

A person in a military uniform is shown in profile, wearing a headset and glasses, focused on several computer monitors. The monitors display various data, including maps and technical information. The scene is dimly lit with blue and purple ambient lighting, suggesting a high-tech or command center environment. A red rectangular box is positioned in the upper right corner of the image.

**BASIC**

Governing Military AI:

# Designing Meaningful Human Control in AI-Enabled Decision-Making

Eva-Nour Repussard

OCTOBER 2026

■ **BASIC is an independent, non-profit think tank working to safeguard humanity and Earth's ecosystem from nuclear risks and interconnected security threats for generations to come.**

We have a global reputation for convening distinctive, empathic dialogues that help states overcome complex strategic and political differences.

Our established networks and expertise, developed since 1986, enable us to get the right people in the room and facilitate effective, meaningful exchange between siloed and often hostile political communities.

---

**The British American  
Security Information  
Council (BASIC)**

Work + Play  
111 Seven Sisters Rd  
Finsbury Park  
London N7 7FN

**Charity Registration No:**  
1001081

**T:** +44 20 3488 6974  
[www.basicint.org](http://www.basicint.org)

© The British American Security Information Council (BASIC), 2026

Image credits: U.S. Army Reserve photos by Spc. William Kuang, via DVIDS. The appearance of U.S. Department of War (DoW) visual information does not imply or constitute DoW endorsement.

The opinions expressed in this publication are the responsibility of the author and do not necessarily reflect the views of BASIC.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage or retrieval system, without any prior written permission of the copyright holder.

Please direct all enquiries to the publishers.

---

## Author

---



### **Eva-Nour Repussard**

Eva-Nour Repussard is a Policy Fellow and Project Manager at BASIC, where she works on the Responsibilities and Global Governance programme. She is a leading expert on emerging and disruptive technologies and ballistic missile proliferation, with a particular focus on their implications for strategic stability and international security.

She holds an MA from King's College London in Intelligence and International Security and a BA (Hons) in International Relations from the University of Birmingham and Fudan University in Shanghai.

Eva-Nour's analysis has notably been presented at the United Nations Office in Vienna, the Austrian MFA (BMEIA), the French MFA (MEAE), and at the Hague Code of Conduct (HCoC) Regional Meeting in Djibouti. She is regularly invited to speak at high-level international forums, including NATO's 5th Early-Career Nuclear Strategists Workshop, the EU Non-Proliferation and Disarmament Conference (EUNPDC), CSIS PONI Winter Conference, and the UK PONI Annual Conference, where she has shared policy-relevant research on topics ranging from cyber threats to strategic stability to missile proliferation dynamics.

## Acknowledgements

---

This report is generously funded by the Department of National Defence of the Government of Canada (DND), and is part of BASIC's Responsibilities and Global Governance and Non-Proliferation and Disarmament Programmes. The report does not directly reflect the views of the Canadian Government. BASIC is grateful for the financial support received for this project.

The author would like to thank Ching Wei Sooi for his invaluable contribution to the project and Dr Chiara Cervasio, Dr Manuel Herrera and Professor Nicholas J. Wheeler for commenting on previous drafts of this report. Moreover, the author would like to thank the interviewees who decided to take part in this project.

# Contents

---

|                          |          |
|--------------------------|----------|
| <b>Executive Summary</b> | <b>5</b> |
|--------------------------|----------|

---

|                     |          |
|---------------------|----------|
| <b>Introduction</b> | <b>7</b> |
|---------------------|----------|

---

|                                                                                |          |
|--------------------------------------------------------------------------------|----------|
| <b>Part I:</b>                                                                 |          |
| <b>Context: Risks, Opportunities and Governance Challenges for Military AI</b> | <b>9</b> |
| A. Opportunities: Enhancing Efficiency, Precision, and Strategic Advantage     | 9        |
| B. Risks: Unintended Consequences and Ethical Dilemmas                         | 10       |
| C. Governance Challenges                                                       | 11       |

---

|                                                            |           |
|------------------------------------------------------------|-----------|
| <b>Part II:</b>                                            |           |
| <b>Policy Recommendations and Governance Opportunities</b> | <b>14</b> |
| A. Lifecycle Policies                                      | 14        |
| B. Use-Case Policies                                       | 16        |
| C. Governance Opportunities                                | 18        |

---

|                   |           |
|-------------------|-----------|
| <b>Conclusion</b> | <b>20</b> |
|-------------------|-----------|



# Executive Summary

Militaries are rapidly integrating artificial intelligence (AI) into operational and strategic decision-making.

AI can process information at a scale and speed beyond human ability, enhance situational awareness, support targeting and intelligence analysis, and assist decision-making in time-critical environments.

However, these capabilities also create significant operational, legal, ethical, and strategic risks, including cognitive bias, system opacity, adversarial manipulation, unreliable outputs, and inappropriate reliance on AI-generated recommendations.

Central to both these opportunities and risks is the relationship between humans and machines. The key governance challenge is not simply whether humans remain involved in military decision-making, but whether that involvement enables meaningful judgement and control. An operator may formally remain part of a decision-making process while lacking the information, time, authority, or training

necessary to critically assess an AI-generated recommendation. Human judgement is itself imperfect: military personnel bring assumptions, cognitive biases, experiences, and perceptions of risk to their interaction with AI systems. Effective military AI governance must therefore address both the limitations of AI systems and the ways in which humans interpret, trust, challenge, and act upon their outputs.

Drawing on 32 semi-structured interviews with stakeholders across 14 countries, alongside policy and academic research, this report argues for a context-sensitive and risk-proportionate approach to Human-Machine Interaction (HMI). Rather than prescribing a uniform level of human involvement across all military AI applications, the appropriate form of oversight and decision

authority should reflect the operational context, the level of uncertainty, the consequences of error, and the time available for human assessment. Current AI systems remain vulnerable to error, adversarial manipulation, and performance degradation when operating outside the conditions in which they were developed and tested. The report focuses on how AI can augment human capabilities and improve military decision-making while ensuring that human involvement remains meaningful. Decisions involving the use of lethal force should remain with an appropriately authorised human decision-maker.

The report translates this approach into a two-tier policy framework. The first tier, *Lifecycle Policies*, establishes foundational requirements for the lifecycle of military AI systems:

- **Policy 1: Transparency and Explainability Standards for Procurement**  
Requiring levels of transparency, explainability, traceability, and auditability proportionate to the risks and operational consequences associated with an AI system.
- **Policy 2: Human-Machine Teaming (HMT) Training**  
Preparing military personnel to work effectively with AI while shaping organisational cultures and practices that support appropriate judgement over AI-generated outputs.

The second tier, *Use-Case Policies*, establishes requirements for the use of AI in different operational contexts:

- **Policy 3: AI-Enabled Data Analysis for Decision Support**  
Using AI to process, correlate, and prioritise large volumes of operational data to improve situational awareness and help personnel identify relevant patterns, anomalies, and information.
- **Policy 4: AI-Enabled Operational Decision Support**  
Using AI-enabled decision-support systems (AI-DSS) to support choices about operational action by synthesising relevant analysis, comparing options, and generating possible courses of action, including in time-sensitive and high-consequence contexts, while preserving meaningful human decision authority, particularly over the use of lethal force.

