness levels.**[224] To ensure that humans have been appropriately considered throughout the lifecycle of the system, standards akin to human readiness levels, just as technical standards, may help ensure technical systems are ready to be used by humans.

[220] Zena Assaad, "A Risk-Based Trust Framework for Assuring the Humans in Human-Machine Teaming," *Proceedings of the Second International Symposium on Trustworthy Autonomous Systems* (New York, NY, USA), TAS '24, Association for Computing Machinery, September 16, 2024, 1–9, https://doi.org/10.1145/3686038.3686045.

[221] Mun-eon Park, *Restrictions on AI Weapons in Specific Situations* (GC REAIM, 2025), https://hcss.nl/wp-content/uploads/2025/04/Park.pdf; Dorsey, *Proportionality under Pressure: AI-Based Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*.

[222] Qiang, *Key Lessons from Existing Governance Initiatives for the Governance of AI in the Military Domain*.

[223] Raska, *Mitigating the Risks of AI-Driven OODA Loops in Military Decision-Making*.

[224] George Salazar et al., "Understanding Human Readiness Levels," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, no. 1 (2020): 1765–69, https://doi.org/10.1177/1071181320641427.

Spotlight Box 8: Techniques for responsibility by design

The Team Design Moral Patterns (TDMP) approach structures moral task allocation within human-machine teams by ensuring humans make morally significant decisions. Rather than assigning specific actions, humans set goals for AI systems, preserving operational flexibility, supporting ex-ante human control in dynamic environments.[225]

The enveloping within a moral operational design domain (MODD) approach sets predefined moral boundaries within which AI systems must operate.[226] Inspired by aviation safety mechanisms, "enveloping" restricts AI from acting outside of pre-defined moral limits. This ensures predictable behavior, aligned with ethical and legal constraints.

The defeaters and design for operator contestability approach introduces 'defeaters', system-generated signals or external evidence that allow human operators to question or override AI outputs.[227] Ensuring contestability when operationally viable can help maintain human accountability and prevent overreliance on AI.[228]

The hard choice flagging approach enables AI systems to identify morally complex decisions, "hard choices", where no option is clearly superior.[229] Rather than resolving such decisions algorithmically, the system flags them for human judgment and helps ensure that human agents retain normative agency in ethically ambiguous situations.

Recommendation 4: ***Ensure ongoing TEVV, monitoring, hardware system or software updates, interoperability, and maintenance throughout the AI system lifecycle***.

The capacity to conduct TEVV for AI-enabled military systems should occur at the same level of robustness as existing TEVV for other safety-critical systems.

4.1 **Commit to investing in TEVV for AI at a level that ensures decision-makers can have confidence in AI-enabled systems equivalent to that placed in comparable non-AI systems.**

4.2 **Introduce review beyond predefined TEVV milestones.** This should include monitoring and post-deployment legal audits to account for shifting jurisprudence, emerging norms, and newly ratified international instruments.

[225] Jurriaan van Diggelen et al., "Team Design Patterns for Meaningful Human Control in Responsible Military Artificial Intelligence," in *Bridging the Gap Between AI and Reality*, ed. Bernhard Steffen (Springer Nature Switzerland, 2025), https://doi.org/10.1007/978-3-031-75434-0_4.

[226] Luciano Cavalcante Siebert et al., "Meaningful Human Control: Actionable Properties for AI System Development," *ResearchGate*, ahead of print, July 24, 2025, https://doi.org/10.1007/s43681-022-00167-3.

[227] Herman Veluwenkamp and Stefan Buijsman, "Design for Operator Contestability: Control over Autonomous Systems by Introducing Defeaters," *AI and Ethics* 5, no. 4 (2025): 3699–711, https://doi.org/10.1007/s43681-025-00657-0.

[228] K. L. Mosier et al., "Automation Bias: Decision Making and Performance in High-Tech Cockpits," *The International Journal of Aviation Psychology* 8, no. 1 (1997): 47–63, https://doi.org/10.1207/s15327108ijap0801_3; Mary Cummings, "Automation Bias in Intelligent Time Critical Decision Support Systems," *AIAA 1st Intelligent Systems Technical Conference*, American Institute of Aeronautics and Astronautics, September 20, 2004, https://doi.org/10.2514/6.2004-6313.

[229] Ruth Chang, "Human in the Loop!," in *AI Morality*, ed. David Edmonds (Oxford University Press, 2024), 222–34, https://doi.org/10.1093/oso/9780198876434.003.0021.

4.3 **Develop data quality requirements, specifications, and standards related to training, post-deployment learning, protection, and decommissioning.** This should include reviewing datasets to detect unwanted biases and protection against adversarial manipulation and data spoofing.[230]

4.4 **Conduct TEVV more regularly for systems that can learn and update as a result of real-world actions and interactions.** This should include establishing a repeatable validation process to ensure proper functioning of the system, including checks after significant updates. This will maximize system effectiveness and minimize the risk of hazardous outcomes.[231]

Recommendation 5: ***Conduct risk assessments.***

Militaries always conduct risk assessments on systems, and AI systems should not be an exception. The depth and extent of risk assessments should mirror the level of potential concern surrounding an AI system. Risk assessments should be conducted as appropriate throughout the AI system lifecycle.[232]

5.1 **Develop standardized risk assessments for military AI systems.** These must be interoperable with civilian risk standards (such as the National Institute of Standards and Technology [NIST]) and adaptable to the specificities of the military domain.[233] Military risk assessments must be designed to ensure systems behave as expected.

