NATO
C2COE

HUMAN OVERSIGHT IN AI-DRIVEN DEFENCE

At what positions do we need the Human in the Loop?

Marcel Scherrenburg MSc MBA BEng
1 Civil and Military Interaction Command Netherlands

Introduction

Picture the battlefield of today: Artificial Intelligence (AI) enables military leaders to analyse information rapidly and direct weaponry with exceptional accuracy. This technology is not coming, it is already here. From intelligence gathering to strategic planning, AI is transforming the operations of modern armed forces. Consider, for example, an AI-powered drone recognizing targets. Its accuracy is impressive, but can we really rely on it to make choices about human lives? This goes beyond precision, it involves judgement. When an AI highlights a potential threat in a busy area, a human operator must weigh civilian safety, international regulations, and mission objectives. These considerations exceed simple data; they involve complex moral judgments and require human insight and ethical understanding.

That is why prevailing military doctrine ensures humans are still "in the loop" for most activities. While AI acts as a formidable tool, crucial choices rest with human operators. Imagine it as a collaboration where each contributes their unique strengths: the speed and accuracy of AI along with the moral reasoning and judgment of humans.

The essential questions are deceptively simple yet deeply significant: What decisions can we entrust to AI, and in which areas should humans keep authority? How can we effectively integrate AI's potential with military capabilities while keeping humanitarian values and a clear sense of accountability? What are the dynamics between human judgment and artificial intelligence in critical situations. In other words, at what positions do we need the Human in the Loop?

1. AI in defence

Artificial Intelligence (AI) is essential in modern-day military operations, facilitating improved decision-making, autonomous systems, and real-time data insights. Military forces globally are adopting AI to improve various aspects, from strategic decisions to combat strategies. Today's AI can analyse intelligence more swiftly than ever and control weapons with remarkable accuracy, envisioning targeting systems that enhance airstrike precision to unprecedented levels.

While AI is transforming modern warfare, it raises important concerns. AI shows substantial potential in military settings, but there is a struggle to the level of autonomy to give to the technology and entrust complete control to machines. AI systems use algorithms to analyse data, find patterns, solve problems, and make decisions. However, these models can become so complex that their creators struggle to explain how decisions are made. This "black box" issue raises transparency concerns, making it hard to understand or trust AI results, and raising questions about accountability and bias in AI decision-making.

Defence applications present unique challenges that do not exist in the civilian world. Every action must align with international humanitarian law, keep clear ethical standards, and follow a transparent chain of command. These are not just bureaucratic wishes; they are fundamental safeguards that protect human lives and uphold our values. Unlike civilian AI applications, where errors may lead to minor issues like a misdirected package or a buggy app, military AI involves critical situations where the stakes are life and death.

The military cannot view AI as just another piece of technology to boost efficiency. We need to carefully consider each step towards automation and weigh it against our ethical duties and the critical importance of human judgment in wartime decisions. The crucial aspect is not merely recognizing ethical dilemmas, but understanding how the use of AI in military operations could align with fundamental values and principles of humanitarian law. It is not just about whether AI can perform these tasks, but whether we want to have human oversight in AI-driven defence and at what positions do we need the human interventions to balance the potential of AI with our ethical responsibilities and humanitarian values?

1.1 A bit of context on AI

In modern defence contexts, using AI is essential for enhancing operational effectiveness, supporting decision-making, and boosting situational intelligence. The following are key principles and applications from a defence perspective:

- Machine Learning (ML) systems can learn from data (Goodfellow, Bengio, & Courville, 2016) , as does Deep Learning, a subset of ML that utilises neural networks to model complex patterns (LeCun, Bengio, & Hinton, 2015), this technology is crucial for predictive analytics. For instance, machine learning can detect threats or anomalies by analysing extensive data from sensors, surveillance, and communications (Defense Science Board, 2016). It aids in cybersecurity by detecting malicious activities and network intrusions (Sommer & Paxson, 2010) In a military context; ML algorithms can analyse data from multiple sources to predict potential enemy movements and enhance battlefield awareness.

- Narrow AI refers to algorithms that specialize in specific, well-defined tasks within a particular context, such as image recognition, language translation, or playing chess. It excels in its programmed domains, providing high efficiency and accuracy (Bostrom, 2014). Generative AI, a type of narrow AI, creates text or images based on its training data. In the military, narrow AI aids in automated threat detection in surveillance footage, enhancing speed and accuracy. Generative AI is useful for creating detailed simulation environments or generating strategic plans. By focusing on specialized tasks, narrow AI significantly enhances military operations and improves situational awareness and decision-making.

-

- Artificial General Intelligence (AGI) refers to a type of intelligence capable of performing any intellectual task that a human can. Unlike narrow AI, AGI has broad cognitive abilities, allowing it to understand, learn, and apply knowledge across diverse fields. This versatility mirrors human intelligence, making AGI a significant focus in discussions about potential superintelligence and the existential risks associated with AI advancement (Bostrom, 2014). While AGI is still mostly speculative, its potential applications in the military could revolutionize decision-making, strategy planning, and autonomous operations.

- Machine Learning-based Object Detection, or Computer Vision, enables real-time object recognition and target tracking using drones, satellites, or surveillance systems. This technology is crucial for detecting enemy assets, overseeing borders, and directing autonomous weapons (Szeliski, 2011). In reconnaissance missions, computer vision helps drones spot and follow enemy vehicles or personnel, delivering vital information to ground forces.

- Autonomous Systems, AI-powered drones and non-crewed ground or underwater vehicles are employed for reconnaissance, logistics, and combat operations (Scharre, 2018)AI allows these systems to operate remotely with minimal human intervention, improving response times and reducing risks to personnel (Morgan, et al., 2020) (Sawant, et al., 2023). For instance, autonomous drones can conduct surveillance over hostile territories, relaying real-time data back to command centres without putting human lives at risk.

- Natural Language Processing (NLP) plays a crucial role in intelligence gathering by enabling the processing and analysis of communications, documents, and social media across various languages. This technology efficiently extracts relevant information from vast datasets, aiding military staff in responding to security threats. By automating the analysis of multilingual content, NLP improves the speed and accuracy of intelligence assessments, contributing to more informed decision-making in defence operations (Stahl, 2021). For example, NLP can help translate intercepted communications in real-time, offering valuable insights into enemy plans.

2. Military applications

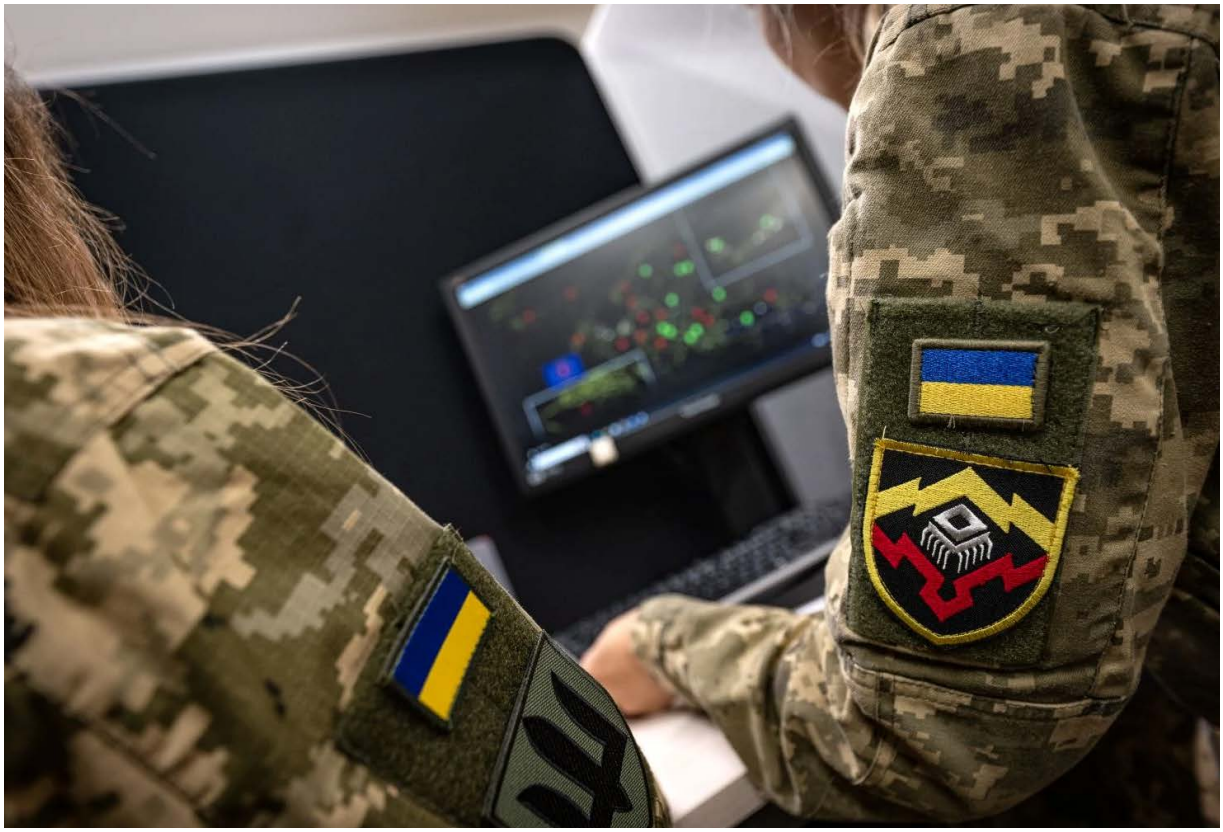
By introducing AI into military systems, armed forces can use advanced technologies to improve processes, automate complex tasks, and enhance situational awareness. These innovations not only streamline operations but also ensure greater precision and effectiveness in mission execution. Advancements in AI will lead to the development of highly sophisticated systems which can significantly impact both lethal and non-lethal actions.