For Canada, these challenges also create opportunities for international leadership. Canada can contribute to greater convergence among NATO Allies by advancing 'responsible-by-design' approaches to military AI, supporting common standards and exercises for HMT training, and investing in research on explainable, traceable, and auditable AI. Such initiatives would not only strengthen responsible AI governance but also support interoperability by helping Allied forces develop compatible expectations about how AI-enabled systems should be designed, assessed, and used in multinational operations.

Ultimately, effective military AI governance should not treat human involvement as a binary requirement or assume that simply keeping a person 'in the loop' guarantees meaningful control.

Instead, the human-machine relationship should be designed so that AI can be used extensively for analysis and decision support while personnel retain the information, authority, and capacity necessary to exercise meaningful judgement.

# Introduction

**In an era of rapid technological change and compressed decision-making timelines, militaries are under increasing pressure to improve operational efficiency and maintain operational and strategic advantage.**

At the same time, the accelerating adoption of artificial intelligence into military operations has intensified strategic competition, creating pressure on militaries to adopt these technologies or risk falling behind adversaries.<sup>1</sup> While such integration has the potential to improve targeting, enhance situational awareness, and accelerate decision-making,<sup>2</sup> integration without proper safeguards introduces significant operational, legal, ethical, and strategic risks.<sup>3</sup>

Central to these opportunities and risks is the relationship between humans and machines. As AI systems become more capable and complex, the way in which military personnel interact with these systems becomes increasingly important. Effective Human-Machine Interaction (HMI) should support meaningful human oversight and control, enabling operators to understand, critically assess, and appropriately respond to AI-generated outputs.<sup>4</sup> However, human judgement is not inherently unbiased or objective. Military personnel's assumptions, experiences, cognitive biases, and perceptions of risk shape how they interpret and act

on AI outputs.<sup>5</sup> Poorly designed HMI can reinforce existing biases and encourage automation bias, over-reliance, or inappropriate delegation of decision-making authority.

The key challenge in HMI is not simply whether humans remain involved in military decision-making, but how that involvement is designed and governed so that it remains meaningful. Some AI-enabled systems can perform specific functions without real-time human intervention. However, current AI systems remain limited in their ability to reliably interpret complex, unfamiliar, or rapidly changing operational contexts and remain vulnerable to error, adversarial manipulation, and performance degradation when operating outside the conditions in which they were developed and tested.<sup>6</sup> This report therefore argues for a context-sensitive and risk-proportionate approach to HMI, in which the form and intensity of human oversight should reflect the operational environment, level of uncertainty, potential consequences of error, and time available for assessment. AI systems may be given greater autonomy to perform bounded, lower-risk supporting functions, while higher-consequence decisions require stronger human scrutiny and decision authority. Decisions involving the use of lethal force should remain with an appropriately authorised human decision-maker.

<sup>1</sup> Interviews 3, 10, 14, and 29; Sheera Frenkel, Paul Mozur, and Adam Satariano, "Mutually Automated Destruction: The Escalating Global A.I. Arms Race," New York Times, April 12, 2026, <https://www.nytimes.com/2026/04/12/technology/china-russia-us-ai-weapons.html>; UK Ministry of Defence, "Putting Artificial Intelligence (AI) at the Heart of UK Defence," June 10, 2026, <https://www.gov.uk/government/publications/putting-artificial-intelligence-ai-at-the-heart-of-uk-defence/putting-artificial-intelligence-ai-at-the-heart-of-uk-defence>; Dario Amodè, "Policy on the AI Exponential," June 2026, <https://darioamodei.com/post/policy-on-the-ai-exponential>.

<sup>2</sup> Interviews 7, 11, 15, and 18.

<sup>3</sup> Interviews 3 and 10; NATO, "Summary of NATO's Revised Artificial Intelligence (AI) Strategy," July 10, 2024, <https://www.nato.int/fr/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy>.

<sup>4</sup> This report uses Human-Machine Interaction (HMI) to describe how military personnel interact with, understand, oversee, and control AI-enabled systems. Human-Machine Teaming (HMT) refers more specifically to collaboration between personnel and AI systems in carrying out tasks.

<sup>5</sup> Interview 7; Defence Science and Technology Laboratory, Human-centred Ways of Working with AI in Intelligence Analysis, July 26, 2023, <https://www.gov.uk/government/publications/human-centred-ways-of-working-with-ai-in-intelligence-analysis/human-centred-ways-of-working-with-ai-in-intelligence-analysis>.

<sup>6</sup> UK Ministry of Defence, "Ambitious, Safe, Responsible: Our Approach to the Delivery of AI-Enabled Capability in Defence," June 15, 2022, <https://www.gov.uk/government/publications/ambitious-safe-responsible-our-approach-to-the-delivery-of-ai-enabled-capability-in-defence/ambitious-safe-responsible-our-approach-to-the-delivery-of-ai-enabled-capability-in-defence>.



To ground the recommendations in real-world insights, this report draws on a multi-method approach. We conducted 32 semi-structured interviews with stakeholders across 14 countries: Australia, Canada, the Czech Republic, Denmark, Finland, France, Germany, Italy, the Netherlands, Peru, Sweden, Ukraine, the United Kingdom, and the United States. Participants were selected for their diverse expertise, spanning military personnel, government officials, industry practitioners, think tank experts, and academics. To protect participants' confidentiality, interviews are cited by number, without identifying individual interviewees. These interviews were complemented by extensive desk-based research, including a review of policy documents and academic literature, to contextualise findings within military AI.

While developed in a Canadian defence context, the recommendations are deliberately framed for military institutions more broadly, reflecting governance challenges shared by Canada and its allies and the importance of developing interoperable approaches to military AI.

#### The report proceeds in two parts:

- **Part I** examines the opportunities, risks, and governance challenges associated with military AI and HMI.
- **Part II** presents a two-tier policy framework comprising *Lifecycle Policies* and *Use-Case Policies*, followed by governance opportunities for Canada to help shape allied standards and norms for responsible military AI.