5.2 **Develop a general AI capabilities database to highlight potential risks that may be associated with individual capabilities or systems of combined capabilities, especially with respect to potential contexts of use.** Acknowledging the general-purpose nature of the technologies, discerning hierarchies of risk to utilizing AI in different types of military operations and contexts is important to establish thresholds permitting or prohibiting specific systems, use cases and/or applications.[234]

5.3 **Conduct pre-deployment evaluations as a category of structured and anticipatory assessment protocols that aim to systematically identify potential ethical, legal, and operational risks of AI-enabled military systems.** In furthering a responsibility by design approach, states must build on Article 36 of AP I by recognizing that, while valuable, legal weapons reviews alone may be insufficient given industry's expanding role in developing military technologies. Responsibility, legality, or compliance by design seek both to foster voluntary industry engagement with legal and ethical standards and encourage states to mandate early demonstration of risk mitigation through procurement and contracting processes.[235]

[230] Raska, *Mitigating the Risks of AI-Driven OODA Loops in Military Decision-Making*.

[231] Boulanin, *The Risks of Integrating Generative AI into Weapon Systems*.

[232] Assaad and Hepworth, *A Systems Engineering Lifecycle Approach to Responsible AI*.

[233] "AI Standards: Federal Engagement," National Institute of Standards and Technology, March 14, 2019, https://www.nist.gov/artificial-intelligence/ai-standards-federal-engagement.

[234] Conn et al., *An Approach for Assessing Autonomous and AI-Enabled Capabilities within Weapons Systems*; Meerveld and Lindelauf, *Context Is Everything: Policy Implications of the Military AI Responsibility Contextualization Framework*; Qiang, *Key Lessons from Existing Governance Initiatives for the Governance of AI in the Military Domain*.

[235] "Expert Meetings on the Legal Review of Autonomous Weapons Workshops," Article 36 Legal – Dr Damian Copeland, accessed August 21, 2025, https://www.article36legal.com/aws-workshops; Afina, "Combat Code Compliance."

Spotlight Box 9: AI impact assessments, examples from Canada and UNESCO

Some governments, like Canada, have introduced the requirement for Algorithmic Impact Assessments (AIA) during procurement processes.[236] This assessment framework, applicable generally outside the military context, examines how automated systems might affect individual rights, health, economic well-being, and environmental conditions, and categorizes systems into four impact levels, from I (minimal impact) to IV (very high impact). Based on the assigned level, appropriate risk mitigation strategies should be implemented, especially for systems deemed higher risk. Additionally, the Canadian directive mandates the public disclosure of AIA results via an open governmental portal, thus reinforcing a principle of public accountability and transparency.

Meanwhile, the AI Ethical Impact Assessment, developed by UNESCO for AI systems used in civilian contexts, is another step-by-step process to help government officials and private sector AI deployers make sure that their procured and in-house developed AI systems are aligned with ethical values and principles, as outlined in the UNESCO 2021 Recommendation on the Ethics of AI. The tool aims to ensure that AI systems comply with fundamental principles such as human rights, labor rights, and environmental considerations. An integral part of the process is open discussions with AI providers and other stakeholders.[237]

4.4 Specific Guidance for States and Industry

Given the complexity and distributed nature of AI integration into the military domain, guiding principles and cross-cutting recommendations must be complemented by guidance tailored to specific actors. The responsibilities associated with the development and use of AI systems are not evenly shared; states, militaries, and industry hold distinct obligations. This section provides targeted recommendations for these actors, helping to ensure that the above recommendations are meaningfully implemented across all levels of decision-making. In an increasingly multipolar international environment such guidance serves a broader purpose: supporting the implementation of transparency measures, confidence-building mechanisms, and risk mitigation strategies. These tools are essential for fostering trust and reinforcing a collective commitment to responsible governance of AI in the military domain.

4.4.1 Specific Guidance for States

The commitment to responsible AI in the military domain requires a layered governance approach that involves the national, regional, and international levels.[238] States can foster responsible practices at the national level by adopting policies and adapting military practices, while international efforts aim to harmonize global standards and norms. The growing alignment of UN resolutions with initiatives like REAIM, regional declarations, and national policies demonstrates a broad interest among states in keeping human judgment and oversight

[236] Government of Canada, "Algorithmic Impact Assessment Tool," 2025, https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html.

[237] United Nations Educational, Scientific and Cultural Organization, "Recommendation on the Ethics of Artificial Intelligence."

[238] UNIDIR Security and Technology Programme, *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security: An Evidence-Based Road Map for Future Policy Action* (2025), https://unidir.org/publication/artificial-intelligence-in-the-military-domain-and-its-implications-for-international-peace-and-security-an-evidence-based-road-map-for-future-policy-action/.

in military AI decisions. The following recommendations are intended to allow for synergies between these ongoing initiatives while leaving space for diverse perspectives.

4.4.1.1 Guidance for Militaries

Beyond overarching lifecycle recommendations, there are additional actions that militaries should take to ensure the responsible adoption of AI for specific military activities. This section provides targeted guidance for militaries to operationalize comprehensive and principled lifecycle approaches. While not exhaustive, the points outlined here represent considerations that should, at a minimum, be addressed to ensure AI-enabled systems are integrated in a manner that is safe, ethically sound, legally compliant, and operationally viable.

1. **Consider the commander's intent throughout the AI system lifecycle, if not already addressed, in demonstrable and proactive ways**.

    1.1. Determine mechanisms for documenting intent across the various stages of the AI system lifecycle.

    1.2. Ensure a military's intent for deploying an AI system is aligned with relevant legal obligations.

    1.3. Ensure the intended use of AI systems in specific operational contexts is clearly captured during the pre-employment lifecycle stages.[239]

    1.4. Ensure clear command responsibility at different stages of the system's use in operations such as planning and execution.[240]

2. **Update and adapt ROEs, SOPs, and tactics, techniques, and procedures (TTPs), to reflect intended applications and operational context in which AI systems are to be used.**[241]

    2.1. Update SOPs and TTPs to clearly demarcate and oversee intended and non-intended uses of AI systems.

    2.2. Document and utilize learning insights from AI system operations to inform or refine SOPs and TTPs for subsequent operations.

    2.3. Employ ongoing assessments of AI systems. In the event of unexpected or unforeseen outputs, investigate whether the system was used in line with doctrine or SOPs, and reassess the suitability of deploying the system.[242]

    2.4. Establish clear escalation procedures for AI systems to alert human operators when they encounter unexpected scenarios involving potential threats to the life of protected persons.

