

# **DISEC COMMITTEE**



**REGULATING**  
*the use of*

**AI-DRIVEN  
WEAPONRY**

*in*

**MODERN WARFARE**

# **DISEC: Regulating the Use of AI-Driven Weaponry in Modern Warfare**

## **Table of Contents**

### **Key Terms**

- 1. Introduction to the Committee: DISEC**
  - 1.1 Mandate of the First Committee (Disarmament and International Security)**
  - 1.2 Role of DISEC in Arms Regulation**
  - 1.3 Previous UN Efforts on Emerging Military Technologies**
- 2. Background Information**
  - 2.1 What Are AI-Driven Weapons?**
  - 2.2 Lethal Autonomous Weapons Systems (LAWS)**
  - 2.3 Difference Between Automated and Autonomous Systems**
  - 2.4 Military Applications of Artificial Intelligence**
    - 2.4.1 Target Identification and Surveillance**
    - 2.4.2 Autonomous Drones and Robotics**
    - 2.4.3 Cyber and AI-Enhanced Warfare**
- 3. Legal and Ethical Framework**
  - 3.1 International Humanitarian Law (IHL) Principles**
    - 3.1.1 Distinction**
    - 3.1.2 Proportionality**
    - 3.1.3 Military Necessity**
  - 3.2 Accountability and Responsibility Gaps**
  - 3.3 Human Control vs Fully Autonomous Decision-Making**
- 4. Current Global Situation**
  - 4.1 States Developing AI Military Capabilities**
  - 4.2 Technological Arms Race and Great Power Competition**
  - 4.3 UN Discussions on Autonomous Weapons (CCW Process)**
- 5. Security Implications**
  - 5.1 Risks of Escalation and Rapid Conflict**
  - 5.2 Lowering the Threshold for War**
  - 5.3 Cyber Vulnerabilities and AI Manipulation**
  - 5.4 Proliferation to Non-State Actors**
- 6. Challenges in Regulation**
  - 6.1 Defining “Autonomous Weapon”**
  - 6.2 Verification and Enforcement Mechanisms**
  - 6.3 Dual-Use Technology Issues**
  - 6.4 Balancing Innovation and Security**
- 7. Questions to Ponder**
- 8. Bibliography**

## Key Terms

- **Confidence-Building Measures (CBMs)** : Actions like transparency, communication, and military notifications that reduce mistrust between states and prevent accidental conflict.
- **Autonomous Weapons Systems (AWS)** : Weapons that can select and engage targets with little or no human involvement after activation.
- **Lethal Autonomous Weapons Systems (LAWS)** : A subset of AWS designed specifically to use lethal force without human control over targeting decisions.
- **Military-Civil Fusion (China)** : A strategy where civilian technology development is directly integrated into military modernisation.
- **Defence AI Strategy (UK)** : A policy focusing on the responsible use of AI in military systems while maintaining human oversight.
- **ISR (Intelligence, Surveillance, Reconnaissance)** : Military systems that gather and analyse information about enemy activity using sensors, satellites, and AI.
- **Group of Governmental Experts (GGE)** : A UN-mandated group composed of selected state experts who study specific security issues (such as cyber warfare or LAWS).
- **Open-Ended Working Group (OEWG)** : A UN General Assembly-based group where **all member states can participate** (unlike GGE, which is limited in membership).
- **Convention on Certain Conventional Weapons (CCW)** : A major international treaty framework that restricts or bans weapons deemed excessively harmful or indiscriminate.
- **UNIDIR (United Nations Institute for Disarmament Research)** : A UN research institute that provides **independent analysis on disarmament issues**.
- **Article 36 Weapons Review** : A legal obligation under Additional Protocol I of the Geneva Conventions.