Modern military operations largely depend on narrow AI. Over six years of expert discussions at NATO's Command and Control Centre of Excellence (NATO C2COE) have shown that these systems are transforming modern warfare, especially in three main areas.

First, AI has revolutionized situational awareness and data integration. On today's battlefield, commanders face an overwhelming flood of information from satellites, drones, and sensors. AI processes this massive data stream in (near) real-time, identifying patterns, anomalies and threats and predicting enemy actions with unprecedented clarity. This capability has become crucial in modern warfare, where traditional analysis methods simply cannot keep pace with the volume and speed of incoming data.

The second transformation comes through AI's analytical capabilities. These systems do not just process current data – they dive deep into historical information, uncovering patterns and insights that help explain past events and predict future developments. This descriptive analytical power helps to make sense of adversaries' activities and develop most-likely and most-dangerous scenarios, turning data into actionable military intelligence.

The third area where AI has significant impact is in military decision-making through advanced



support systems. Commanders now have tools capable of simulating various scenarios, conducting enhanced wargaming within synthetic environments, evaluating potential strategies, and assessing their probability of success. These systems mitigate human bias by generating data-driven insights, thus ensuring decisions are based on factual analysis rather than personal judgment.

In the dynamic landscape of modern military operations, AI is revolutionizing both tactical and strategic capabilities. This chapter explores, without being exhaustive, the diverse ways in which AI is being integrated into military applications, significantly enhancing efficiency, accuracy, and safety across multiple domains.

Improving Logistics and Resource Management - is crucial for military operations. AI significantly enhances these processes, ensuring that NATO forces are still prepared and well-equipped. By refining supply chains, AI predicts demand, manages inventory, and identifies efficient transportation routes. (COL Lacroix, 2023) , (MAJ Tilley, 2024) Machine learning algorithms analyse factors such as historical demand patterns, weather conditions, and geopolitical events to forecast future needs and adjust supply chains accordingly. For instance, during military exercises, AI can predict the consumption of critical supplies like fuel and ammunition, ensuring timely delivery to the right locations.

De-mining Operations - AI revolutionizes demining operations by accurately finding and mapping the locations of landmines. Machine learning algorithms analyse data from sources like satellite imagery and ground sensors to detect mines with high precision. For example, in conflict zones, AI speeds up the clearance of unexploded mines, reducing the risk to human deminers and civilians (Vivoli, Bertini, & Capineri, 2024). AI-driven drones and robotic systems already navigate hazardous areas, detect mines, and even neutralize them autonomously (Vivoli, Bertini, & Capineri, 2024) This decreases human exposure to danger and accelerates the demining process, ensuring safer and more effective military operations (Vivoli, Bertini, & Capineri, 2024).

A notable example of AI in demining operations is the collaboration between the US-based company Palantir and the Ministry of Economy of Ukraine. This partnership aims to clear 80% of potentially mined land within ten years by using advanced AI-enabled software. Palantir's Artificial Intelligence Platform (AIP) will be employed for decision-making in demining efforts (Palantir Technologies, 2024).

Predictive Maintenance - AI-powered predictive maintenance systems analyse data from sensors embedded in military equipment to predict potential failures and schedule maintenance activities proactively. By finding issues before they become critical, these systems reduce downtime and extend the lifespan of valuable assets. For instance, AI monitors aircraft engines, detecting subtle signs of wear and tear that may show an impending failure. Maintenance teams then take corrective actions before problems escalates, ensuring that planes are still operational and mission ready. ML can predict asset failures and improve the Mean Time Between Failures (MTBF), enabling new maintenance strategies to reduce integrated logistic support costs (NATO Science & Technology Organization., 2023). This minimizes the risk of shortages and supports operational readiness.

Enhancing Cybersecurity - As NATO relies increasingly on digital technologies, cybersecurity becomes a critical concern. AI can play a vital role in enhancing cybersecurity by detecting and mitigating threats in real-time. (Sarker, Furhad, & Nowrozy, 2021) (IBM, n.d.). AI-powered cybersecurity systems can analyse network traffic, find suspicious activities, and respond to threats in real-time. Machine learning algorithms can detect patterns indicative of cyberattacks, such as unusual login attempts or data transfers, and start automated responses to mitigate the threat. For example, an AI-based intrusion detection system can find a phishing attack by recognizing the unique characteristics of phishing emails. The system can then block the email and alert security teams, preventing the attack from spreading.

3. Decision-making and Command and Control

In modern military operations and decision-making, AI offers benefits across multiple levels, each uniquely improved by advancements in technology. These levels are tactical, operational, and strategic. Each level has specific characteristics, capabilities, and requirements that are crucial for effective military operations.

At the tactical level, decisions are centred on the immediate needs of the battlefield, requiring swift and adaptive responses to dynamic and often unpredictable combat situations. The primary focus here is on interpreting real-time information to make informed adjustments to tactics and plans. AI plays a crucial role by analysing vast amounts of data to offer actionable insights that enhance situational awareness. For instance, AI systems can rapidly assess the battlefield, find emerging threats, and suggest best suited responses. This capability ensures that commanders can make quick, effective decisions that directly affect mission success. Tactical decisions demand rapid responses to dynamic and often unpredictable combat situations. The key characteristic of this level is the need for immediate, real-time data analysis to adapt to changing circumstances on the battlefield. AI's real-time data processing capabilities are invaluable here; for instance, computer vision algorithms can quickly find threats and optimize battlefield responses. This enhances situational awareness and provides crucial tactical advantages during engagements. Requirements for this level include advanced AI systems capable of processing and analysing vast amounts of data swiftly and accurately.

The operational level involves the coordination of resources and missions, bridging the gap between high-level strategic goals and on-the-ground tactical actions. The primary characteristic of the operational level is its focus on logistical efficiency and resource allocation. AI supports operational decisions by processing vast amounts of logistical and geospatial data, enabling military planners to allocate resources effectively and optimize supply chains and troop movements. For example, AI-assisted mission planning systems can simulate various scenarios to help commanders make informed decisions. The requirements for this level include comprehensive data integration systems and AI tools designed for operational planning and resource management (Cole, Howard, Latiff, Lucas, & Roy, 2024)

Strategic decision-making shapes the long-term objectives and alliances of military operations. The defining characteristic of this level is its focus on overarching goals and the broader geopolitical landscape. AI plays a pivotal role by modelling geopolitical scenarios and assessing risks, thus providing strategic foresight. By evaluating the potential impacts of alliances or conflicts on national security, AI aids in planning and decision-making at the highest levels. The requirements for this level include advanced modelling tools and AI systems capable of synthesizing complex geopolitical data to support long-term strategic planning.