[239] Maarten Schadd et al., "How a Machine Can Understand the Command Intent," The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology 22, no. 1 (2025): 41–58, https://doi.org/10.1177/15485129221115736.

[240] Centre for Humanitarian Dialogue, *Code of Conduct on Artificial Intelligence in Military Systems*, 2021, para. 9. https://hdcentre.org/wp-content/uploads/2021/08/AI-Code-of-Conduct.pdf.

[241] Giacomo Persi Paoli and Yasmin Afina, *The Tactical Governance of Artificial Intelligence in the Military Domain*, GC REAIM Expert Policy Note Series, (GC REAIM, 2025), 9, https://hcss.nl/wp-content/uploads/2025/05/Dr-Giacomo-Persi-Paoli-and-Dr-Yasmin-Afina.pdf.

[242] Miles Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," arXiv preprint, December 1, 2024, https://doi.org/10.48550/arXiv.1802.07228.

3. **Develop and implement requirements around after-action reviews (AARs), decommissioning, and disposal in demonstrable and proactive ways.**[243]

   3.1. Ensure decommissioning considerations are captured early on for all appropriate lifecycle stages, reflecting the various reasons why a system may need to be decommissioned at each stage.

   3.2. Determine the parameters that make an AI system fit for purpose and what the limitations are of system updates against those parameters. [244]

   3.3. Ensure regular maintenance, such as system updates, is conducted to avoid the obsolescence of AI systems and minimize the subsequent risks of vulnerability, system degradation, or exploitation.

   3.4. Determine what formal procedures are needed to dispose of AI systems, including data sanitization, hardware dismantling, software retirement, and the potential reuse or repurposing of system components.[245]

   3.5. Conduct regular AARs to assess system performance and identify lessons learnt and areas for improvement.

   3.6. Encourage the use of, as appropriate, secure channels within the military chain of command for operators to report concerns about AI system misuse or malfunction. Ensure transparency by providing relevant documentation, including the legal reviews and other materials that support system explainability.

4.4.1.2 Guidance for Policymakers

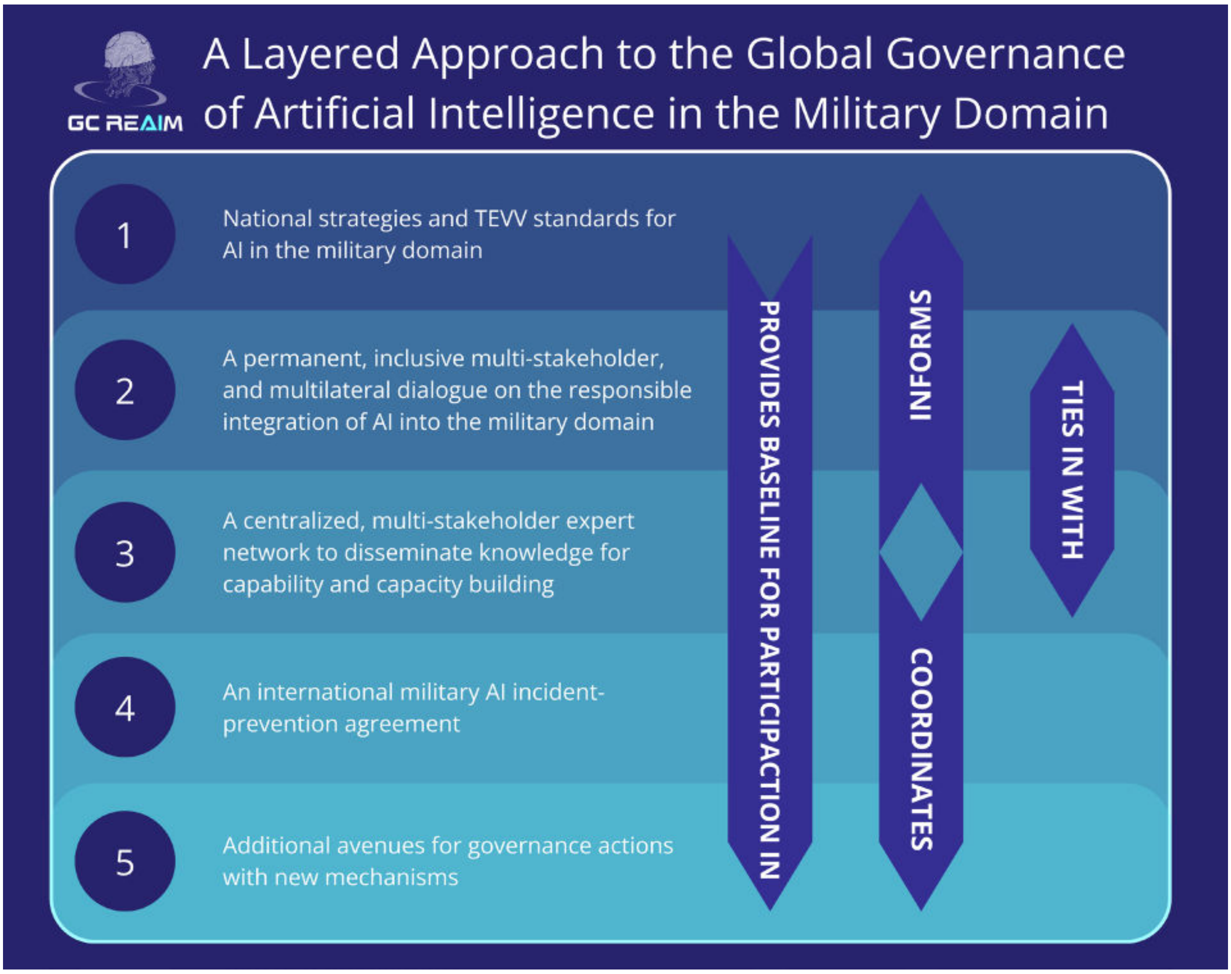
Targeted guidance for policymakers operating at national and international levels is also required, given their role in setting strategic priorities and enabling institutional coordination. This section expands on the interrelated Core Recommendations 4: **Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain.** and 5: **Develop a centralized, multi-stakeholder expert network on AI in the military domain to disseminate knowledge for capability and capacity building** by outlining key actions to support effective governance as well as coherence across sectors and jurisdictions. The guidance highlights foundational steps, escalating from baseline requirements to more aspirational actions to advance the governance of AI in the military domain.