# Context: Risks, Opportunities and Governance Challenges for Military AI

**Military AI is increasingly reshaping modern warfare. Ukraine's AI-enabled drones have been used in precision strikes, surveillance, and electronic warfare,<sup>7</sup> while Russia has also reportedly deployed experimental AI-guided drones capable of autonomously selecting specific targets once directed towards a target area.<sup>8</sup>**

The US, meanwhile, has invested heavily in AI, notably through the Project Maven initiative.<sup>9</sup> Project Maven includes the Maven Smart System (MSS), an AI-enabled software platform that fuses intelligence, surveillance, and reconnaissance data into a single interface to speed targeting and decision-making. MSS has been deployed across US combatant commands and multiple components of the US military.<sup>10</sup> More recently, the Pentagon has begun expanding MSS beyond intelligence and operations data to incorporate force readiness, logistics, supply-chain, and budgeting information.<sup>11</sup> Israel has also made extensive use of AI-enabled intelligence and targeting-support systems, including 'Lavender',

reportedly used to rate individuals according to suspected militant affiliation, and 'The Gospel', used to support the generation of potential targets.<sup>12</sup> China's People's Liberation Army is pursuing greater AI integration across its operations, including to enhance targeting processes. Its Air Force reportedly uses an AI system to develop plans for large-scale strike operations.<sup>13</sup>

These examples illustrate a broader trend: AI is increasingly being integrated across military operations, from intelligence analysis and surveillance to targeting and autonomous systems, often to increase the speed, scale, and precision of military activities.

## A. Opportunities: Enhancing Efficiency, Precision, and Strategic Advantage

Artificial intelligence offers opportunities to improve how militaries process information, allocate resources, and make decisions. Its potential extends across tactical, operational, and strategic levels, with different applications addressing different operational needs. Some involve automating routine tasks, while others provide support in complex situations where personnel must assess large amounts of information under pressure. Although this report focuses on AI's role in operational decision-making,

7 Interviews 1, 4, 6, and 25; Andrew E. Kramer, "Enter the Killer Robots: The Ukrainian Forging the Future of Warfare," New York Times, May 15, 2026, <https://www.nytimes.com/2026/05/15/world/europe/mykhailo-fedorov-ukraine-ai.html>.

8 Andrew E. Kramer, "A Drone Killed Three Ukrainians. It Was Guided Entirely by A.I.," New York Times, August 24, 2026, <https://www.nytimes.com/2026/08/24/world/europe/russia-drones-autonomous-ai-kill-ukraine-war.html>.

9 Interviews 6, 9, and 12; U.S. Department of War, Artificial Intelligence Strategy for the Department of War, memorandum for Senior Pentagon Leadership, Commanders of the Combatant Commands, Defense Agency and DoD Field Activity Directors, January 9, 2026, <https://media.defense.gov/2026/Jan/12/2003855671/-1/-1/o/ARTIFICIAL-INTELLIGENCE-STRATEGY-FOR-THE-DEPARTMENT-OF-WAR.PDF>.

10 Katrina Manson, Project Maven: A Marine Colonel, His Team, and the Dawn of AI Warfare (New York: W. W. Norton & Company, 2026); Matt Mande and Gregory C. Allen, "What Is Maven Smart System, and What Does It Do?," Center for Strategic and International Studies, June 2, 2026, <https://www.csis.org/analysis/what-maven-smart-system-and-what-does-it-do>.

11 John Croxton, "Maven Is Becoming the Pentagon's Everything App," Defense One, September 9, 2026, <https://www.defenseone.com/technology/2026/09/maven-becoming-pentagons-everything-app/415882>.

12 Interview 28; Noor Hammad, "The Proliferation of AI-Enabled Military Technology in the Middle East," International Institute for Strategic Studies (IISS), April 2, 2026, <https://www.iiss.org/online-analysis/charting-middle-east/2026/04/the-proliferation-of-ai-enabled-military-technology-in-the-middle-east/>.

13 Austin Horng-En Wang et al., The People's Liberation Army's Perspectives on Artificial Intelligence: Highlighting Integration as Key to "Intelligitization" Goals (RAND Corporation), 2026, <https://www.rand.org/pubs/perspectives/PEA4574-1.html>; Gerald Mako, "China's Military Is Now Using AI to Plan Strike Operations," The Diplomat, August 7, 2026, <https://thediplomat.com/2026/08/chinas-military-is-now-using-ai-to-plan-strike-operations/>.

its benefits also extend to supporting functions. For example, AI-enabled software can assist with logistics planning and predictive maintenance, helping militaries anticipate equipment needs and use resources more efficiently.<sup>14</sup>

In intelligence analysis, AI can enable militaries to process vast amounts of data at speeds unattainable by humans alone.<sup>15</sup> By helping analysts identify relevant information and patterns across multiple sources, these systems can improve situational awareness and support a more informed assessment of the operational environment.<sup>16</sup> Such support can also reduce cognitive overload in data-intensive tasks, allowing personnel to devote more attention to contextual interpretation and higher-stakes decision-making.<sup>17</sup>

These capabilities can be particularly valuable in time-sensitive military environments, where personnel have limited time to interpret incoming information and determine how to respond. AI can contribute by processing information and presenting relevant assessments more rapidly than would otherwise be possible.<sup>18</sup>

Across these applications, AI's potential lies in complementing human capabilities. AI can provide speed, scale, and pattern recognition, while humans contribute contextual understanding and judgement. Effective human-machine interaction can bring these capabilities together, enabling military personnel to benefit from AI-generated analysis while retaining the ability to assess its relevance and limitations.

## B. Risks: Unintended Consequences and Ethical Dilemmas

The integration of AI into military systems introduces new opportunities alongside new dimensions of complexity and risk in modern warfare. AI integration can create significant operational, legal, ethical, and strategic risks, particularly when systems are deployed without adequate safeguards.<sup>19</sup> AI can amplify existing challenges, including errors in target identification and decision-making biases, while introducing new vulnerabilities associated with the reliability, predictability, and security of increasingly complex systems.

A key risk stems from biases, limitations, and errors that can arise in AI systems. Biases and systematic errors can arise from training data, model design, problem formulation, or the contexts in which AI systems are deployed, potentially producing discriminatory or ineffective outcomes.<sup>20</sup> AI systems may also produce behaviour that is inconsistent with the objectives, intentions, or constraints of their human operators.<sup>21</sup> This can occur when the objectives specified during development do not fully capture the operational goals or when a system identifies strategies that technically satisfy its task but produce unintended outcomes. For example, in August 2026, OpenAI reported that AI agents operating under reduced safeguards during internal cybersecurity tests in July had circumvented controls and accessed external systems in ways that were misaligned with their assigned tasks.<sup>22</sup>

Adversarial manipulation and data poisoning also stand out as critical concerns.<sup>23</sup> Adversarial manipulation can occur at different stages of the AI lifecycle. Adversaries may attempt to manipulate inputs at the point of use, for example through spoofed sensor data or deceptive imagery, or poison training or update datasets in ways that degrade a system's performance. These vulnerabilities are particularly significant in military settings, where adversaries have strong incentives to identify and exploit weaknesses in AI-enabled systems.<sup>24</sup>

14 Interviews 2 and 5.

15 Interviews 5, 7, 11, and 12.

16 Max Hunder, "Ukraine Sees 'Priceless' Digital Battlefield Data Trove as Key to West's Support," Reuters, August 27, 2025, <http://reuters.com/business/aerospace-defense/ukraine-sees-priceless-digital-battlefield-data-trove-key-west-s-support-2025-08-27>.