3.1 Binary decisions

In military and defence contexts, digital and autonomous technologies have made it possible to automate decisions based on real-time feedback from systems. When each part of a complex, integrated system shows a "go" or "green" signal on a dashboard, it means the system has met certain operational standards. This confirmation often triggers autonomous actions without needing human input, enabling rapid responses in critical situations where quick decisions are essential.

These decision support systems rely on automation, following pre-programmed actions that are activated in a set sequence. Typically, automated decision support uses rule-based algorithms that produce the same outcome every time for a given input. Command-and-control centres commonly employ this approach to reduce decision-making delays, enhance response times, and ease the cognitive burden on operators. For example, if all subsystems in a missile defence system show a green signal

when detecting an incoming threat, the system can autonomously activate and respond within established parameters. (Scharre, 2018). This type of automation improves operational response times and increases the ability of systems to handle multiple simultaneous inputs, which would overwhelm a human operator.

Although binary decision automation offers clear benefits in terms of efficiency and speed, ethical concerns are still substantial. The lack of human oversight in binary automated systems might lead to risks when unforeseen circumstances arise that fall outside the predefined binary parameters or the trained boundaries of a narrow AI algorithm. Additionally, concerns about accountability and transparency in automated systems highlight the importance of robust safeguards and ethical frameworks to mitigate unintended consequences (Horowitz, 2018). (Williams, et al., 2022)

The evolving nature of these systems raises questions about where and how human judgment should intervene in decision-making processes. The debate centres on balancing operational efficiency with accountability, particularly in life-and-death scenarios where binary decisions may oversimplify complex ethical considerations. These decisions with potential lethal or strategic consequences are more likely to be subjected to ethical considerations. Non-lethal decisions might be validated in a synthetic environment with AI, but the impact of decisions and actions must always be carefully weighed.

3.2 Command cycle

A study by the NATO Command and Control Centre of Excellence(NATO C2COE) highlighted that the C2-cycle used in the Command and Control Capstone Concept Version (ACT, 2018) is better suited for defining the different steps in Command and Control (C2) for a Commander compared to the well know OODA (Observe-Orient-Decide-Act)-loop (MAJ Scherrenburg, LTC Clemente Clemente, & Streefkerk, 2019). This C2-cycle offers a more suitable representation of relevant steps.

The C2-cycle outlines a sequential order of steps and suggests uniformity in priority and time commitment for a headquarters. Although the steps appear equal in size, not all parts of the C2-cycle hold the same level of relevance.

Like the OODA-loop, a more rapid C2-cycle can provide an operational advantage. A command-and-control system designed to support operational decision-making, real-time data integration, and collaboration among diverse partners. The future, as envisioned by NATO C2COE, involves distributed and dispersed headquarters using information superiority and data-driven decisions (MAJ Scherrenburg, LTC Clemente Clemente, & Streefkerk, 2019). This requires trust in AI-human teaming.

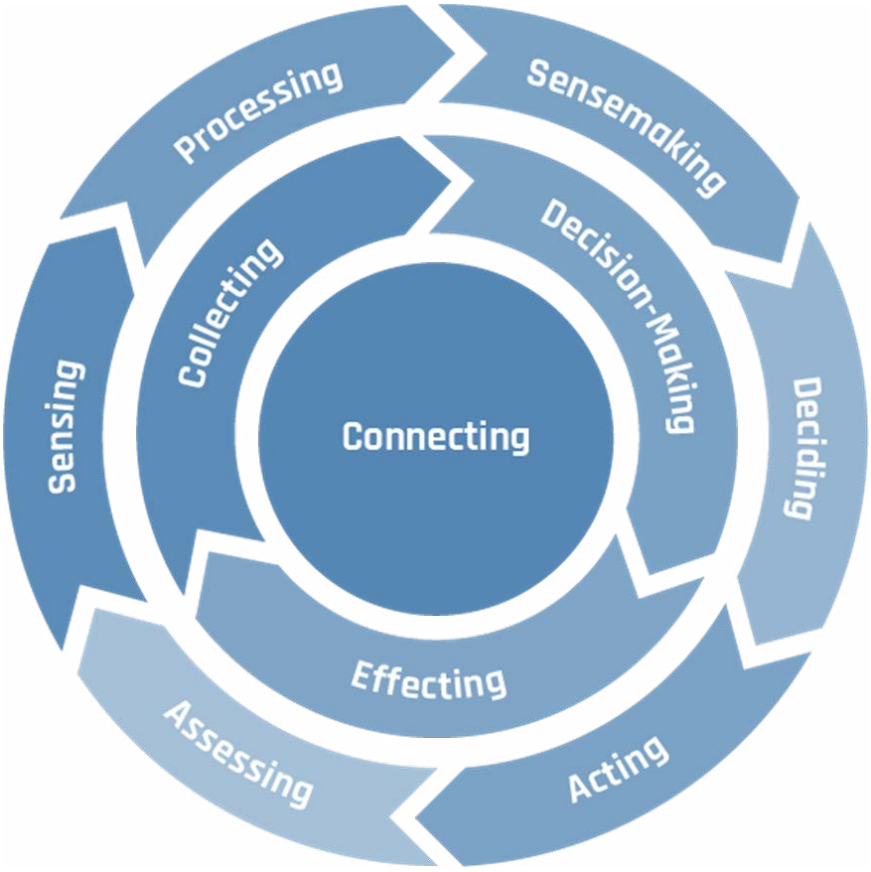


Figure 1: Command and Control – Cycle (ACT, 2018)

The C2-cycle includes the following elements:

1. Connecting: In the Connecting phase, the focus is on setting up a global network that integrates a vast array of sensors and people to support the three phases and six sub-phases of the C2 cycle. This phase is about creating a robust communication and information infrastructure that links various actors, platforms, and systems worldwide. It includes manual, semi-automated, and automated communication and information capabilities needed to connect actors and platforms in a network-enabled environment, providing an agile, secure, and resilient plug-and-play infrastructure for end-to-end connectivity. The goal is to ensure seamless collaboration and data flow across all domains. This integrated network enables real-time connectivity and interoperability, allowing for efficient data collection, processing, and decision-making. Visual representations might include communication pathways, network topologies, or integration systems, showing how different elements connect within this global network. For example, interfaces could depict how data flows between sensors, field units, and command centres, illustrating the connections that support situational awareness and decision-making. The essence of the Connecting phase is to build an interoperable, resilient, and agile infrastructure that supports cross-domain collaboration across the entire C2 cycle, emphasizing the importance of global real-time connectivity, real-time data sharing, and effective, ease-of-use, human-machine teaming. This ensures that all components work together seamlessly to enhance operational effectiveness and decision-making.

Example: AI-powered communication systems can automatically route messages and data to the right recipients, prioritize urgent communication/messages, and translate languages in real-time to ensure seamless collaboration across multinational forces.

2. Collecting: In the Collecting phase, the focus is on managing the collection of data within the operation's area of interest. This involves gathering information from various sources, such as sensors and intelligence feeds, and processing it to support the next steps to develop situational awareness and support decision-making. The collected data, which can be extensive, plays a crucial role in informing a commander's decisions. During this phase, data is gathered from the operational environment, including sensor inputs and raw data, which are then processed to provide a clear picture of the situation. If visual representations show data input, sensor networks, or information processing, they illustrate this part of the cycle. For instance, images with symbols for sensor nodes, satellite imagery, or data processing systems can demonstrate how information is collected and processed in a battlefield or operational context. This helps in understanding the link between data collection and decision-making, especially for real-time intelligence. The Sensing aspect involves using sensor fusion, data aggregation, and platform-agnostic data collection to improve situational awareness and intelligence. The Processing aspect focuses on applying advanced techniques like anomaly detection, pattern recognition, visualization, and cognitive computing to effectively process the relevant data.

   Example: AI algorithms can analyse data from multiple sensors to detect patterns and anomalies, providing commanders with real-time insights and alerts about potential threats or changes in the operational environment.