[243] Afina and Paoli, *Governance of Artificial Intelligence in the Military Domain*, 24–25

[244] Giacomo Persi Paoli and Yasmin Afina, *The Tactical Governance of Artificial Intelligence in the Military Domain*, 9.

[245] Anthony King, "Digital Targeting: Artificial Intelligence, Data, and Military Intelligence," *Journal of Global Security Studies* 9, no. 2 (2024), https://doi.org/10.1093/jogss/ogae009.

Graphic 2. **A Layered Approach to the Global Governance of Artificial Intelligence in the Military Domain**


GC REAIM
A Layered Approach to the Global Governance of Artificial Intelligence in the Military Domain
1
National strategies and TEVV standards for AI in the military domain
2
A permanent, inclusive multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain
3
A centralized, multi-stakeholder expert network to disseminate knowledge for capability and capacity building
4
An international military AI incident-prevention agreement
5
Additional avenues for governance actions with new mechanisms
PROVIDES BASELINE FOR PARTICIPACTION IN
INFORMS
COORDINATES
TIES IN WITH

1. **Develop, adopt, and publish national strategies and TEVV standards for AI in the military domain**.[246] States should release national military AI strategies that outline their commitment to best practices, including TEVV, to ensure the responsible development and use of AI. The responsible integration of AI into the military domain is impossible without proven TEVV methods and standards to build trusted confidence inside and outside of militaries in the behavior of military AI systems. Only sovereign governments can establish the dense institutional requirements of effective management for the development and use of military AI, and do so in a manner which fulfills their binding international legal obligations under the applicable ethical principles. National strategies should:[247]

   1.1. Align military AI development and use to the lifecycle and military-specific recommendations above.

   1.2. Adapt existing or develop, as required, new doctrines, SOPs, TTPs, logbooks, and AARs to account for the integration of AI. Special attention should be given to ROEs to ensure clear chains of accountability and operations conducted in compliance with international and national law.

[246] Boulanin, *The Risks of Integrating Generative AI into Weapon Systems*; Paoli et al., *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security – An Evidence-Based Road Map for Future Policy Action*.

[247] Yasmin Afina and Giacomo Persi Paoli, *The Nuts and Bolts of the Governance of Military Artificial Intelligence: A Balancing Act Between Technical Depth and Practicality*, GC REAIM Expert Policy Note Series (GC REAIM, 2025).

1.3. Draw on the international best practices and lessons learned.

1.4. In line with national procedures, ensure accountable chains of command have the necessary information to ensure the responsible and lawful use of AI in military operations.

2. **Establish a permanent, inclusive, multi-stakeholder, and multilateral dialogue on the responsible integration of AI into the military domain.**[248] Continuing dialogue through diplomatic processes such as REAIM is a building block for more sustained cooperation and progressive alignment with the civilian domain and evolving industry standards. These dialogues represent the first step to effective confidence-building measures and capacity-building that is an essential contribution to norm development. Novel dialogue mechanisms would also help create synergies and convergence around already agreed-upon principles and serve as a cooperative tool for advancing global governance of AI if states wish to move towards a more universal regulatory system in this critical domain.[249] This dialogue could, at minimum, contribute to the following functions:

2.1. **Host capacity and capability-building engagements**. State and multi-stakeholder meetings could serve as a means by which states can build knowledge and expertise around the responsible integration of AI in the military domain. The exchange of best practices, as well as scenario planning and crisis management simulations, can help identify vulnerabilities, make projections, and enhance communication protocols nationally and internationally.[250]

2.2. **Serve as a platform for interaction with civilian AI governance initiatives.** Cross-sector collaboration, especially with private industry, and shared learning are essential for developing responsible and robust AI systems. Forming structured mechanisms for regular knowledge exchange and dissemination of best practices from the civilian sector is critical for fostering talent development. This could occur within the UN context, in interaction with the Independent International Scientific Panel on AI and the Global Dialogue on AI Governance, among others.

2.3. **Leading to institutional paths and formal governance mechanisms.** As and when states reach agreements over particular red lines or areas of cooperation in AI in this dialogue,[251] there may be a role for developing (technical) mechanisms for governance, in concert with recommendation 5 in this section.

[248] Mohammed Soliman, *The Strategic Imperative of Inclusion: Global South and Middle Power Perspectives on AI in the Military Domain*, GC REAIM Expert Policy Note Series (GC REAIM, 2025).

[249] Afina, *Draft Guidelines for the Development of a National Strategy on AI in Security and Defence – A Policy Brief*.

[250] Paoli et al., *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security – An Evidence-Based Road Map for Future Policy Action*.

[251] Ben Bucknall et al., "In Which Areas of Technical AI Safety Could Geopolitical Rivals Cooperate?," *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA), FAccT '25, Association for Computing Machinery, June 23, 2025, 3148–61, https://doi.org/10.1145/3715275.3732201.

Spotlight Box 10: The Independent International Scientific Panel on AI and the Global Dialogue on AI Governance[252]

Following "The Pact for the Future" and its annex, "The Global Digital Compact", the UNGA has established, by consensus, an Independent International Scientific Panel on AI and initiated a Global Dialogue on AI Governance. The panel is intended to serve as a multidisciplinary and geographically representative group of experts who will provide evidence-based scientific assessments on policy-relevant impact of AI through annual summary reports. These reports are to be presented at the annual meetings of the Global Dialogue, taking place in the margins of existing UN conferences. The Global Dialogue will bolster global capacity-building efforts on AI, supported by the UNSG Secretariat. While these activities are limited to the non-military domain, GC REAIM supports the creation of a similar panel focused on AI in the military domain, sharing the goals of capacity-building and providing credible and actionable information for policy-makers.