17 Interviews 2 and 7; Manson, *Project Maven*.

18 Interviews 2, 11, and 13; U.S. Department of Defense, "DOD Releases AI Adoption Strategy," November 2, 2023, <https://www.defense.gov/News/News-Stories/Article/Article/3578219/dod-releases-ai-adoption-strategy/>.

19 Interviews 3, 6, 8, 10, and 26.

20 Interview 4; Ninareh Mehrabi et al., "A Survey on Bias and Fairness in Machine Learning," *ACM Computing Surveys* 54, no. 6 (2021): 1–35, <https://doi.org/10.1145/3457607>.

21 Interviews 4 and 12.

22 OpenAI, "The Hugging Face Incident and the Road Ahead," August 26, 2026, <https://openai.com/index/hugging-face-incident-and-the-road-ahead/>.

23 Interview 30.

24 Andrew Lohn, *Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity* (Washington, DC: Center for Security and Emerging Technology, December 2020), <https://doi.org/10.51593/2020CA006>.

These risks are compounded by the opacity of many AI systems, particularly those relying on complex models whose internal decision-making processes are difficult to interpret. Such systems can make it more difficult for operators to understand, justify, or audit AI-generated recommendations.<sup>25</sup> If an AI-assisted system misclassifies a civilian structure as a military target, limited insight into how the decision was reached or why the error occurred can undermine accountability and erode confidence in the system. Moreover, insufficient transparency regarding an AI system's data can make it more difficult for operators to recognise when the system has been compromised or manipulated, including through adversarial inputs or poisoned training data.

AI systems may also perform unpredictably when deployed in environments that differ materially from the conditions in which they were developed and tested.<sup>26</sup> In rapidly changing or adversarial environments, incomplete information and compressed decision timelines can make failures more difficult for human operators to detect or correct before consequential actions are taken.

## C. Governance Challenges

The pace of military AI development is placing pressure on existing governance processes. This challenge is particularly acute where rapidly evolving commercial AI capabilities enter high-consequence security domains faster than existing governance processes can adapt.<sup>27</sup> Strategic competition can incentivise states to prioritise rapid development and deployment, creating tension between maintaining technological advantage and developing appropriate safeguards.<sup>28</sup> Ukraine illustrates how wartime pressures can complicate military AI governance, where immediate operational needs may take precedence over detailed domestic safeguards, even as it continues to engage in international initiatives on responsible military AI.<sup>29</sup> Without coordinated standards, the rapid proliferation of AI-enabled autonomous and decision-support systems risks producing inconsistent practices and weakening shared expectations around responsible use.

Military AI governance is spread across multiple international processes, including discussions within the Group of Governmental Experts on lethal autonomous weapons systems under the Convention on Certain Conventional Weapons (CCW), alongside initiatives such as the Responsible AI in the Military Domain (REAM) process.<sup>30</sup> Existing instruments also differ substantially in scope and legal force. NATO's AI Strategy establishes principles and governance objectives for responsible military AI, while the EU AI Act excludes AI systems used exclusively for military, defence, or national-security purposes from its scope.<sup>31</sup> National approaches differ substantially in the degree and form of constraints they place on military AI, reflecting different strategic priorities, legal frameworks, approaches to autonomy and risk tolerance. Disagreements also persist over the scope of international governance and the extent to which decisions about military AI should remain under national control.<sup>32</sup>

Even among allies that broadly endorse responsible AI principles, publicly stated approaches differ in how they frame human control, explainability, and accountability. These differences may complicate efforts to establish common expectations for multinational operations. The table below compares these approaches across Canada, the United States, Australia, the United Kingdom, and France.

25 Scott Sullivan and Iben Ricket, "Targeting in the Black Box," in CyCon 2024: Over the Horizon, 16th International Conference on Cyber Conflict (Tallinn: NATO Cooperative Cyber Defence Centre of Excellence, 2024), [https://ccdcoe.org/uploads/2024/05/CyCon\\_2024\\_Sullivan\\_Ricket-1.pdf](https://ccdcoe.org/uploads/2024/05/CyCon_2024_Sullivan_Ricket-1.pdf).

26 Interviews 7 and 18.

27 Stephen Herzog, Allison Berke, Yanliang Pan, William C. Potter and Douglas B. Shaw, "AI is changing biological and nuclear risks; governance must change accordingly," *Bulletin of the Atomic Scientists*, June 18, 2026, <https://thebulletin.org/2026/06/ai-is-changing-biological-and-nuclear-risks-governance-must-change-accordingly/>.

28 Interviews 8, 11, and 12.

29 Interviews 1 and 2.

30 Tobias Vestner and Simon Cleobury, "AI in the Military Domain: Multilateralism at a Crossroad," *Just Security*, July 30, 2026, <https://www.justsecurity.org/144022/ai-military-domain-multilateralism-crossroad/>.

31 NATO, "Summary of NATO's Revised Artificial Intelligence (AI) Strategy"; European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), Official Journal of the European Union L 2024/1689 (July 12, 2024), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

32 Interview 22.

## Table 1: Selected National Approaches to Military AI Governance

|                                      | Canada                                                                                                                                                                                                                          | United States                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Australia                                                                                                                                                                                                                                                                                                                                                                                                                                           | United Kingdom                                                                                                                                                                                                                                                                                                                                                                                                                    | France                                                                                                                                                                                                                                                                                                                                                                             |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Human Involvement and Control</b> | Human involvement should be calibrated to the risk and impact of the AI system. For low-risk applications, human involvement may be minimal; applications involving lethal force must retain a human in the loop. <sup>33</sup> | DoD's AI Ethical Principles emphasise human responsibility and system governability rather than prescribing a uniform level of human involvement. Personnel must exercise appropriate judgement and remain responsible for the development, deployment, and use of AI capabilities, while systems should be designed to detect and avoid unintended consequences and allow deployed capabilities to be disengaged or deactivated if they demonstrate unintended behaviour. <sup>34</sup> | Australian Defence applies risk-based, proportionate controls to military AI. Its policy states that control models should define how human control is enacted throughout a system's operations, with the appropriate model determined by the system's purpose, capabilities, and limitations. An Accountable Officer is responsible for ensuring that an appropriate control model is applied and that users are adequately trained. <sup>35</sup> | UK Defence takes a context-dependent approach to human control. The appropriate degree of autonomy and form of human involvement should be determined case by case according to system design, mission objectives, operational context, and risk. Human control need not always take the form of real-time supervision, but personnel must retain the ability to exercise human judgement over AI-enabled outcomes. <sup>36</sup> | France's Defence Ethics Committee recommends a context-dependent approach to human control, allowing commanders to modulate the degree of automation according to operational circumstances while preserving human decision-making in critical moments. It also calls for rigorous and proportionate assessment of AI systems according to their benefits and risks. <sup>37</sup> |

<sup>33</sup> Department of National Defence and Canadian Armed Forces, "Guiding Principles," The Department of National Defence and Canadian Armed Forces Artificial Intelligence Strategy, 2024, <https://www.canada.ca/en/department-national-defence/corporate/reports-publications/dnd-caf-artificial-intelligence-strategy/guiding-principles.html>.