3. Decision Making: In the Decision-Making phase, the focus is on understanding the situation and making informed choices. This involves analysing incoming information, considering different courses of action, and making decisions based on the data available. The emphasis here is on sensemaking and deciding. A suitable human-machine interface for this phase might include tools like decision trees, analysis models, or visualization aids that help interpret complex data. These interfaces can show how commanders and their staff analyse the situation and make decisions. Decision-support systems, machine-learning analytics, and command workflows are crucial in this phase, as they help prioritize threats, assess risks, and simulate outcomes, thereby speeding up the decision-making process and making it more dynamic. The essence of this phase is to provide relevant data in context to offer alternatives for decision-makers and to base decisions on superior information and the art of command. This ensures that decisions are well-informed and timely, enhancing the overall effectiveness of the C2 cycle.

   Example: AI-driven decision-support systems can simulate various scenarios (Course of Action - COAs) and outcomes, provide 'red-teaming'[1], 'wargame' the given alternatives in a synthetic environment, and help commanders evaluate different strategies and make informed decisions quickly.

4. Effecting: In the Effecting phase, the focus is on implementing decisions to bring about measurable changes in conditions to resolve unwanted situations. This can involve a combination of human, semi-automated, and automated activities across political, military, economic, social, and informational domains.

1. Red teaming in the context of military Course of Action (COA) analysis involves using a dedicated team to critically challenge and evaluate plans, strategies, and assumptions from the perspective of an adversary. This process helps finding potential weaknesses, vulnerabilities, and unforeseen consequences in operational plans.

This phase includes two main activities: acting and assessing. Acting involves executing decisions by delivering both lethal and non-lethal effects. This could mean deploying forces on the battlefield and coordinating with non-military instruments of power. Assessing measures the effectiveness of these actions about the intended outcomes and shows any unintended consequences. Depending on the level of warfare[2], AI interfaces for operators or commanders might show campaign oversight, specific force movements, actions, or outcome assessments. Tools like effect-based operations charts, outcome measurement tools, or feedback loops for assessing unintended consequences are often visualized here. For example, a screen displaying lethal and non-lethal effects, their projected impacts, and methods for assessing these impacts would be crucial for commanders to understand the effectiveness of their decisions. The essence of this phase is to ensure that actions are conducted effectively and efficiently, with a clear understanding of their impacts. This involves coordinating between various military and non-military actors to achieve the desired operational effects.

Example: AI can enhance the precision of both lethal and non-lethal effects. For instance, AI-driven targeting systems can improve the accuracy of airstrikes, while AI algorithms can optimize electronic warfare tactics to disrupt enemy communications effectively.

Example: AI can aid in assessing the effectiveness of actions by analysing data from various sources to measure the impact of operations. AI-powered analytics can detect unintended outcomes and suggest changes for current operations.

3.3 Military Commanders

A fundamental aspect of military operations within NATO, is the concept of the single chain of command. This hierarchical structure ensures that military orders flow from the highest-ranking officer down to subordinate personnel in a clear, linear fashion, with each layer of command responsible for executing directives from the level above. At the top of this chain is the commander, who has traditionally made critical decisions in high-stakes environments such as combat operations. This model has long been effective in traditional warfare scenarios, ensuring that all units stay coordinated and focused on the same objectives. While the hierarchical structure provides clarity, accountability, and efficiency in executing orders, it also places significant responsibility on a single individual.

The NATO-tenet of mission command encourages decentralized decision-making, empowering subordinates to act decisively within the intent of the commander. The introduction of AI capabilities can relieve some of this dependency on a single commander by distributing decision-making responsibilities more evenly across the chain of command. This is particularly beneficial in complex, fast-moving environments where traditional hierarchical decision-making can be slow or rigid.

2. NATO recognizes three primary levels of warfare, which help in organizing and synchronizing military operations:

- Strategic Level: This level involves the highest level of decision-making, focusing on national or coalition objectives. It encompasses the overall direction and use of military power to achieve political goals. Strategic decisions are made by national leaders and senior military commanders.
- Operational Level: This level bridges the gap between strategic objectives and tactical actions. It involves planning and conducting campaigns and major operations to achieve strategic goals within a theatre of operations. The operational level focuses on the coordination and movement of forces to achieve specific objectives.
- Tactical Level: This is the level where individual battles and engagements are planned and executed. It involves the direct application of combat power to achieve specific, immediate objectives on the battlefield. Tactical decisions are made by lower-level commanders and involve the detailed management of troops and resources.

Mission command, as defined in NATO's AJP-01 doctrine, emphasizes decentralized execution and the importance of subordinate initiative within the framework of the commander's intent. Considering this, AI's ability to analyse massive datasets and provide actionable intelligence in real time could enable multiple actors within the military hierarchy to contribute to decisions, potentially leading to a more decentralized decision-making process. The integration of AI into military operations can be seen as enhancing mission command by providing lower-level commanders with real-time, data-driven insights that inform their decisions while still aligning with the overarching strategic objectives. AI can support this by offering actionable intelligence and predictive analytics, enabling more distributed decision-making processes that enhance responsiveness and effectiveness on the battlefield.

For instance, when creating a Comprehensive Understanding of the Operations Environment (CUOE)[3], what biases do we accept? Every digital system simplifies reality to some extent. Filtering data excludes information that could be vital at various levels. Who decides what selection is used? Can an AI tool make these decisions, or should we rely on human viewpoints and search queries on big data lakes? The risk of over-confidence in single-commander leadership and complex risk-taking must be mitigated by a diverse and well-informed oversight team.

Operationally, headquarters prioritize planning and equipping commanders with top-notch insights for decisions. CUOE is vital for effective decision-making and planning, helping commanders anticipate challenges, spot opportunities, and coordinate efforts across various domains. This comprehensive understanding supports the

3. The Comprehensive Understanding of the Operations Environment (CUOE), as outlined in NATO's Allied Joint Publication (AJP)-01F, refers to a holistic approach to understanding the operational environment. This involves integrating various factors such as political, military, economic, social, information, and infrastructure aspects to create a detailed and nuanced picture of the environment in which operations are conducted.

execution of joint operations by ensuring that all relevant factors are considered and addressed. Execution is conducted by tactical or component commanders.

In contrast to a single commander making all critical decisions, AI could empower lower-level commanders and even individual soldiers to make informed decisions based on real-time data feeds, battlefield conditions, and predictive models. Moreover, ensuring that AI systems are transparent and explainable can enhance trust and facilitate better human-AI collaboration (David Gunning, 2019). While the commander would still have ultimate authority, AI could facilitate more collaborative decision-making, which may be beneficial in complex, fast-moving environments where traditional hierarchical decision-making can be slow or rigid (Scharre, 2018).

While the benefits of integrating AI into the command line are evident, there are several challenges to consider. Firstly, the current command structure may resist shifting decision-making power from commanders to AI systems or lower-ranking personnel. Commanders typically make decisions based on their experience, judgment, and intuition, and relying heavily on AI recommendations may be viewed as a threat to their authority.

Secondly, some studies indicate that the integration of AI in military operations must consider the psychological impacts on soldiers. AI systems can reduce cognitive load, but reliance on these systems may result in skill degradation among personnel (Cummings, 2017).

Thirdly, trust in AI systems is crucial, as military leaders need to trust the data and insights provided by AI and ensure that their subordinates are confident in using these tools (LTC Gangi, MAJ Bartko, & MAJ van der Veer, 2019) Without trust in the reliability and accuracy of AI, commanders may be hesitant to incorporate AI into their decision-making processes, preferring traditional methods that rely solely on human judgment (Arkin, 2010).

4. Challenges and Considerations

The necessity for lethal AI measures arises from the urgent need to safeguard military assets and respond to significant operational threats. The advancements in AI technology for lethal measures by strategic competitors and terrorist groups underscore its potential, raising important ethical considerations.

In its Artificial Intelligence Strategy (NATO, 2024), NATO underscores six fundamental principles for the responsible use of AI in defence. These principles include lawfulness, which ensures adherence to international law; responsibility and accountability, which assigns definitive responsibility for AI systems; explainability and traceability, which mandates that AI decisions be transparent and traceable; reliability, which guarantees the dependability of AI systems; governability, which maintains control over AI systems; and bias mitigation, which involves preventing and addressing biases within AI systems. These principles are intended to guide NATO's integration of AI technologies while keeping ethical standards and operational effectiveness.