3. **Develop a centralized, multi-stakeholder expert network on AI in the military domain to disseminate knowledge for capability and capacity building.** National approaches must be backstopped through knowledge developed and maintained through an expert network. This network should be composed of geographically diverse experts with deep knowledge of military operations, AI technologies, legal frameworks, ethics, and other relevant dimensions. Its primary function would be to assist states in navigating complex policy challenges and to contribute to informed guidance of international governance efforts. Supported by multi-stakeholder actors, this network can build on the dialogue described above to promote norms of responsible behavior for the military development and use of AI aligned with civilian AI governance, when appropriate. This network could, at minimum, contribute to the following functions:[253]

    3.1. **Maintain international repositories of best practices, including AI concepts and definitions,[254] AI capabilities in specific contexts of application, TEVV standards,[255] and military AI policies.[256]** This continuously updated and developed body of knowledge could be maintained through different kinds of mechanisms. A future international agency focused on AI could also help promote these standards, much like the International Civil Aviation Organization.[257]

[252] Stéphane Dujarric, Spokesman for the Secretary-General, *Statement on The General Assembly Decision on New Artificial Intelligence Governance Mechanisms Within the United Nations*, 26 August 2025, United Nations Secretary-General, https://www.un.org/sg/en/content/sg/statement/2025-08-26/statement-attributable-the-spokesperson-for-the-secretary-general-%E2%80%93-the-general-assembly-decision-new-artificial-intelligence-governance-mechanisms-within-the-united.

[253] This recommendation builds upon multi-national cooperation recommendations provided by states and organizations such as the African Union, Australia, Singapore, the ROK, and UNDIR.

[254] A centralized and trustful repository of machine-readable ontologies discerning AI concepts is an important step towards this. Prestes et al., *Effective Governance Through Precise Common Understanding*; Conn et al., *An Approach for Assessing Autonomous and AI-Enabled Capabilities within Weapons Systems*; Meerveld and Lindelauf, *Context Is Everything: Policy Implications of the Military AI Responsibility Contextualization Framework*.

[255] Johnson, "Artificial Intelligence and Nuclear Stability: Understanding AI's Impact on Military Escalation Dynamics and Strategic Deterrence."

[256] A notable example is the "OECD's Live Repository of AI Strategies & Policies," accessed August 20, 2025, https://oecd.ai/en/dashboards.

[257] Mette Eilstrup-Sangiovanni, "Ordering Global Governance Complexes: The Evolution of the Governance Complex for International Civil Aviation," *The Review of International Organizations* 17, no. 2 (2022): 293–322, https://doi.org/10.1007/s11558-020-09411-z.

3.2. **Maintain an AI incident repository.** As described above in Section 4.3, centrally defining AI system incident categories and tracking incidents is crucial for minimizing risk and enabling transparency and accountability.[258]

3.3. Introduce a "**Military Readiness Assessment Methodology**" self-assessment and peer-review cycle for states and militaries to benchmark progress and steer capacity-building,

Spotlight Box 11: Potential parameters of entries for the military AI capabilities database[259]

A dynamic database detailing risk assessments of various capabilities could contain:

1. The name and technical description of the capability/tool
2. The intended use cases
3. The types of programming techniques used in design
4. Known risks of the tool. Aside from autonomous capabilities, specific weapon systems should also be included. Weapon system entries should contain the aforementioned 4 pieces of information, but should also provide information on
5. Descriptions of the internal components, such as code and algorithmic function
6. Descriptions of the external components, such as inputs or user prompts
7. A critical description of the design, research, and development process
8. A critical description of the procurement, acquisition, and manufacturing process
9. A critical description of the testing, evaluation, verification, and validation process.
10. An explanation of necessary human training and risks
11. A critical description of tactical deployment.

These categories are based on the Chain of Responsibility (COR) model and the IEEE-SA Lifecycle Framework and should be considered a starting point rather than a definitive list. The COR model is a mechanism for attributing liability in human-machine teams across the AI lifecycle.[260]

4. **Develop an international military AI incident-prevention agreement.** Some of the most prominent risks to international peace and security come from the potential for AI to increase the risks of accidents or unintended escalation.[261] All states should be in agreement that military AI systems should function in stable, predictable ways. States should draft and negotiate a legally binding AI Incidents Agreement that would include degrees of transparency on AI during peacetime.[262]

[258] Ren Bin Lee Dixon and Heather Frase, *An Argument for Hybrid AI Incident Reporting: Lessons Learned from Other Incident Reporting Systems* (Center for security and emerging technologies, 2024), https://cset.georgetown.edu/publication/an-argument-for-hybrid-ai-incident-reporting/; Ren Bin Lee Dixon and Heather Frase, "AI Incidents: Key Components for a Mandatory Reporting Regime," *Center for Security and Emerging Technology*, 2025, https://cset.georgetown.edu/publication/ai-incidents-key-components-for-a-mandatory-reporting-regime/.

[259] Conn et al., *An Approach for Assessing Autonomous and AI-Enabled Capabilities within Weapons Systems*.

[260] Brendan Walker-Munro and Zena Assaad, "The Guilty (Silicon) Mind: Blameworthiness and Liability in Human-Machine Teaming," *Cambridge Law Review* 8 (2023): 1, https://law.uq.edu.au/files/98673/Walker-Munro_Assaad_Human-Machine_Teaming.pdf.