<sup>34</sup> U.S. Department of Defense, "DOD Adopts 5 Principles of Artificial Intelligence Ethics," February 25, 2020, <https://www.defense.gov/Explore/News/Article/Article/2094085/dod-adopts-5-principles-of-artificial-intelligence-ethics/>.

<sup>35</sup> Australian Department of Defence, *Policy Settings for Responsible Use of Artificial Intelligence in Defence*, <https://www.defence.gov.au/sites/default/files/2026-03/Policy-Settings-for-Responsible-Use-of-Artificial-Intelligence-in-Defence-%5BOFFICIAL%5D.pdf>.

<sup>36</sup> UK Ministry of Defence, *Ambitious, Safe, Responsible*.

<sup>37</sup> French Ministry of the Armed Forces, "Comité d'éthique de la défense : pour un usage de l'IA éthique et proportionné," February 12, 2025, <https://www.defense.gouv.fr/actualites/comite-dethique-defense-usage-lia-ethique-proportionne>.

|                                                      | Canada                                                                                                                                                                                                                                                                                                                                   | United States                                                                                                                                                                                                                                                                                                  | Australia                                                                                                                                                                                                                                                       | United Kingdom                                                                                                                                                                                                                                                                                                                                                                                                                      | France                                                                                                                                                                                                                                                                                                                                  |
|------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Transparency, Explainability and Traceability</b> | DND/CAF states that AI-enabled decisions and outcomes should be explainable and transparent where possible. Decisions that augment human judgement should be justified and documented, with mechanisms to ensure that ultimate decisions are traceable and explainable and subject to appropriate accountability measures. <sup>38</sup> | DoD requires AI capabilities to be developed and deployed so that relevant personnel have an appropriate understanding of the technology, development processes, and operational methods, including transparent and auditable methodologies, data sources, design procedures, and documentation. <sup>39</sup> | The function and relationship between inputs and outputs of AI technologies must be traceable, and the technologies must be able to meaningfully accommodate human-machine interactions. <sup>40</sup>                                                          | UK Defence requires AI-enabled systems and their outputs to be appropriately understood by relevant personnel, with mechanisms for interpretation built into system design. Systems or their outputs should be auditable to an appropriate level, while information about what systems do, how they are used, and how outcomes are assured should be as transparent as possible within national-security constraints. <sup>41</sup> | French Defence states that AI tools integrated into military operations should be trustworthy, including by being traceable and explainable. French defence R&D priorities also identify explainable AI, including its acceptability in human interaction, alongside evaluation, verification, and qualification methods. <sup>42</sup> |
| <b>Accountability and Responsibility</b>             | DND/CAF requires appropriate accountability mechanisms for AI-enabled decisions and outcomes. Personnel involved in developing, procuring, and using AI must clearly understand their roles, responsibilities, and authority with respect to AI systems. <sup>43</sup>                                                                   | DoD personnel remain responsible for the development, deployment, and use of AI capabilities and are expected to exercise appropriate levels of judgement and care. DoD's Responsible AI framework also establishes governance structures for oversight and accountability. <sup>44</sup>                      | Australian Defence policy explicitly preserves individual accountability for AI-enabled decisions and outcomes. Responsible AI use must be underpinned by individual accountability and comply with domestic and international legal obligations. <sup>45</sup> | Human responsibility for AI-enabled systems must be clearly established, ensuring accountability for their outcomes. Responsibility and risk ownership must be defined throughout the system lifecycle, with no deployment or use without clear lines of responsibility and accountability. <sup>46</sup>                                                                                                                           | France's Defence Ethics Committee emphasises that responsibility must remain human. Chains of responsibility should be formally defined across command, control, and execution, including the responsibilities of industry, programmers, users, and decision-makers when AI contributes to military decisions. <sup>47</sup>            |

Note: The documents compared differ in scope and policy status. The table summarises publicly stated approaches and does not establish how consistently they are implemented in practice.

<sup>38</sup> Department of National Defence and Canadian Armed Forces, "Guiding Principles"; Department of National Defence and Canadian Armed Forces, "Line of Effort 3: Ethics, Safety, and Trust," The Department of National Defence and Canadian Armed Forces Artificial Intelligence Strategy, 2024, <https://www.canada.ca/en/department-national-defence/corporate/reports-publications/dnd-caf-artificial-intelligence-strategy/line-of-effort-3.html>.

<sup>39</sup> U.S. Department of Defense, "DOD Adopts 5 Principles of Artificial Intelligence Ethics."

<sup>40</sup> Australian Department of Defence, *Policy Settings for Responsible Use of Artificial Intelligence in Defence*.

<sup>41</sup> UK Ministry of Defence, *Ambitious, Safe, Responsible*.

<sup>42</sup> French Ministry of the Armed Forces, "Pas de cyberdéfense sans intelligence artificielle et pas d'IA sans cyberdéfense," November 20, 2024, <https://www.defense.gouv.fr/comcyber/actualites/pas-cyberdefense-intelligence-artificielle-pas-dia-cyberdefense-general-bonnetmaison-comcyber>; French Ministry of the Armed Forces, Artificial Intelligence: Defence Research and Innovation Orientation Document (DrOID 2023), 2023, <https://www.defense.gouv.fr/sites/default/files/aid/DrOID%202023-EN.pdf>.

<sup>43</sup> Department of National Defence and Canadian Armed Forces, "Line of Effort 3: Ethics, Safety, and Trust."

<sup>44</sup> U.S. Department of Defense, Responsible Artificial Intelligence Strategy and Implementation Pathway (Washington, DC: Department of Defense, June 2022), <https://www.ai.mil/Portals/137/Documents/Resources%20Page/DoD%20Responsible%20AI%20Strategy%20and%20Implementation%20Pathway.pdf>; U.S. Department of Defense, "DOD Adopts 5 Principles of Artificial Intelligence Ethics."

<sup>45</sup> Australian Department of Defence, *Policy Settings for Responsible Use of Artificial Intelligence in Defence*.

<sup>46</sup> UK Ministry of Defence, *Ambitious, Safe, Responsible*.

<sup>47</sup> French Ministry of the Armed Forces, "Comité d'éthique de la défense."

# Policy Recommendations and Governance Opportunities

**The following recommendations address how military institutions can benefit from AI while preserving meaningful human judgement and control.**

They combine *Lifecycle Policies* on procurement and training with *Use-Case Policies* on data analysis and operational decision support, followed by opportunities for Canada to advance responsible military AI governance with its allies.

## A. Lifecycle Policies

To ensure responsible and effective AI adoption, states should first establish foundational standards across the lifecycle, from development and procurement through deployment, operation, updating and retirement. The following *Lifecycle Policies* address the core principles of transparency, accountability, and human-AI collaboration, supporting reliable AI systems that can be appropriately understood, scrutinised, and governed.