## 1. Introduction to the Committee: DISEC

### 1.1 Mandate of the First Committee (Disarmament and International Security)

The First Committee of the United Nations General Assembly, formally known as the **Disarmament and International Security Committee (DISEC)**, is responsible for addressing global security, disarmament, and threats to international peace. Established under the authority of the **United Nations**, the First Committee operates in accordance with the principles outlined in the **United Nations Charter**, particularly those concerning the maintenance of international peace and security.

DISEC's primary mandate is to consider matters related to arms control, non-proliferation, and the regulation of conventional and unconventional weapons. This includes nuclear disarmament, preventing an arms race, regulating emerging military technologies, and reducing global military tensions. The committee also addresses regional security concerns, confidence-building measures, and transparency in military capabilities.

Although DISEC lacks binding enforcement authority, it plays a critical normative role in shaping international standards and influencing global policy through resolutions, recommendations, and multilateral dialogue. Many major arms control frameworks and global disarmament initiatives have been debated within the First Committee before being formalised in international agreements.

In the context of regulating AI-driven weaponry, DISEC serves as a key platform for discussing the security, legal, and ethical implications of emerging technologies in warfare, facilitating international cooperation and promoting responsible state behaviour in the development and deployment of advanced military systems.

## **1.2 Role of DISEC in Arms Regulation**

The Disarmament and International Security Committee (DISEC), as the First Committee of the United Nations, plays a central role in shaping international discussions on disarmament and arms regulation. While it lacks enforcement authority, DISEC serves as the primary forum within the UN General Assembly for debating global arms control measures, promoting transparency, and advancing norms to reduce military risks.

DISEC examines a broad spectrum of weapons-related issues, including nuclear, chemical, biological, and conventional arms, as well as emerging technologies such as autonomous weapon systems and artificial intelligence in warfare. Through resolutions and recommendations, the committee contributes to the development of international legal frameworks and encourages states to adopt confidence-building measures, non-proliferation commitments, and responsible arms transfer policies.

In addition, DISEC facilitates multilateral dialogue among member states, enabling negotiations that may later evolve into binding international agreements. It supports global efforts connected to treaties such as the Treaty on the Non-Proliferation of Nuclear Weapons and discussions under the Conference on Disarmament. By fostering cooperation and encouraging consensus, DISEC helps shape the international regulatory environment governing weapons development and deployment.

In the context of AI-driven weaponry, DISEC's role becomes increasingly significant, as it provides a structured platform for addressing the legal, ethical, and security implications of emerging military technologies and for exploring potential global regulatory mechanisms.

## **1.3 Previous UN Efforts on Emerging Military Technologies**

One of the most significant initiatives concerns Lethal Autonomous Weapons Systems (LAWS). Since 2014, discussions have been conducted under the Convention on Certain Conventional Weapons (CCW), which established a Group of Governmental Experts (GGE) to examine the legal, ethical, and military dimensions of autonomous weapons. These discussions focus on ensuring compliance with International Humanitarian Law (IHL), particularly the principles of distinction, proportionality, and accountability. While no binding treaty on autonomous weapons has yet been adopted, the CCW process has produced guiding principles emphasising the importance of meaningful human control over weapons systems.

The UN has also addressed related issues such as cyber warfare and information security. Through resolutions adopted in the General Assembly, member states have examined norms of responsible state behaviour in cyberspace, including the work of various Groups of Governmental Experts and Open-Ended Working Groups on information and telecommunications in the context of international security.

Additionally, the UN Secretary-General has repeatedly called for stronger regulation of autonomous weapons, warning that machines with the power to take human life without human control are “politically unacceptable and morally repugnant.” These statements have reinforced growing international awareness of the need to establish governance mechanisms for rapidly evolving military technologies.

Overall, while progress toward binding legal regulation remains limited, the UN has played a crucial role in initiating dialogue, shaping global norms, and building consensus around the responsible development and use of emerging military technologies.