[261] Maas et al., "10. Military Artificial Intelligence as a Contributor to Global Catastrophic Risk"; Wendell Wallach, *A Dangerous Master: How to Keep Technology from Slipping Beyond Our Control* (Sentient Publications, 2015); Michael C Horowitz and Paul Scharre, *AI and International Stability – Risks and Confidence-Building Measures* (Center for a New American Security, 2021), https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/AI-and-International-Stability-Risks-and-Confidence-Building-Measures.pdf.

[262] Matthijs M. Maas, "How Viable Is International Arms Control for Military Artificial Intelligence? Three Lessons from Nuclear Weapons," *Contemporary Security Policy* 40, no. 3 (2019): 285–311, https://doi.org/10.1080/13523260.2019.1576464. This recommendation also builds upon the multilateral recommendations provided by policy documents such Paoli et al., *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security – An Evidence-Based Road Map for Future Policy Action* among others.

Spotlight Box 12: Lessons learned from the Incidents at Sea Agreement

One of the most successful confidence-building measures from the Cold War was the Incidents at Sea Agreement between the US and the Soviet Union, which dramatically reduced the risk of unintended peacetime engagements between US and Soviet forces at sea, decreasing the probability of nuclear war.[263] The Incidents at Sea Agreement was signed in 1972 by the US and the Soviet Union as part of the Strategic Arms Limitations Talks/Treaty (SALT) I process. It created rules of the road for how US and Soviet ships, and aircraft would interact after a series of accidents in the 1960s. The agreement adopted civilian sector standards and best practices to clarify expectations during peacetime encounters. The Incidents at Sea Agreement was successful because it focused specifically on a pathway for escalation (a spiraling conflict due to an accidental encounter between naval and air assets) that would have resulted in a war that no one wanted. This achievement demonstrates that confidence-building measures are not only feasible in times of high uncertainty but also offer a concrete path towards their reduction.[264]

5. **Consider establishing additional avenues for governance actions with new mechanisms.** From codes of conduct, plans of action, and industry standards to a framework convention on AI in the military domain. For example, in addition to the formal dialogue platform recommended above, states could create novel institutions and mechanisms to support convergence around already agreed-upon principles, reaffirm, and strengthen their obligations under existing international legal frameworks. These additional governance mechanisms could serve as a building block for advancing regulatory approaches if states wish to move toward a more universal regulatory system, such as a Framework Convention, to bolster governance.[265] Alternatively, states could develop a “soft law” framework with an implementation mechanism to (1) highlight existing IHL responsibilities and emerging norms around AI in the military domain and (2) build on the momentum generated by the REAIM Summits.[266] The first step in such a process would be to launch preparatory discussions for a soft-law framework to transform shared principles into measurable practices.

These recommendations underscore the critical need for a layered governance framework to ensure the responsible integration of AI into the military domain worldwide. They emphasize aligning national strategies with international standards, establishing global institutionalized ways to achieve cooperation, and promoting multi-stakeholder dialogue to foster norms and transparency. The overall goal is to align military AI policies with international law and foster global cooperation to ensure the safe and ethical integration of AI into the military domain.

[263] Michael C. Horowitz and Lauren Kahn, “Leading in Artificial Intelligence Through Confidence Building Measures,” *The Washington Quarterly*, 44:4 (2021), https://doi.org/10.1080/0163660X.2021.2018794

[264] For additional recommendations on confidence building measures see Sofia Romansky, *Lessons from the EU on Confidence-Building Measures Around Artificial Intelligence in the Military Domain*, No. 97, Non-Proliferation and Disarmament Papers (Stockholm International Peace Research Institute, EU Non-Proliferation and Disarmament Consortium, 2025), https://www.sipri.org/publications/2025/eu-non-proliferation-and-disarmament-papers/lessons-eu-confidence-building-measures-around-artificial-intelligence-military-domain.

[265] As stated by the UNSG High-level Advisory Body on AI, “If the risks of AI become more serious, and more concentrated, it might become necessary for Member States to consider a more robust international institution with monitoring, reporting, verification, and enforcement powers.” High-level Advisory Body on artificial intelligence, *Governing AI for Humanity*, § 195; Reinhold and Schörnig, *Armament, Arms Control and Artificial Intelligence*.

[266] Modelled on UNESCO’s non-binding yet universal Recommendation on the Ethics of AI and its implementation mechanism the Readiness Assessment Methodology that has received broad international support as documented in the 2025 AI and Democratic Values Report. United Nations Educational, Scientific and Cultural Organization, “Recommendation on the Ethics of Artificial Intelligence.”

4.4.2 Specific Guidance for Industry

To complement the above guidance for states, it is equally important to address the distinct responsibilities and capacities of industry developers and technology providers. Their decisions significantly influence whether downstream users can integrate these technologies in ways that align with ethical, legal, and operational standards. The following section sets out specific guidance for industry, with a view to fostering more transparent, accountable, and collaborative development processes.

1. **For industry and providers working on general-purpose AI with applications in both civilian AI and military AI systems, address questions about dual-use implications in demonstrable and proactive ways.**[267]

    1.1. Create guidelines and safeguards to ensure systems cannot be repurposed for unlawful uses or in inappropriate contexts.

    1.2. Create means of tracking potential military repurposing of AI systems that governments can use as part of end use monitoring of systems, as appropriate.

    1.3. At plausible classification levels, systematically select performance metrics and assess AI system error rates across protected categories as part of documented ex-ante human rights impact assessment.

    1.4. Consider and institute requirements for responsible disclosure to clients around AI systems developed for the military domain and dual-use applications.

2. **Address design issues around human-machine interaction, user interfaces, and experiences to enhance user understanding in clear and proactive ways that boost system effectiveness and lessen risks associated with bias.**[268]

    2.1. Ensure developers of military systems partner with end users early in the system lifecycle and throughout the development process to understand operational needs and ensure the system is appropriately aligned with how it will be used in practice.

    2.2. Establish testing procedures that account for the system's technical performance and how the system may be used across varied operational contexts.

    2.3. Incorporate design choices and invest in training that enable and encourage users to exercise agency and control over decisions that have legal or safety implications.