### Policy 1: Transparency and Explainability Standards for Procurement

#### Context

As many military AI capabilities are developed or supplied by private-sector actors,<sup>48</sup> states need to engage with industry and establish clear requirements for the systems they acquire. Commercial incentives and proprietary development practices do not necessarily align with military operational, legal and accountability requirements. Without sufficient transparency, military personnel may struggle to assess the basis and limitations of AI outputs, while

commanders and auditors may be unable to reconstruct how those outputs informed a decision.<sup>49</sup> This can allow errors to go undetected and make it harder to establish accountability when failures occur. Procurement represents a critical point at which high-level commitments to responsible military AI can be translated into concrete and testable requirements for suppliers.<sup>50</sup>

#### Policy Recommendation

To mitigate these risks, states should establish risk-proportionate transparency, explainability, traceability and auditability requirements for military AI through procurement. Procurement requirements should extend beyond delivery. Suppliers should keep documentation up to date, support audits, and report incidents. Significant changes to the system or how it is used should trigger a new assessment.

Where operationally feasible, inherently interpretable models should be preferred. Where more complex models are necessary, stronger assurance, testing and traceability requirements should apply, alongside appropriate explanation mechanisms and record-keeping sufficient to reconstruct relevant decisions. Additionally, clear and usable user-interface and user-experience (UI/UX) design is critical in military AI systems, with interfaces that communicate system status, uncertainty, limitations, and any degradation in capability, as well as the distinction between AI outputs and authorised decisions.<sup>51</sup> Alongside these measures, where operationally appropriate, embedding or forward-deploying engineers with military users could provide developers with a

<sup>48</sup> Alexander Blanchard, Vincent Boulanin, and Laura Bruun, "Mapping the Military AI Industry," Stockholm International Peace Research Institute (SIPRI), 2026, <https://www.sipri.org/commentary/topical-background/2026/mapping-military-ai-industry>.

<sup>49</sup> Interviews 11 and 21.

<sup>50</sup> Netta Goussac and Vincent Boulanin, Responsible Procurement of Military Artificial Intelligence (Stockholm: Stockholm International Peace Research Institute, February 2026), <https://doi.org/10.55163/YOLG1827>.

<sup>51</sup> Interview 22; Rupert Barrett-Taylor, Anna Knack, and Natasha Kerner, Resilient Defence AI: Sustainable and Operationally Effective Capabilities by Design (London: The Alan Turing Institute, May 2026), <https://www.turing.ac.uk/news/publications/resilient-defence-ai-sustainable-and-operationally-effective-capabilities-design>.

deeper understanding of the challenges operators face when using AI systems in real-world scenarios.<sup>52</sup>

## Addressing Trade-Offs

There can be a trade-off between explainability and predictive performance. Simpler, highly interpretable models are often easier for humans to understand, whereas complex deep-learning models can offer strong performance in analysing unstructured data such as imagery, text, and signals, but are generally more difficult to interpret.<sup>53</sup> Requirements should reflect both the consequences of error and the time available for assessment:

- High-consequence uses should require strong traceability and sufficient information for personnel to assess the basis and limitations of AI outputs.
- Where time limits the detail that personnel can assess, greater emphasis should also be placed on testing before deployment, system monitoring, and post-action auditability.

## Why This Matters

Transparency and explainability are important enablers of accountability and can support compliance with applicable legal obligations.<sup>54</sup> When AI systems provide meaningful information about the basis, confidence, and limitations of their outputs, operators can better assess and, where appropriate, challenge, disregard or, where applicable, override AI recommendations. This can also enable auditors and commanders to reconstruct decisions, identify errors, and assess whether the use of AI was consistent with applicable legal and operational requirements. However, transparency and explainability cannot themselves determine whether an AI-supported decision is legally compliant. Applying principles such as distinction and proportionality involves highly contextual assessments that explainability alone cannot resolve. Moreover, an AI system may produce an explainable output while still relying on problematic assumptions or data. Explainability therefore enables greater scrutiny, but does not itself guarantee legal compliance.

<sup>52</sup> Interview 31.

<sup>53</sup> Defense Advanced Research Projects Agency (DARPA), "Explainable Artificial Intelligence (XAI)," <https://www.darpa.mil/research/programs/explainable-artificial-intelligence>.

<sup>54</sup> Interview 11.

## Policy 2: Human-Machine Teaming (HMT) Training

### Context

As AI systems become more widely used in military operations, personnel increasingly encounter AI-generated analysis and recommendations. However, this collaboration introduces new challenges, including automation bias, over-reliance, and the tendency to attribute human-like understanding or judgement to AI systems.<sup>55</sup> Without specialised training, the operational value of even technically capable AI systems can be undermined. HMT training is therefore a critical investment to help ensure that AI systems appropriately support, rather than displace, human judgement and responsibility.

### Policy Recommendation

Militaries should develop specialised HMT training programmes that go beyond technical proficiency to address the psychological, ethical, and operational dimensions of working with AI-driven systems. Training should be tailored to personnel's roles and levels of decision-making, extending beyond system operators to commanders and senior leaders who rely on AI-generated analysis or recommendations.

Training should help personnel recognise when to rely on AI outputs, when to verify them, and when to challenge or disregard them. It should also address how their own biases, assumptions, and experiences influence their interpretation of those outputs. These skills should be supported by organisational cultures, incentives, and procedures that encourage critical assessment rather than automatic deference to AI or its blanket rejection.

### Implementation Approach

Militaries should integrate HMT training throughout personnel development rather than treating it as a one-time technical qualification. To achieve this, they should:

- Require developers or suppliers to provide training on the capabilities, limitations, intended uses, and known failure modes of their systems. This should be supplemented by military-led training that

<sup>55</sup> Interviews 4, 12, 18, and 26; Jonathan P. Wong et al., *One Team, One Fight, Volume I: Insights on Human-Machine Integration for the U.S. Army* (Santa Monica, CA: RAND Corporation, 2025), [https://www.rand.org/pubs/research\\_reports/RRA2764-1.html](https://www.rand.org/pubs/research_reports/RRA2764-1.html).

places system use within operational contexts and command structures.

- Conduct scenario-based and red-teaming exercises that expose personnel to system failures, adversarial manipulation, conflicting or incomplete information, and conditions that differ from those in which the system was tested. Exercises should challenge both the AI system and the human-machine team, requiring personnel to question recommendations and make decisions when AI outputs conflict with other evidence.<sup>56</sup>
- Assess personnel's ability to identify unreliable outputs, exercise independent judgement, and recognise when a system should not be relied upon, alongside their technical proficiency, as part of relevant qualification processes.
- Introduce training when personnel first work with AI-enabled systems and reinforce it through regular exercises and assessments. Training should be updated as systems and operational uses evolve, drawing on feedback from personnel and lessons from exercises and deployments.