## 2. Background Information

### 2.1 What Are AI-Driven Weapons?

AI-driven weapons are military systems that use **artificial intelligence (AI)** to perform tasks that would normally require human judgment — such as identifying targets, analysing threats, navigating environments, or even deciding when to fire. The key idea is that AI allows a weapon system to **process data, learn patterns, and make decisions faster than humans**, sometimes with reduced human involvement

### 2.2 Lethal Autonomous Weapons Systems (LAWS)

Lethal autonomous weapon systems (LAWS) are a special class of weapon systems that use sensor suites and computer algorithms to independently identify targets and employ onboard weapons to engage and destroy them without human control. It is important to note that while LAWS capable of independent targeting are AI-enabled, not all AI-enabled weapons systems qualify as LAWS; many operate with human oversight or limited autonomy and perform complex functions without fully independent targeting. Some operate using pre-programmed, rule-based logic

### 2.3 Difference Between Automated and Autonomous Systems

A clear distinction between “automated” and “autonomous” systems lies in the degree of adaptation, learning capacity, and decision-making authority embedded within the system.

Automated systems operate within predefined, narrowly structured sets of parameters. Their functions are limited to specific tasks, and the decisions or actions they take are based on predetermined rules, algorithms, or heuristics. Such systems are designed to perform repetitive operations efficiently under predictable and controlled conditions. Their behaviour does not change substantially in response to new or unforeseen environmental inputs.

In contrast, autonomous systems can learn, adapt, and evolve in response to dynamic, uncertain environments. Rather than relying solely on fixed instructions, they analyse incoming data, adjust their responses accordingly, and may operate beyond the parameters initially anticipated at deployment. These systems can process and incorporate large, expanding datasets at speeds and levels of complexity that exceed human cognitive limitations. As a result, their decision-making processes are not strictly deterministic but responsive to changing operational contexts.

It is useful to conceptualise automation and autonomy as existing along a continuum rather than as mutually exclusive categories. Systems originally developed as automated platforms may gradually incorporate autonomous features to enhance functionality, extend operational lifespan, and increase adaptability. While automated systems execute clearly defined tasks within known constraints, autonomous systems contribute to identifying appropriate courses of action in complex, evolving, and non-deterministic environments.

## 2.4 Military Applications of Artificial Intelligence

Artificial intelligence (AI) has rapidly become a transformative force in modern military systems, reshaping how armed forces operate across the full spectrum of defence activities. AI is already embedded in a range of military applications — from logistics, intelligence, surveillance, and reconnaissance (ISR) to autonomous or semi-autonomous systems and electronic warfare — enhancing both efficiency and effectiveness on and off the battlefield. AI's integration into military systems began decades ago, with early experimentation in basic reasoning and decision support; for example, the U.S. Defence Advanced Research Projects Agency (DARPA) supported projects such as the Dynamic Analysis and Replanning Tool (DART) in the early 1990s to optimise logistics operations. In recent years, advances in machine learning and data processing have enabled more complex applications, including AI-enabled command and control, target identification, and the coordination of unmanned vehicles across land, sea, and air domains. Militaries around the world are pursuing AI technologies both to improve traditional tasks — such as threat detection and battlefield planning — and to develop new capabilities that may autonomously process information, support human decision-making, or, in some cases, act independently under human supervision. These developments reflect AI's potential to revolutionise military strategy and operations. Still, they also raise concerns about accountability, escalation risks, and the balance between human control and machine autonomy in warfare in the 21st century. AI now supports a wide range of military functions:

- ***Unmanned and Autonomous Vehicles:*** AI enables unmanned aerial vehicles (UAVs), uncrewed ground systems, and autonomous naval vessels to perform surveillance, reconnaissance, and even tactical operations with reduced risk to human operators.
- ***Intelligence, Surveillance, and Reconnaissance (ISR):*** Machine learning improves pattern detection and target identification, rapidly analysing large data streams from satellites, sensors, and signals intelligence.
- ***Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR):*** AI assists commanders by synthesising information, recommending courses of action, and enabling faster decision cycles.
- ***Electronic Warfare and Cyber Operations:*** AI tools help detect and counter adversary electronic attacks or conduct extensive cyber operations with greater speed and adaptability.
- ***Logistics and Force Management:*** AI optimises supply chains, maintenance schedules, and force deployment, enhancing readiness and reducing logistical bottlenecks.