3. **Address questions around organizational design for innovation, research, and development in demonstrable and proactive ways**.

    3.1. Ensure that teams are sufficiently diverse, meaning they possess in-depth knowledge across all relevant disciplines, including legal and ethical experts to support the development of AI systems in compliance with IHL, Human Rights Law, and other applicable standards.[269]

[267] Edson Prestes and Michael A. Houghtaling, "How IEEE 7007-2021 Is Influencing GeoPolitics.," *Forthcoming: IEEE Robotics and Automation Magazine*, December issue (2025); Davinder Kaur et al., "Trustworthy Artificial Intelligence: A Review," *ACM Computing Surveys* 55, no. 2 (2023): 1–38, https://doi.org/10.1145/3491209.

[268] Lama H. Nazer et al., "Bias in Artificial Intelligence Algorithms and Recommendations for Mitigation," *PLOS Digital Health* 2, no. 6 (2023): 278, https://doi.org/10.1371/journal.pdig.0000278; Institute of Electrical and Electronics Engineers, "IEEE Ontological Standard for Ethically Driven Robotics and Automation Systems."

[269] Aizenberg and van den Hoven, "Designing for Human Rights in AI." This principle also builds upon AI policy documents published by states and organizations such as the African Union (AU), Australia, the Republic of Korea, and the United Kingdom. John Gerard Ruggie, "The Social Construction of the UN Guiding Principles on Business and Human Rights," in *Research Handbook on Human Rights and Business*, ed. Surya Deva and David Birchall (Edward Elgar Publishing, 2020), https://doi.org/10.4337/9781786436405.00009.

3.2. Ensure all team members are empowered to raise ethical concerns and questions around necessity, purpose, and alternatives through clear escalation pathways.

3.3. Ensure consistent information and knowledge transfer across all team members to maintain alignment, support informed decision-making, and uphold continuity throughout the AI system's lifecycle.

3.4. Establish appropriate human expert oversight during development stages, supported by rigorous early-stage testing, to ensure AI-generated outputs are safe, responsible, and effective.

4. **Develop forward-looking standards that meet the pace of innovation**.

4.1. Establish flexible frameworks and standards that can be updated alongside governance as required.

4.2. Ensure responsible innovation by actively adhering to existing IEEE and International Organization of Standardization (ISO) standards on AI, supporting the creation of new ones, and participating in ongoing development and implementation across various fields. [270]

4.3. Promote accountability across the AI system lifecycle and support efforts towards consistent AI governance across industries and beyond borders.[271]

[270] This includes ISO/IEC 22989, Artificial Intelligence – Concepts and Terminology, ISO/IEC 23053, Framework for Artificial Intelligence Systems Using Machine Learning; ISO/IEC CD 23894 Information Technology – artificial Intelligence – risk management; ISO/IEC TR 24027:2021, Information technology — Artificial intelligence (AI) — Bias in AI systems and AI-aided decision making; IEEE 7010-2020, IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being; and IEEE 7000-2021 – IEEE Standard Model Process for Addressing Ethical Concerns during System Design, and IEEE 7001-2021 – IEEE Standard for Transparency of Autonomous Systems.

[271] Vincent Boulanin et al., "AI Missteps Could Unravel Global Peace and Security," IEEE Spectrum, 2024, https://spectrum.ieee.org/ai-missteps-unravel-global-security.

Spotlight Box 13: Three initiatives to inspire industry engagement, responsibility, and ethical innovation

Roundtable for AI, Security, and Ethics

The United Nations Institute for Disarmament Research (UNIDIR) Roundtable for AI, Security and Ethics (RAISE), founded in 2024, provides a platform for multi-stakeholder dialogue, with a specific focus on inclusive industry engagement. In partnership with Microsoft, RAISE has:

- Hosted the inaugural RAISE Global Conference in 2025, "for the diplomatic ecosystem in Geneva and the wider multi-stakeholder community, including academic experts, civil society organizations, industry representatives, and research labs"". [272]
- Published a policy brief highlighting different dimensions of the multi-stakeholder perspective and priority areas for governance.[273]

The Partnership on AI

Founded in 2016 with original members including Amazon, Facebook, Google, DeepMind, Microsoft, and IBM – now with more than 100 partners from industry and civil society. The Partnership is a cross-sector hub for convening new ideas, practices with actionable guidance. Two major outcomes are:

- The AI Incident Database to anticipate and mitigate risks and harms and solve real-world problems.[274]
- The Guidelines for AI and Shared Prosperity, which create a platform for dialogue on how to amplify the uses of AI for the common good of all.

Responsible Innovation for Peace and Security

In April 2023, the United Nations Office for Disarmament and the Stockholm International Peace Research Institute (SIPRI) launched a three-year initiative on responsible innovation in AI for peace and security:[275]

- The initiative collaborates closely with industry and creators and includes a strong educational component focused on STEM students, focusing on AI.
- It reinforces the idea of responsible innovation for industry within UN activities and aligns with GC REAIM's emphasis on design for values and responsibility by design.

[272] "RAISE: The Roundtable for AI, Security and Ethics", United Nations Institute for Disarmament Research, August 15, 2024, https://unidir.org/raise/.

[273] Afina and Paoli, *Governance of Artificial Intelligence in the Military Domain*.

[274] "Welcome to the Artificial Intelligence Incident Database,"Artificial Intelligence Incident Database, accessed August 20, 2025, https://incidentdatabase.ai/.

[275] Supported by the Council of the European Union to mitigate the risks that the misuse of civilian AI technology can pose to international peace and security. Council Decision (CFSP) 2022/2269 of 18 November 2022 on Union Support for the Implementation of a Project 'Promoting Responsible Innovation in Artificial Intelligence for Peace and Security,' Pub. L. No. OJ L 300/11 (21.11.2022), 2022/2269 11 (2022). https://eur-lex.europa.eu/eli/dec/2022/2269/oj/eng.