## Why This Matters

By prioritising HMT training, militaries can better prepare personnel to work effectively with AI-enabled systems while reducing the risks of inappropriate reliance, automation bias, and uncritical acceptance of AI-generated outputs. Effective training can strengthen the quality and resilience of human-AI decision-making by helping personnel understand both the capabilities and limitations of the systems they use.<sup>57</sup>

## B. Use-Case Policies

Military AI is not one-size-fits-all: its risks and governance requirements vary depending on the function, operational context, stakes, and degree of autonomy involved.<sup>58</sup> While *Lifecycle Policies* establish foundational standards for responsible AI adoption, *Use-Case Policies* address how AI should be used for different functions within military decision-making. The following policies distinguish between the use of AI to analyse and prioritise operational information, and its use to support choices about operational action. The two functions raise different questions about reliance,

human judgement, and decision authority, but may overlap within a single system and both influence high-consequence decisions.

---

## Policy 3: AI-Enabled Data Analysis for Decision Support

### Context

Modern military operations generate vast amounts of data, from satellite imagery and signals intelligence to real-time sensor feeds, cyber data, and intelligence reports. The scale, speed, and complexity of these information streams can exceed personnel's ability to process the information efficiently.<sup>59</sup> AI-enabled analytical systems can process and correlate large datasets at machine speed, identifying patterns, anomalies, relationships, and competing interpretations that might otherwise be missed.<sup>60</sup> By automating data-intensive analysis and prioritisation, AI can reduce cognitive overload and allow personnel to focus their attention on contextual interpretation, judgement, and high-stakes decisions.

### Policy Recommendation

Militaries should integrate AI-enabled analytical tools into operational workflows to help personnel process, correlate, and prioritise large volumes of information at the required speed and scale. These tools should identify relevant patterns, anomalies, relationships, and gaps, helping analysts determine what requires further investigation or contextual interpretation. Where appropriate, they may also generate alternative hypotheses for analysts to examine.

AI-generated assessments should remain inputs into a broader intelligence and analytical process. Personnel should assess these outputs against other available information and interpret them in their operational context. Systems should communicate uncertainty, limitations, and sources of information wherever feasible so that analysts can identify when additional verification is required.

<sup>56</sup> Interviews 31 and 32.

<sup>57</sup> UK Ministry of Defence, Defence Artificial Intelligence Strategy, 2022, <https://www.gov.uk/government/publications/defence-artificial-intelligence-strategy/defence-artificial-intelligence-strategy>.

<sup>58</sup> Herwin Meerveld, Lonneke Peperkamp, Marie Šafář Postma, and Roy Lindelauf, "Operationalising Responsible AI in the Military Domain: A Context-Specific Assessment," *Ethics and Information Technology* 27 (2025): 48, <https://doi.org/10.1007/s10676-025-09865-y>.

<sup>59</sup> NATO Defence Innovation Accelerator for the North Atlantic (DIANA), "Data Overload to Operational Clarity: Supporting Decision Advantage through Data-Assisted Decision Making," July 14, 2026, <https://www.diana.nato.int/connect/data-overload-to-operational-clarity-supporting-decision-advantage-through-dataassisted-decision-making.html>.

<sup>60</sup> Interviews 5, 9, 11, 12, 30, and 32.

## Implementation Approach

To implement AI-enabled data analysis effectively, militaries should:

- Integrate AI with existing command-and-control and intelligence systems, enabling it to draw appropriately on relevant data sources, including satellite imagery, signals intelligence, sensor networks, cyber data, and other intelligence feeds.
- Design systems to support critical assessment rather than automation bias. Interfaces should make clear that AI outputs are analytical assessments, not authoritative decisions, and should provide operators with sufficient information to assess their reliability.
- Provide operators with relevant information about uncertainty, confidence, limitations, and data provenance, enabling them to identify situations in which AI outputs require additional verification.<sup>61</sup>
- Ensure that unusual results, conflicting information, and low-confidence assessments are clearly flagged for further review by analysts.
- Regularly test and audit AI systems for accuracy, reliability, bias, model degradation, and vulnerabilities to adversarial manipulation, including data poisoning and spoofing.

## Why This Matters

The scale and speed of information generated during modern military operations make AI-enabled data analysis a potentially transformative capability. Used appropriately, AI can reduce the cognitive burden placed on military personnel, accelerate the identification of relevant information, improve situational awareness, and allow human expertise to be concentrated on decisions where contextual understanding and judgement are most important.

## Policy 4: AI-Enabled Operational Decision Support

### Context

Military decision-making often takes place under conditions of uncertainty, incomplete information, rapidly changing circumstances, and significant time pressure. By processing large volumes of information more quickly than personnel could unaided, AI-DSS can help decision-makers assess operational

options within compressed timelines.<sup>62</sup> In these environments, AI-DSS can build on analytical outputs by helping personnel compare interpretations, assess operational options, and develop possible courses of action. This can be particularly valuable where the volume or speed of incoming information makes it difficult for personnel to assess all relevant factors within the available time.

## Policy Recommendation

Militaries should integrate AI-DSS into operational decision-making where these systems can improve the speed, scale, and relevance of decision support available to personnel. AI-generated outputs should, however, remain advisory rather than authoritative: personnel should retain the authority to contextualise, challenge, reject, or disregard AI-generated recommendations. While AI may provide information and recommendations relevant to the use of lethal force, such decisions should remain with an appropriately authorised human decision-maker.

The level of human involvement need not be identical across all uses of AI-DSS. In lower-risk, well-bounded applications, systems may be given greater autonomy to process information, prioritise options, or carry out predefined supporting functions without requiring human approval at every intermediate stage. As uncertainty, potential consequences, or operational complexity increase, stronger human scrutiny and verification should be required.<sup>63</sup>

Personnel should have sufficient information to understand the basis, limitations, and uncertainty of AI-generated recommendations, alongside the authority and practical ability to reject or disregard them. As AI systems generate and prioritise recommendations at greater speed and scale, however, human review can itself become a bottleneck, reducing the time available for meaningful scrutiny before action is taken.<sup>64</sup> Decision-making procedures should allow sufficient time for personnel to assess AI-generated outputs before acting. Where this is not possible, the fallback procedures outlined below should apply.

62 Interviews 5, 11, 15, and 21; UK Ministry of Defence, *Ambitious, Safe, Responsible*.

63 Meerveld et al., "Operationalising Responsible AI in the Military Domain"; Canadian Nuclear Safety Commission, UK Office for Nuclear Regulation, and U.S. Nuclear Regulatory Commission, "Considerations for Developing Artificial Intelligence Systems in Nuclear Applications" (September 2024), <https://www.nrc.gov/docs/ML2424/ML24241A252.pdf>.

64 Alison Killing, John Thornhill, Charles Clover, Chris Campbell, and Fabrice Deprez, "The Era of AI Warfare Has Arrived," *Financial Times*, September 17, 2026, <https://www.ft.com/content/686429c0-daf3-42a5-9b7c-7ff06eb291ef>.

61 Interviews 4, 31, and 32.

## Implementation Approach

To support effective AI-assisted operational decision-making, militaries should:

- Ensure that interfaces clearly distinguish AI-generated recommendations from authorised decisions and communicate relevant uncertainty and limitations.
- Design systems to prioritise information relevant to the decision when time is limited. The volume and pace of recommendations should not overwhelm personnel's ability to assess them.
- Ensure that low-confidence assessments, unusual results, and conflicting information are clearly flagged for additional review, particularly where they may influence consequential decisions.
- Establish clear procedures for when personnel cannot adequately assess an AI recommendation, such as seeking further information, referring the decision to a commander, or limiting use of the system.