### 2.4.1 Target Identification and Surveillance

Artificial intelligence plays a critical role in modern target identification and surveillance operations by significantly enhancing the speed, accuracy, and scale of data analysis in military environments. Contemporary armed forces rely on vast streams of information collected from satellites, unmanned aerial vehicles (UAVs), ground sensors, radar systems, and signals intelligence platforms. AI-driven systems, particularly those based on machine learning and computer vision, can process this data in real time, identify patterns, detect anomalies, and classify objects with an efficiency that exceeds human analytical capacity.

In target identification, AI assists in recognising potential threats by analysing imagery and sensor inputs to distinguish between military objectives and civilian objects. Advanced algorithms can be trained to detect vehicles, weapons systems, infrastructure, and troop movements, thereby supporting decision-makers in selecting lawful, strategically relevant targets. This capability is especially significant in high-tempo operations, where rapid assessment may be required to respond to emerging threats.

In surveillance operations, AI enhances persistent monitoring by continuously scanning and evaluating large geographic areas. It can flag unusual activities, track movement across time, and correlate multiple data sources to build a comprehensive operational picture. This improves situational awareness and supports intelligence, surveillance, and reconnaissance (ISR) missions.

However, reliance on AI in target identification and surveillance raises important legal and ethical considerations. Systems must comply with the principles of International Humanitarian Law, particularly the principles of distinction and proportionality. Errors in data interpretation, biased training datasets, or adversarial manipulation of AI systems could result in misidentification, increasing the risk of civilian harm or unintended escalation. Consequently, many states emphasise maintaining meaningful human oversight in AI-assisted targeting processes to ensure accountability and adherence to established legal norms.

### **2.4.2 Autonomous Drones and Robotics**

Autonomous Mobile Robots (AMRs) navigate dynamic environments using LiDAR, cameras, and SLAM (Simultaneous Localisation and Mapping). They adapt to changing layouts and perform zone picking and goods-to-person operations. Automated Guided Vehicles (AGVs) follow predefined routes and are well-suited for repetitive, fixed-path material transport.

Collaborative Robots (Cobots) are used near humans and support tasks such as assembly, packing, or workstation help. Autonomous Case-handling Robots (ACRs) manage heavier loads and multiple cases per cycle. Automated Storage and Retrieval Systems (ASRS) increase vertical density and reduce space requirements. Autonomous aerial systems, commonly known as unmanned aerial vehicles (UAVs), can navigate complex environments, avoid obstacles, and identify objects of interest through real-time data processing. Some systems operate under a “human-in-the-loop” model, where human authorisation is required before the use of force. In contrast, others may operate under a “human-on-the-loop” model, allowing the system to act autonomously within predefined parameters while being monitored by an operator. Advances in swarm technology further enable multiple drones to coordinate their movements and tasks through AI-driven algorithms, increasing operational flexibility and resilience.

Ground-based and maritime robotic systems are also increasingly integrated into military operations. Autonomous ground vehicles can be deployed for reconnaissance in high-risk environments, explosive ordnance disposal, or logistical resupply in contested areas. Similarly, autonomous naval systems may conduct mine detection, surveillance patrols, or anti-submarine operations. These robotic platforms reduce personnel risks while extending operational reach.

Despite their strategic advantages, autonomous drones and robotics raise significant legal, ethical, and strategic concerns. Questions persist regarding accountability for decisions made by AI-enabled systems, particularly in lethal contexts. There are also concerns about system reliability in unpredictable environments, vulnerability to cyber interference, and the potential lowering of the threshold for armed conflict. As these technologies continue to develop, debates within international forums increasingly focus on ensuring meaningful human control, compliance with international humanitarian law, and preventing destabilising arms races driven by autonomous capabilities.