5 Conclusion and Next Steps

This report addresses the critical challenges posed by the integration of AI into the military domain. Drawing from global consultations across five continents, engagement with numerous international initiatives, input from its Commissioners and Experts, as well as associated epistemic networks, the report examines how AI systems integrated throughout the military domain are reshaping the landscape of international peace and security, the conduct of armed conflict, and beyond. The Global Commission affirms that while AI offers significant opportunities to enhance military effectiveness, it simultaneously introduces unprecedented risks related to human agency, responsibility, and accountability, particularly when systems operate with opacity and unpredictability in high-stakes environments.

The report establishes three guiding principles for the responsible governance of AI in the military domain and operationalizes these principles in five core recommendations, along with overarching lifecycle recommendations, and specific guidance for states, militaries, and industry. The Global Commission emphasizes a responsibility by design approach that spans the entire AI system lifecycle with continuous attention to legal compliance, human training, risk assessment, and accountability mechanisms.

The first time that AI in the military domain reached the international diplomatic stage was at the 2023 REAIM Summit in The Hague. This milestone was followed by the 2024 REAIM Summit in Seoul. The Call for Action and Blueprint for Action provide an innovative opportunity to develop new global governance that incorporates diverse perspectives.[276] Central to this process has been the creation of the GC REAIM to foster awareness and understanding of AI in the military domain. Through its work and the recommendations found within this report, GC REAIM galvanizes an interdisciplinary, multi-stakeholder approach that contributes to an inclusive and human-centered conception of the responsible integration of AI into the military domain.

The REAIM process has evolved from initial awareness-building at REAIM I to norm development at REAIM II, with REAIM III expected to focus on operationalizing these norms into actionable frameworks. GC REAIM recognizes the potential synergies between the REAIM process, existing institutions within the UN system and beyond, to harness and harmonize ongoing efforts towards global governance of AI across domains.

There are three notable diplomatic areas to pursue further collaboration. The first is the UN General Assembly resolution process. On December 24th, 2024, the UN General Assembly passed a resolution titled “Artificial Intelligence in the Military Domain and its Implications for International Peace and Security,” with a strong showing of 159 votes in favor.[277] Building on

[276] Tim Sweijs and Sofia Romansky, “International Norms Development and AI in the Military Domain”; Tobias Vestner and Juliette François-Blouin, “Globalizing Responsible AI in the Military Domain by the REAIM Summit,” *Just Security*, March 13, 2023, https://www.justsecurity.org/85440/globalizing-responsible-ai-in-the-military-domain-by-the-reaim-summit/; Eleonore Fournier-Tombs et al., “A Global Architecture for Artificial Intelligence,” United Nations University, August 20, 2025, https://unu.edu/publication/global-architecture-artificial-intelligence.

[277] UNGA, *Resolution Adopted by the General Assembly on 24 December 2024 Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security,* A/RES/79/239 (United Nations, 2024), https://documents.un.org/doc/undoc/gen/n24/426/98/pdf/n2442698.pdf.

their roles as global AI leaders through REAIM Summits, the Netherlands and the ROK will continue to lead international efforts, along with partners like Spain and others, to establish norms that will encourage broader international participation in multi-stakeholder discussions.[278] The resolution prompted the UNSG to seek views and opinions from member states and civil society, resulting in a highly productive exchange on the challenges and opportunities of applying AI in the military and the narrower area of autonomous weapons.[279] The second area is discussions surrounding the implementation of the Summit of the Future's Global Digital Compact, creation of a UN Intergovernmental Process for an Independent International Scientific Panel on AI, and Global Dialogue on AI Governance, co-facilitated by Costa Rica and Spain. [280] The third, while a subset of the area of engagement by GC REAIM, is to continue working on issues around autonomous weapons systems, including with the UN and ICRC. Collaboration and shared learning across sectors are vital for creating resilient and ethical AI systems, given the interconnected nature of military and civilian AI technologies.

The priority in global governance processes should be to continue advancing multi-stakeholder approaches to deliberations around AI in the military domain. GC REAIM foresees the immediate priority as paving the way for long-term global cooperation efforts in moving from the Blueprint for Action to a 'Plan for Action' at REAIM III. This follow-up would help formulate AI governance through close cooperation among governments, policymakers, militaries, industry, academia, and civil society.

GC REAIM supports a layered governance approach that includes iteratively developed national strategies, regional confidence-building measures, permanent platforms for dialogue, as well as international institutions equipped with sufficient subject-matter expertise, best practice repositories, and incident tracking systems. This comprehensive framework serves as both a technical guide for military and industry practitioners and a diplomatic roadmap for policymakers. It highlights that the responsible integration of AI into the military domain requires a systematic focus on the design of socio-technical systems to ensure that technological progress supports the core goals of international peace and security and upholds human agency and dignity. GC REAIM calls upon all stakeholders to help ensure that AI in the military domain will be developed and used responsibly, anchored in ethics and international law, to preserve peace for present and future generations.

[278] Andrew Yeo, "South Korea as a Global Pivotal State," *Brookings*, 2023, https://www.brookings.edu/articles/south-korea-as-a-global-pivotal-state/.

[279] *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security – Report of the Secretary-General*, A/80/78 (United Nations General Assembly, 2025), https://documents.un.org/doc/undoc/gen/n25/107/66/pdf/n2510766.pdf.

[280] "Joint Call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to Establish New Prohibitions and Restrictions on Autonomous Weapon Systems" October 4, 2023, https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems. The call builds upon the seminal ICRC position on autonomous weapons that advised on new rules that are urgently needed to clarify and specify how IHL applies to autonomous weapons, International Committee of the Red Cross, "ICRC Position on Autonomous Weapon Systems."

GC REAIM
GLOBAL COMMISSION ON RESPONSIBLE
AI IN THE MILITARY DOMAIN

HCSS
Lange Voorhout 1
2514 EA The Hague

Follow us on social media:
@hcssnl

The Hague Centre for Strategic Studies

POWERED BY The Hague Centre for Strategic Studies