## Why This Matters

AI can strengthen military decision-making by helping personnel synthesise complex information, compare operational options, and consider possible courses of action under demanding operational conditions. Its value should not, however, be measured solely by whether it enables decisions to be made faster. The objective is to improve the quality of military decision-making while preserving human judgement and keeping decision authority within the chain of command. This approach allows militaries to make extensive use of AI-DSS, including in time-sensitive and high-consequence contexts, without allowing AI-generated assessments to replace human authority.

## C. Governance Opportunities

The policies above focus on how military AI should be governed within national defence institutions. Beyond implementing these measures domestically, Canada can help shape the broader international framework for responsible military AI. The following opportunities build on the policy recommendations, identifying ways for Canada to work with its allies on common standards, human-machine teaming training, and research.

### OPPORTUNITY 1

#### Advance 'Responsible by Design' Standards for Military AI

Procurement provides an opportunity to incorporate legal, ethical, and operational safeguards into military AI before deployment.<sup>65</sup> Canada could build on this by working with its allies to develop common procurement standards that support a responsible-by-design approach.

For Canada, this could mean establishing common requirements for transparency and explainability, human oversight, testing and evaluation, and resilience to adversarial manipulation. By working with NATO Allies to develop and promote such standards, Canada could also support greater interoperability by ensuring that AI-enabled systems are developed and assessed according to compatible requirements. This would provide Canada with a practical avenue to contribute to the development of responsible military AI governance across the Alliance.

<sup>65</sup> Goussac and Boulanin, *Responsible Procurement of Military Artificial Intelligence*.

## OPPORTUNITY 2

### Lead HMT Training for NATO

As AI systems become increasingly integrated into military operations, HMT training can help personnel work effectively with AI while reducing automation bias and over-reliance. It should also prepare them to recognise signs that outputs may be unreliable because of system failures or adversarial manipulation.<sup>66</sup> Canada could take a leading role in developing and promoting common HMT training approaches across NATO, helping prepare personnel for the cognitive, ethical, and operational demands of working with AI-enabled systems.

These programmes should assess human readiness alongside technical readiness, including whether personnel understand a system's capabilities and limitations and can judge when its outputs should be questioned.<sup>67</sup> Training should develop critical thinking and operational adaptability through exercises that require personnel to challenge, verify, or disregard AI recommendations. Canada could support a NATO HMT training initiative or hub that brings together Allied expertise, develops shared exercises and assessment approaches, and exchanges lessons across the Alliance. This cooperation could support interoperability by building a shared understanding of how AI-enabled systems should be assessed and used in multinational operations, while allowing training to be adapted to national requirements.

<sup>66</sup> Interviews 7 and 21.

<sup>67</sup> Interview 24.

## OPPORTUNITY 3

### Support Explainable, Traceable, and Auditable AI Research

To address the challenges posed by opaque and complex AI systems, Canada can support research and development aimed at making military AI more explainable, traceable, and auditable. The level and form of explainability required should depend on how a system is used and the consequences associated with its outputs, rather than requiring all AI systems to be fully interpretable. The importance of this issue has already been recognised within Canadian defence thinking. Canadian Army analysis has cautioned against 'black box' AI in decision support and emphasised that AI-generated recommendations should be sufficiently explainable for human decision-makers to understand and assess them.<sup>68</sup>

This is particularly important in high-consequence applications where operators and commanders need sufficient information to understand the basis, limitations, and uncertainty of AI-generated outputs, and to determine when those outputs require additional scrutiny. Research should focus on improving explainability, and on developing mechanisms that allow AI outputs, system behaviour, and relevant decision processes to be effectively scrutinised and reconstructed. By supporting research and common standards in these areas, Canada can contribute to the development of military AI that combines operational capability with the transparency and traceability needed to support effective oversight, accountability, and meaningful human control.

<sup>68</sup> Geoffrey Priems and Peter Gizewski, "Leveraging Artificial Intelligence for Canada's Army: Current Possibilities and Future Challenges," *The Canadian Army Journal* 19, no. 2 (2021): 40-51.

# Conclusion

Military AI offers significant opportunities to improve operational effectiveness, from processing large volumes of information to enhancing situational awareness and supporting operational decision-making.

At the same time, these capabilities introduce important risks, including bias, unreliable outputs, adversarial manipulation, system opacity, and inappropriate reliance on AI-generated recommendations. As AI systems become increasingly capable and integrated into military operations, the central governance challenge is not simply whether humans remain involved, but whether their involvement enables meaningful judgement, oversight, and decision authority.

This report has argued that a uniform approach to HMI is insufficient. The appropriate form of human involvement should instead reflect the operational context, the level of uncertainty, the consequences of error, and the time available for human assessment. AI can provide substantial value by processing information at a scale and speed beyond human capacity and by supporting personnel in complex and time-sensitive decision-making. AI-generated outputs should support rather than replace human decision-making.

The four policies proposed in this report translate this approach into practical governance measures. Transparency and explainability standards help ensure that military AI systems can be appropriately understood, scrutinised, and audited. HMT training can prepare personnel to critically assess AI outputs and recognise when they should be relied upon, challenged, or disregarded. AI-enabled data analysis can strengthen situational awareness by helping personnel process and prioritise large volumes of operational information. AI-DSS can further support operational choices by synthesising that analysis, comparing options, and generating possible courses of action, while preserving human authority over consequential decisions.

For Canada, these policies also create opportunities to contribute to responsible military AI governance beyond the national level. By advancing responsible-by-design approaches, supporting common HMT training and exercises across NATO, and investing in research on explainable, traceable, and auditable AI, Canada can help strengthen Allied interoperability and contribute to greater convergence in how AI-enabled systems are designed, assessed, and used. These efforts would complement NATO's broader work to operationalise its Principles of Responsible Use through standards, testing and evaluation, and the development of an AI-ready workforce.<sup>69</sup>

Ultimately, effective military AI governance should neither reject AI because of its risks nor assume that human presence alone guarantees meaningful control. The objective should be to integrate AI in ways that make military decision-making more informed, effective, and resilient while ensuring that personnel retain the information, authority, and capacity necessary to exercise meaningful judgement. A context-sensitive and risk-proportionate approach to HMI offers a way to capture the operational benefits of AI while preserving human responsibility and authority over decisions where the consequences of error are greatest.

69 NATO, "Summary of NATO's Revised Artificial Intelligence (AI) Strategy."

**BASIC promotes meaningful dialogue amongst governments and experts in order to build international trust, reduce nuclear risks, and advance disarmament.**

---

**The British American Security  
Information Council (BASIC)**

Work + Play  
111 Seven Sisters Rd  
Finsbury Park  
London N7 7FN

---