### **2.4.3 Cyber and AI-Enhanced Warfare**

Artificial intelligence has significantly expanded the scope and effectiveness of cyber operations, giving rise to what is often described as AI-enhanced warfare in the digital domain. Modern military strategies increasingly integrate AI into cyber defence, cyber offence, and broader information operations. By enabling rapid data analysis, automated threat detection, and adaptive responses, AI enhances the speed, precision, and scale of cyber activities beyond traditional human capabilities.

In defensive applications, AI systems continuously monitor networks, detect anomalies, and identify malicious activity in real time. Machine learning algorithms can recognise patterns associated with cyber intrusions, malware, and phishing campaigns, enabling faster mitigation and containment. These capabilities are particularly important in protecting critical military infrastructure, communications systems, satellite networks, and command-and-control platforms.

On the offensive side, AI can assist in identifying vulnerabilities in adversary systems, automating reconnaissance, and optimising the timing and methods of cyberattacks. AI-driven tools may also generate adaptive malware that can modify its behaviour to evade detection. In addition, AI plays a role in information warfare, including automated influence operations, deepfake technologies, and large-scale disinformation campaigns designed to manipulate public opinion or destabilise political systems. The integration of AI into cyber warfare raises significant security concerns. AI systems can accelerate the pace of cyber conflict, reducing decision-making time and increasing the likelihood of miscalculation or unintended escalation. The difficulty of attributing cyberattacks further complicates deterrence and accountability. Moreover, AI tools are often dual-use technologies, widely available in civilian markets, which increases the risk of proliferation to non-state actors.

As cyber and AI-enhanced warfare continues to evolve, international discussions increasingly focus on establishing norms of responsible state behaviour in cyberspace, improving resilience of critical infrastructure, and preventing destabilising uses of AI that could undermine global security.

### 3. Legal and Ethical Framework

#### 3.1 International Humanitarian Law (IHL) Principles

**An oft-cited foundation of international humanitarian law (IHL)** is the Martens Clause, first introduced in the Preamble to the 1899 Hague Convention II and now widely regarded as part of customary international law. The Clause affirms that even where no specific treaty rule applies, civilians and combatants remain protected by the principles of international law derived from custom, the laws of humanity, and the dictates of public conscience.

Its significance lies in ensuring that gaps in codified law do not result in a legal vacuum. Especially in the context of emerging technologies—such as AI-driven weaponry—the Martens Clause reinforces that the absence of explicit regulation does not remove humanitarian constraints. It reflects IHL's overarching purpose: to establish minimum, non-derogable standards of restraint applicable in all armed conflicts.

#### *Sources of Core IHL Principles*

The fundamental principles of IHL derive from both customary international law and treaty law, particularly the four Geneva Conventions and the two Additional Protocols. Together, they form the backbone of the modern law of armed conflict.

##### 3.1.1 Distinction

The principle of distinction is a core rule of international humanitarian law (IHL) requiring parties to an armed conflict to distinguish at all times between combatants and civilians, and between military objectives and civilian objects. Attacks may only be directed against legitimate military targets; civilians and civilian property must not be intentionally targeted. This rule also forbids indiscriminate means and methods of warfare that cannot differentiate between lawful and unlawful targets.