This paper will later discuss the ethical and legal restrictions surrounding these technologies. One of the primary advantages of semi-autonomous systems is their ability to enhance targeting precision, reduce human error, and enable rapid, data-driven decision-making on the battlefield (Morgan, et al., 2020). In addition, autonomous systems equipped with AI can operate in hostile environments where human presence is either impossible or too risky, providing a critical advantage in combat situations (Trusilo, 2023).

On the other side, AI's potential for non-lethal measures is equally significant. AI can be used in crucial, non-lethal roles like surveillance, cyber defence, and intelligence analysis. For example, AI-powered computer vision systems can monitor and spot potential threats in real time, avoiding human confrontations on the battlefield. Moreover, AI-powered NLP can sift through vast amounts of communication data to detect and mitigate cyber threats before they escalate.

At the tactical level, AI's impact is profound. AI systems autonomously gather and integrate data from diverse sensors like cameras, radars, and satellites (Morgan, et al., 2020). Advanced AI algorithms analyse incoming data instantaneously, offering immediate insights and alerts to operators (Vivoli, Bertini, & Capineri, 2024). AI systems continuously learn from historical data, enhancing their performance over time and adapting to new threats or environmental changes (Morgan, et al., 2020). Additionally, AI generates intuitive visualizations and dashboards, presenting complex data in an easily digestible format for decision-makers. Multiple AI systems collaborate, sharing data and insights to create a more cohesive and comprehensive situational understanding (Morgan, et al., 2020).

Despite the tremendous benefits AI offers, its deployment raises significant challenges, particularly in ethics, data security, technological integration, policies, and trust. These challenges must be addressed to ensure responsible AI implementation and prevent unintended consequences. In the rapidly evolving landscape of AI-driven defence systems, the question of human oversight is more pressing than ever. Do we need humans in the loop or in the system? This question is not just about technological capability but also about ethical responsibility and strategic prudence. This section provides an overview rather than a detailed analysis of all challenges. Other publications offer more extensive discussions on these points.

4.1 Ethical Concerns in AI Deployment

One of the most significant challenges in the military use of AI is the ethical implications of autonomous systems, particularly about life-and-death decisions. Autonomous weapons and AI-enabled decision-making systems can act without human intervention, raising serious concerns about accountability, bias, and compliance with international humanitarian law. The use of lethal autonomous weapons systems (LAWS) raises critical ethical concerns, particularly on responsibility and unintended civilian harm.

These issues stem from the delegation of life-and-death decisions to AI systems, which may act unpredictably or without adequate human oversight. The International Committee of the Red Cross (ICRC) has highlighted concerns about the potential loss of human control in such scenarios, emphasizing the risks of unintended consequences, including civilian casualties (ICRC, 2018). Moreover, the principles of international humanitarian law, such as proportionality and distinction, become harder to enforce when autonomous systems are involved, potentially leading to violations of these laws if AI systems malfunction or are not properly monitored. (O'Connell, 2023) This challenge is further exacerbated by the lack of opacity of some AI algorithms, which can make it difficult to understand how decisions are made.

AI systems in defence may inherit biases from their training data, leading to unintended outcomes for certain individuals or groups (Allen & Chan, 2017). To mitigate these risks, AI must be transparent, explainable, and rigorously overseen. Human oversight is crucial, especially in lethal decision-making, to prevent unintended consequences and keep ethical standards.

4.2 Data Privacy and Security

AI's ability to process vast amounts of data in real time presents both opportunities and challenges, particularly in the areas of data privacy and security. AI systems require access to large datasets, which often include sensitive military and intelligence information. Protecting this data from cyberattacks and unauthorized access is critical to ensuring the integrity of AI systems and supporting the security of NATO operations.

Data breaches in AI systems could have catastrophic consequences, as compromised data may allow adversaries to manipulate AI decision-making processes or gain access to classified information. As a result, NATO must implement stringent cybersecurity measures to protect sensitive data and ensure that AI systems are resilient to attacks. Recent developments in cybersecurity technologies, including AI-enhanced defence mechanisms, offer new ways to safeguard AI systems from increasingly sophisticated cyber threats (NSTC, 2020). By investing in secure infrastructure and adopting best practices for data protection, NATO can mitigate the risks associated with AI deployment.

4.3 Data Privacy and Security

Integrating AI into military systems and infrastructure is a complex and resource-intensive process. NATO must invest in new AI technologies and upgrade legacy systems to support AI functionalities. Ensuring interoperability between AI systems and traditional military technologies is essential for seamless operations across different branches of the military and among NATO member states.

Recent efforts to modernize NATO's digital infrastructure, often referred to as the "digital backbone," are crucial for supporting AI integration. This infrastructure enables the collection, processing, and analysis of large-scale data, providing the foundation for AI-driven decision-making. However, NATO must also focus on training its personnel to work with AI systems effectively. Commanders and military staff need to understand how to interpret AI-generated insights, incorporate these insights into decision-making processes, and keep human control over AI-driven operations.

Ongoing research into human-machine teaming in military contexts emphasizes the importance of collaboration between AI systems and human operators. An AI system should adapt dynamically to the decision maker, considering their objectives, preferences, and track record (Bosch & Bronkhorst, 2018). The success of AI in defence relies on ensuring that human operators are in control while using AI's capabilities to augment their decision-making (Kaushal, et al., 2024).

The development of Artificial Superintelligence (ASI) in defence could change strategy and decision-making by enabling autonomous operations and improved situational awareness. However, these advancements introduce risks, particularly on system vulnerabilities in critical applications. ASI and other advanced AI systems

are especially prone to adversarial actions, such as data poisoning, jamming, or hacking, which can result in significant misjudgements or loss of control in crucial operations (Boulanin & Verbruggen, 2017) (Longpre, Storm, & Shah, 2022). Defence experts highlight the necessity for robust validation, safe-fail protocols, and adaptive responses to counter these risks. Pre-deployment measures, continuous testing, and resilience against operational context variability are crucial in safeguarding ASI-driven applications against manipulation and maintaining human oversight (Longpre, Storm, & Shah, 2022) (Michel, 2021).

4.4 Legislation and Policies

Key risks for using AI in a military context include accountability for decisions made by AI systems, ensuring transparency in decision-making processes, and preventing unintended consequences such as discrimination or loss of autonomy. Therefore, for military and defence, mitigating risks associated with automated decision-making systems requires robust policies and legislation. These measures should address potential ethical, legal, and operational challenges.

The governance of AI technologies in military operations requires well-defined legislative frameworks and policies that align with international laws and ethical standards. NATO must develop comprehensive guidelines that govern the deployment of AI in compliance with the law of armed conflict and other international regulations. In recent years, there has been growing pressure to regulate AI in military contexts, with various governments and organizations calling for a global framework that ensures the responsible use of AI technologies in warfare (National Security Commission on Artificial Intelligence, 2021).

NATO has already taken steps to integrate AI into its Command and Control (C2) systems. The C2 framework leverages AI to help in data collection, processing, and decision-making, providing commanders with more prompt and correct information. However, setting up clear rules for the operation of AI within this cycle is essential to ensure that the technology supports military objectives without compromising ethical standards. By developing policies that emphasize human oversight and accountability, NATO can ensure that AI becomes a tool that enhances decision-making rather than replacing human judgment.

4.5 Trust and Human-Machine Collaboration

Trust is one of the most important yet challenging aspects of integrating AI into military operations. Building trust between commanders, staff, and AI systems is essential to ensure the effective use of AI technologies. Current research suggests that trust in AI systems is often hindered by a lack of understanding of how these systems function and concerns about their reliability in high-stakes situations. (Arkin, Ulam, & Wagner, Moral Decision Making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception, 2012)

AI systems can process vast amounts of data quickly and without fatigue, but they are not failsafe. They can be prone to biases embedded in their training data and can make errors that a human might catch. Conversely, humans can make errors due to stress, fatigue, or cognitive biases. The challenge lies in finding the right balance between human oversight and AI autonomy. Human oversight enhances AI accuracy, safety, aligns it with human values, and promotes trust in the technology. This oversight is expected to be essential in guaranteeing that AI technologies are applied responsibly and ethically, resulting in more dependable and trustworthy AI-driven solutions. However, humans may lack the competence or be harmfully incentivized, creating a challenge for effective oversight. Therefore, a balanced approach that combines human expertise with AI capabilities is essential for optimal performance and safety (Laux, 2023).

To build trust, NATO must provide opportunities for military personnel to interact with AI systems in controlled environments, allowing them to gain familiarity and confidence in these technologies. Experimentation and practical application will enable commanders to learn how AI can support decision-making while keeping ultimate control over the process (Pfaff, Lowrance, Washburn, & Carey, 2023). Trust-building is a gradual process, but it is essential for ensuring long-term success of AI deployment in military contexts.

Additionally, research has shown that human trust in AI depends on the system's transparency and the perceived reliability of its outputs (Schmidt, Biessmann, & Teubner, 2020). By designing AI systems that are explainable and ensuring that they perform consistently and reliably, NATO can foster greater trust between humans and machines.

5. At what position do we need the Human in the Loop?

The integration of AI into decision-making processes has become increasingly widespread. Yet, this raises a critical question: at what position do we need the human in the loop? While AI offers significant advantages in efficiency and predictive capabilities, there are some domains where human oversight is still indispensable (Russell & Norvig, 2020). Most publications on the necessity of having humans in the loop discuss the following aspects: Complexity and Predictability, Ethical Considerations, Legal Compliance and Accountability. More detailed information is available in various publications on this topic. The discussions include:

Military decisions often involve processing large amounts of data and identifying patterns, tasks at which AI excels. However, AI's capabilities are best suited for predictable scenarios. In complex, unpredictable situations, human intervention is crucial. Humans bring adaptability, intuition, and quick-thinking that are essential in rapidly changing environments, such as battlefield engagements or natural disaster responses (Goodfellow, Bengio, & Courville, 2016)

Decisions that involve ethical dilemmas and potential harm to civilians require nuanced ethical considerations that AI lacks the empathy to address. Human decision-makers can weigh the humanitarian cost and moral implications, making decisions rooted in empathy and ethics (Tegmark, 2017)

Ensuring adherence to international laws and rules of engagement is a foundational element of ethical military operations. Human oversight is essential in setting and upholding engagement rules in conflict zones to align with international humanitarian law. Unlike AI, which may follow programmed rules but lacks the contextual understanding of nuanced laws, humans can interpret and apply legal standards to ensure compliance with ethical norms (Asaro, 2016).

In military operations, accountability is crucial to keep ethical and operational integrity. Human commanders are responsible for the consequences of their actions, ensuring that there is a clear chain of accountability. While AI can execute actions, only humans can be held accountable for the outcomes, especially in cases of operational failures or collateral damage. This accountability fosters transparency and responsibility within military leadership (Arkin, 2009).

5.1 Human control - supervision

In modern AI applications there are various levels of human interaction with AI systems. These can be categorized into three types: human-in-the-loop, human-on-the-loop, and human-out-of-the-loop. Each system has its own unique characteristics and implications for operational effectiveness and ethical considerations.

In human-in-the-loop systems, AI requires human authorization for critical decisions, such as target identification and engagement, which is essential in combat scenarios. This model upholds ethical and legal standards by ensuring that only human operators authorize actions with significant ramifications. An example includes AI-driven missile defence systems where human-in-the-loop processes ensure that operators make final launch decisions (Defense Innovation Board, 2019)

In human-on-the-loop applications, AI systems perform tasks autonomously but allow human operators to intervene if needed. This approach is common in autonomous military vehicles and weaponized drones, where AI-driven systems work independently but remain under human supervision, ready to override if mission parameters change unexpectedly (Schmitt & Thurnher, 2013). According to recent research, this hybrid model enhances the responsiveness and adaptability of AI systems while maintaining a necessary layer of human oversight (Kott, et al., 2018)

Human-out-of-the-loop systems are the highest level of autonomy, where AI operates independently without human intervention. These

systems can make real-time decisions based on their programming and learning algorithms. While practical applications of fully autonomous AI in military contexts are rare due to the risks of unintended consequences, there are emerging uses in non-lethal areas, such as autonomous surveillance drones that patrol border areas without direct human control (Scharre, 2018). The potential for such systems to revolutionize operations is substantial, but ethical and safety concerns stay paramount (Russell & Norvig, 2020).

Balancing these levels of autonomy is essential in defence applications. Over-reliance on autonomous systems can lead to unpredictable outcomes, while insufficient autonomy may hinder the operational efficiency of AI. Striking this balance is key to using AI's capabilities while ensuring ethical responsibility and strategic control.

At the lower boundary, full autonomy might allow an AI system to operate independently, making real-time decisions based on its programming and learning algorithms. This level of autonomy is limited in military applications due to the risks of unintended consequences in combat situations. However, it can be suitable in non-lethal areas, like surveillance drones that autonomously check border areas (Scharre, 2018).

At the upper boundary, supervisory control allows military operators to set specific parameters and operational rules for AI. Here, AI can assist in tasks such as intelligence analysis, threat detection, and logistical planning. Human oversight remains central, allowing intervention if necessary to prevent potential ethical or strategic risks. For example, in unmanned aerial vehicles (UAVs) deployed for reconnaissance, supervisory control enables humans to monitor AI-driven pattern recognition while retaining the ability to override or adjust AI decisions (Scharre, 2018)

The balance between these boundaries is essential in defence applications. Too much autonomy risks unpredictable and potentially dangerous outcomes, while insufficient autonomy may undercut AI's operational efficiency, slowing down decision-making in critical situations. Striking this balance is key to using AI's capabilities while keeping ethical responsibility and strategic control.

By using AI for specific tasks while maintaining human oversight for critical decisions, military operations can achieve a balance between efficiency and ethical responsibility. AI can perform tasks such as analysing vast amounts of data, finding patterns, and providing actionable insights, but human judgment remains indispensable for ethical, legal, and adaptive decision-making (Brynjolfsson & McAfee, 2014)

Thus, figuring out the best collaboration between human oversight and AI autonomy is essential for the successful integration of AI in military contexts. Incorporating human judgment in AI-driven processes ensures that military operations are conducted responsibly, ethically, and in compliance with international laws and ethical standards. This balanced approach encourages trust in AI systems and enhances their effectiveness in complex, high-stakes environments. (Laux, 2023) (Schmidt, Biessmann, & Teubner, 2020)

5.2 Autonomous AI

In the military, AI excels at handling repetitive, data-intensive tasks that benefit from its speed and processing power. Automated, rule-based expert systems in military operations can perform critical tasks without direct human interference, using well-defined algorithms and rulesets that allow them to operate autonomously. AI is particularly useful for processing large datasets to identify patterns, predict outcomes, and suggest actionable insights ((Russell & Norvig, 2020). For example, satellite imagery analysis for tracking enemy movements is enhanced by AI's ability to quickly detect subtle changes over time (Goodfellow, Bengio, & Courville, 2016)

AI systems can also find potential supply chain disruptions, allowing proactive logistical adjustments to avoid shortages in critical resources (Brynjolfsson & McAfee, 2014). In operational planning, these systems generate

optimized plans by considering multiple variables like terrain, weather conditions, and enemy capabilities. For instance, algorithms can assist in devising efficient routes for troop movements or supply chains, enhancing logistical coordination. This enables more adaptable and responsive operational plans that cater to dynamic field conditions, creating logistical plans for troop movements or supply routes (Arkin, 2009)

AI-powered surveillance and reconnaissance systems are essential in collecting real-time data where human lives are at risk. Using AI-powered drones, military forces can conduct surveillance missions and gather intelligence on enemy positions without risking human lives. These drones monitor vast areas continuously, flagging suspicious activities and providing actionable data in real-time ((Asaro, 2016)

Automated defence systems, such as missile defence systems, respond to threats with minimal delay. AI algorithms in missile defence, for example, can track incoming threats and launch interceptors autonomously, reducing reaction times that are crucial in averting potential damage. AI's responsiveness in these systems is often faster and more accurate than manual responses. Additionally, in the realm of cybersecurity, AI plays a crucial role in protecting sensitive military networks and data from cyber threats. Advanced AI algorithms continuously monitor networks for anomalies, quickly detect potential breaches, and can even initiate responses to mitigate attacks (Russell & Norvig, 2020)

Although these systems operate autonomously, their underlying algorithms or rulesets can be explained on demand to ensure transparency and accountability.

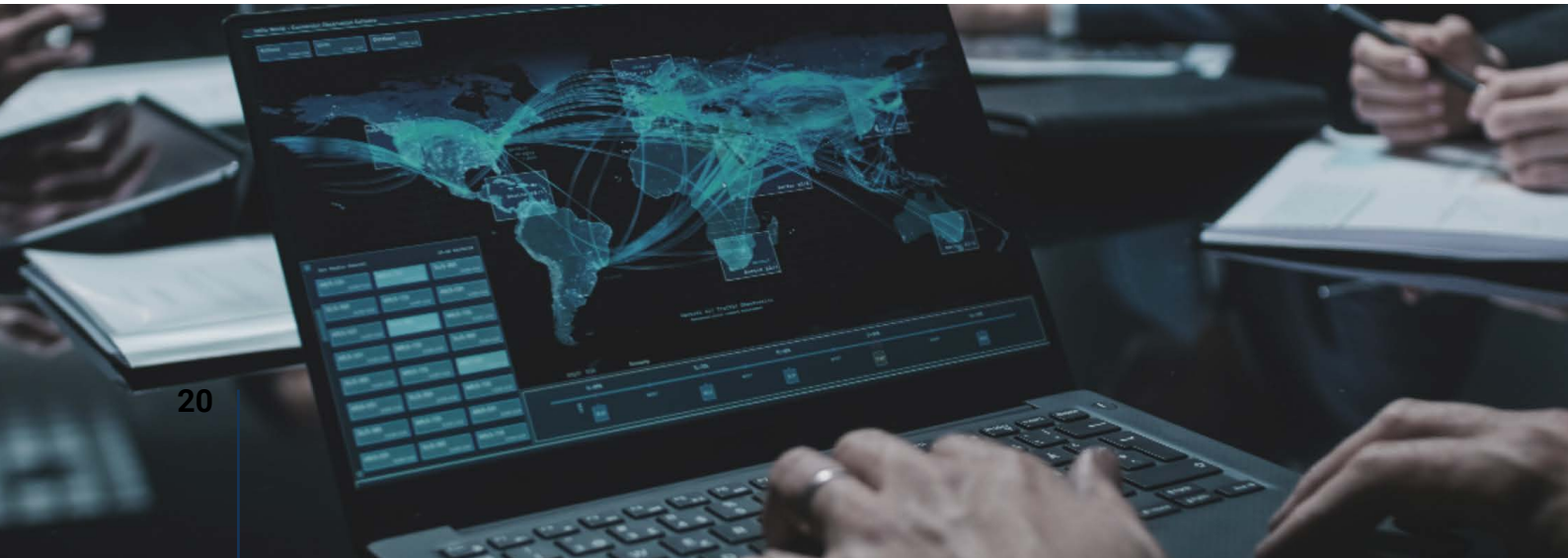
5.3 Human intervention

In high-stakes and rapidly changing situations, such as battlefield engagements or natural disaster responses, real-time human decision-making is essential. Humans bring adaptability, intuition, and quick-thinking that are difficult to replicate in AI, particularly in scenarios that may fall outside of AI's programmed responses. Military commanders, for example, rely on experience and intuition to make real-time adjustments in critical situations, ensuring that actions are adapted to on-the-ground realities (Asaro, 2016).

Humans are known to have reduced decision-making abilities under stress, which can lead to catastrophic errors in high-stakes defence scenarios. This does not mean we should entirely rely on AI. Instead, we should focus on training individuals who can keep composure under pressure. Such individuals need exceptional stress resilience, quick thinking, and a deep understanding of both AI systems and military strategy.

While AI brings efficiency to various operational aspects, there are areas in military decision-making where human judgment remains indispensable. AI's performance is limited to scenarios within its programmed training data, whereas human decision-makers can draw on past experiences, intuition, and real-time environmental cues, allowing them to adapt to novel situations with flexibility and insight (Russell & Norvig, 2020)

Strategic military planning involves formulating long-term goals and adapting to evolving geopolitical landscapes, a task that demands human intuition, experience, and vision. For example, national leaders set military policy and strategic objectives that reflect complex socio-



political goals. Unlike AI, humans can anticipate and navigate the nuanced and unpredictable dimensions of international relations and the diverse interests of multiple stakeholders (Brynjolfsson & McAfee, 2014)

Military decisions often carry profound ethical implications, especially when they involve potential harm to civilians or non-combatants. For instance, deciding to launch an offensive in a populated area requires nuanced ethical considerations that AI lacks the empathy to address. Human decision-makers can weigh the humanitarian cost and moral implications, making decisions rooted in empathy and ethics (Arkin, 2009)

Decisions that involve legal adherence, such as respecting international rules of engagement, also require human judgment to ensure alignment with international laws. Unlike AI, humans understand the broader context and implications of legal standards, making them better equipped to ensure compliance (Russell & Norvig, 2020)

In military operations, accountability is crucial to maintain ethical and operational integrity. Human commanders are responsible for the consequences of their actions, ensuring that there is a clear chain of accountability. While AI can execute actions, only humans can be held accountable for the outcomes, especially in cases of operational failures or collateral damage. This accountability fosters transparency and responsibility within military leadership (Schmidt, Biessmann, & Teubner, 2020).

6. In closing

The future of AI-driven defence lies not in choosing between human and AI control but in finding a complementary balance. Human oversight is still crucial, not to replace AI's advantages but to mitigate its limitations, especially in ethical decision-making and complex strategic judgment. By reserving AI for data-intensive, predictable tasks and retaining human control over strategic, ethical, and crisis-based decisions, military operations can achieve a balance between operational efficiency and ethical responsibility.

In this balanced approach, AI can serve as a powerful tool that augments human capabilities rather than replacing them, ensuring a future where military operations are both effective and ethically grounded. AI does not replace human thinking but augments it. In the art of war, AI will relieve the burden of repetitive tasks within Command and Control functions, but the creativity of the Commander and staff will still be needed.

Non-lethal AI applications can greatly enhance NATO operations by improving situational awareness, decision-making, logistics, and cybersecurity. By using AI's capabilities, NATO can achieve greater operational efficiency, adaptability, and readiness in an increasingly complex and dynamic security environment.

The integration of Artificial Intelligence, Machine Learning, and automation into the C2-cycle has the potential to significantly enhance each phase. AI and ML can improve data collection and processing, providing more accurate and timely situational awareness. Automation can streamline connectivity, communication, and information sharing, making the connecting phase more efficient. In decision-making, AI and ML can help in analysing complex data, finding patterns, and predicting outcomes, thereby supporting commanders in making more informed decisions. During the effecting phase, automation can enhance the precision and effectiveness of actions, as well as improve the assessment of outcomes.

By reserving AI for data-intensive, predictable tasks and retaining human control over strategic, ethical, and crisis-based decisions, military operations can achieve a balance between operational efficiency and ethical responsibility. This balance between leveraging AI's potential and upholding humanitarian values is delicate and requires continuous evaluation. It should also be understood that our adversaries may not share the same values and may make decisions that prioritize operational advantage over moral authority.

Handing over decision-making to AI is practical, but in high stake situations, a person must be accountable for acts of war. Final decisions on whether adversary forces will endure lethal effects should not be handed over to machines, in line with universal values. A human must be accountable for those decisions. The threshold for handing over decisions to machines depends on the impact of the decision at a given moment.

References

Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach. Pearson Education Limited.

Defense Science Board. (2016). Summer study on autonomy. Washington: Office of the Under Secretary of Defense for Acquisition, Technology and Logistics.

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. . Oxford University Press.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. 010 IEEE Symposium on Security and Privacy (pp. 305-316). IEEE.

Szeliski, R. (2011). Computer vision: Algorithms and applications. London: Springer.

Scharre, P. (2018). Army of none: Autonomous weapons and the future of war . New York | London: W. W. Norton & company.

Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K., & Grossman, D. (2020). Military Applications of Artificial Intelligence; Ethical Concerns in an Uncertain World. Santa Monica, Calif.: RAND Corporation.

Sawant, S., Brady, C., Mallick, R., McNeese, N., Chalil Madathil, K., & Bertrand, J. (2023). Human-AI teams in complex military operations: Soldiers' perception of intelligent AI agents as teammates in human-AI teams. Human Factors and Ergonomics Society Annual Meeting, 67, pp. 1122-1124.

Stahl, B. C. (2021). Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies. Cham: Springer.

Boulanin, V., & Verbruggen, M. (2017). Mapping the development of autonomy in weapon systems. SIPRI.

ACT. (2018). Command and Control (C2) Capstone Concept Version 0.5 (draft). Norfolk: Allied Command Transformation. Retrieved April 2018

Vivoli, E., Bertini, M., & Capineri, L. (2024). Deep Learning-Based Real-Time Detection of Surface Landmines Using Optical Imaging. Remote Sens., 16(677).

O'Connell, M. E. (2023). Banning Autonomous Weapons: A Legal and Ethical Mandate. Ethics & International Affairs, , 37, pp. 287–298.

Allen, G. C., & Chan, T. (2017). Artificial Intelligence and National Security. Harvard, Belfer Center for Science and International Affairs.

National Security Commission on Artificial Intelligence. (2021). Final report: National Security Commission on Artificial Intelligence. United States: National Security Commission on Artificial Intelligence (U.S.).

NSTC. (2020). AI and Cybersecurity: Opportunities and Challenges. U.S. National Science and Technology Council (NSTC), Networking and Information Technology Research and Development Program, United States of America.

Kaushal, S., Lynch, J., Suess, J., Lee, J.-J., Vannurden, L., & Bajraktari, Y. (2024). Leveraging Human– Machine Teaming. London: Royal United Services Institute for Defence and Security Studies.

Pfaff, C. A., Lowrance, C. J., Washburn, B. M., & Carey, B. A. (2023). Trusting AI: Integr rusting AI: Integrating Ar ating Artificial Intelligence int tificial Intelligence into the Army o the Army's Professional Expert Knowledge. Carlisle Barracks, PA:: US Army War College.

Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. Journal of Decision Systems,.

MAJ Scherrenburg, M., LTC Clemente Clemente, F., & Streefkerk, J. (2019). The Future of the Command Post - part 1. Utrecht: NATO Command & Control Centre of Excellence.

LTC Gangi, M., MAJ Bartko, F., & MAJ van der Veer, J. (2019). The Future of the Command Post - part 2. Utrecht: NATO Command & Control Centre of Excellence.

Horowitz, M. C. (2018, may). Artificial Intelligence, International Competition, and the Balance of Power. Texas National Security Review, 1(3), 37-57.

ICRC. (2018). Ethics and autonomous weapon systems: An ethical basis for human control? Geneva: International Committee of the Red Cross.

Bosch, K. v., & Bronkhorst, A. (2018). Human-AI Cooperation to Benefit Military Decision Making. NATO IST-160 Specialist' meeting on Big Data and Artificial Intelligence for Military Decision Making. Bordeaux: TNO.

Laux, J. (2023). Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act. AI & Society: Knowledge, Culture and Communication,.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. Retrieved October 2024, from DeepLearningBook.org: https://www.deeplearningbook.org

LeCun, Y., Bengio, Y., & Hinton, G. (2015, may 27). Deep learning. Nature, 521, 436-444.

Longpre, S., Storm, M., & Shah, R. (2022, August). Lethal autonomous weapons systems & artificial intelligence: Trends, challenges, and policies. MIT Science Policy Review, 3, 47-56.

Michel, A. H. (2021). Known Unknowns: Data Issues and Military Autonomous Systems. UNIDIR. Geneva: United Nations Institute for Disarmament Research .

NATO Science & Technology Organization. (2023). A novel predictive maintenance methodology for improving defense logistics processes. NATO STO.

COL Lacroix, E. B. (2023). Future of Army Logistics | Exploiting AI, Overcoming Challenges, and Charting the Course Ahead. Army Sustainment(SUMMER).

MAJ Tilley, S. (2024, October 17). Smart Logistics: Navigating the AI Frontier in Sustainment Operations. Retrieved October 2024, from U.S. Army: https://www.army.mil/article/280377/smart_logistics_navigating_the_ai_frontier_in_sustainment_operations

Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). AI Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions. SN Computer Science, 2(173).

IBM. (n.d.). Artificial intelligence (AI) cybersecurity. Retrieved October 2024, from IBM Security: https://www.ibm.com/ai-cybersecurity?utm_=SRCWW&p1=Search&p4=43700077627821636&p5=e&p9=58700008513382304&gad_=1&gbraid=0AAAAAoS6865NPMQ6FpkMiy6ZWZ3Z-_cP5&gclid=Cj0KCQiA_qG5BhDTARIsAA0UHSJc-sEeg4Kv73tfF07MLBozWO33gYzyLh0D-qSc4TKhQsmeJDFMZ-YaAp

Williams, R., Cloete, R., Cobbe, J., Cottrill, C., Edwards, P., Markovic, M., . . . Pang, W. (2022). From transparency to accountability of intelligent systems: Moving beyond aspirations. Data & Policy, 4(7).

Schmitt, M. N., & Thurnher, J. S. (2013). Out of the loop": autonomous weapon systems and the law of armed conflict. Harvard National Security Journal, 4(2), 231-281.

Palantir Technologies. (2024, April 03). Palantir and Ministry of Economy of Ukraine Sign Demining Partnership. Retrieved October 2024, from Palantir Technologies: https://investors.palantir.com/news-details/2024/Palantir-and-Ministry-of-Economy-of-Ukraine-Sign-Demining-Partnership/

Defense Innovation Board. (2019 , October). U.S. Department of Defence. Retrieved from ations on the Ethical Use of Artificial Intelligence by the Department of Defense: https://media.defense.gov/2019/oct/31/2002204458/-1/-1/0/dib_ai_principles_primary_document.pdf

Cummings, M. L. (2017). Artificial Intelligence and the Future of Warfare. Chatham House.

David Gunning, D. W. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. AI MAGAZINE , 40(2), 44-58.

Trusilo, D. (2023). Autonomous AI Systems in Con ict: Emergent Behavior and Its Impact on Predictability and Reliability. Journal Of Military Ethic, 2-17.

Cole, A., Howard, D., Latiff, R., Lucas, G., & Roy, G. M. (2024). Artificial Intelligence in Military Planning and Operations. Oslo: Peace Research Institute Oslo (PRIO).

Asaro, P. M. (2016). The Liability Problem for Autonomous Artificial Agents. AAAI Spring symposium series. Association for the Advancement of Artificial Intelligence (AAAI).

Tegmark, M. (2017). Life 3.0: Being Human in the Age of Artificial Intelligence. New York: Borzoi Books.

Arkin, R. C. (2009). Governing lethal behavior in autonomous robots. CRC Press.

Arkin, R. C. (2010). The Case for Ethical Autonomy in Unmanned Systems. Journal of Military Ethics.

Arkin, R. C., Ulam, P., & Wagner, A. R. (2012). Moral Decision Making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception. IEEE.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. . W. W. Norton & Company.

Kott, A., Théron, P., Drašar, M., Dushku, E., LeBlanc, B., Losiewicz, P., Rzadca, K. (2018). Autonomous Intelligent Cyber-defense Agent (AICA) Reference Architecture.

NATO. (2024, Jul 10). North Atlantic Treaty Organization. Retrieved from Summary of NATO's revised Artificial Intelligence (AI) strategy: https://www.nato.int/cps/en/natohq/official_texts_227237.htm

Author biography

Marcel Scherrenburg MSc MBA BEng

Marcel Scherrenburg brings 25-plus years of experience in multiple engineering and managerial positions within SMEs in the high-tech manufacturing and product development industry. He holds a master's degree in business administration and bachelor's degrees in both industrial automation and industrial product development. His areas of expertise are (military) system innovation, technology- & product development and operations management. As a reservist (NLD-A) and an engineer, Marcel is bridging the gap between the military and the industrial world. He is a former member and now ambassador of the NATO C2COE.


COMMAND AND CONTROL
CENTRE OF EXCELLENCE