Under this principle:

- *Combatants are lawful targets because they directly participate in hostilities, while civilians are protected unless and for such time as they take a direct part in hostilities.*
- *In non-international armed conflicts, combatant status may not be explicitly defined, but a meaningful distinction must still be maintained to protect civilians.*
- *The evolution of modern conflict—such as the presence of non-state armed groups and blurred civilian/combatant roles—presents practical challenges but does not negate the obligation to distinguish.*

*This principle is foundational for lawful targeting decisions and is especially relevant when evaluating whether AI-assisted or autonomous systems can comply with IHL obligations.*

### 3.1.2 Proportionality

The principle of proportionality governs lawful attacks in armed conflict and limits incidental harm to civilians and civilian objects. It does not prohibit all civilian harm, but it does place a legal boundary on permissible force:

- *It prohibits attacks expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof that would be excessive in relation to the concrete and direct military advantage anticipated from the attack.*
- *This rule is codified in Article 51(5)(b) of the First Additional Protocol to the Geneva Conventions of 1977 and is also recognised as customary international law.*
- *The principle requires commanders to balance anticipated civilian harm against military gains using the information reasonably available at the time of the attack. Civilian harm that is foreseeable and not excessive relative to the expected military advantage may be lawful.*
- *Proportionality is distinct from distinction. An attack targeting civilians is always unlawful, regardless of proportionality; the proportionality assessment applies only to attacks against legitimate military targets where some incidental civilian harm might occur.*

In essence, proportionality seeks to humanise warfare by ensuring that military operations do not impose harm on civilians that is disproportionate to the legitimate military objectives sought.

### 3.1.3 Military Necessity

Military necessity in international humanitarian law (IHL) is the concept used to justify the use of force in armed conflict, but it is strictly limited. It allows a belligerent to apply only the amount and kind of force necessary to defeat the enemy and achieve a legitimate military objective. Any unnecessary or wanton application of force is prohibited.

While there is no single authoritative definition, IHL practitioners generally agree that military necessity must meet these conditions: it must be urgent, necessary for a known military purpose, and consistent with IHL's other rules. Force that is not justified by military necessity—such as violence unconnected to concrete military objectives, indiscriminate attacks, or measures aimed at spreading terror—is prohibited.

Crucially, military necessity cannot justify violations of fundamental humanitarian protections. It operates alongside core principles such as distinction and proportionality; an act is lawful only if it is consistent with them. Where IHL expressly allows derogations for military necessity, they must be carefully justified and are limited by law.

## 3.2 Accountability and Responsibility Gaps

Historically, ICL (International Criminal Law) assumes that human agents commit crimes with intent, knowledge, and control; this anthropocentric framework strains when machines can select and attack targets independently, disrupting the usual causal chain linking human decisions to battlefield conduct. As autonomous systems grow more complex and unpredictable, it becomes increasingly difficult to attribute criminal liability to commanders or programmers under doctrines like command responsibility and the mental-element requirements of the Rome Statute (e.g., intent and knowledge), creating a potential “accountability gap” where serious violations occur without a clear human actor liable for them.

The piece explores doctrinal adaptations proposed by legal scholars to address this gap. These include models of indirect perpetration (treating the machine as an intermediary through which humans commit crimes), collective liability for those involved in development or deployment, and even concepts of algorithmic mens rea where systematic harmful patterns in code resemble intent. The article also highlights the emerging norm of Meaningful Human Control (MHC) as a legal-ethical safeguard, insisting that humans must tangibly influence an autonomous system's lethal decisions, and the idea of Advanced Control Directives to document

commanders' legal and operational intent before deployment. However, it concludes that such mechanisms are incomplete without broader legal reform: accountability must remain anchored in human agency, ensuring that legal responsibility does not become the “first casualty of autonomy.”

### **3.3 Human Control vs Fully Autonomous Decision-Making**

One of the central debates surrounding AI-driven weapon systems concerns the extent to which humans should retain control over decisions involving the use of force. The distinction between systems that operate under meaningful human oversight and those capable of fully autonomous decision-making has significant legal, ethical, and strategic implications.

Human-controlled or “human-in-the-loop” systems require direct human authorisation before engaging a target. In such systems, artificial intelligence may assist in data processing, threat identification, or recommendation of possible actions, but the final decision to use lethal force remains with a human operator. A related model, often described as “human-on-the-loop,” allows systems to operate autonomously within predefined parameters while remaining under human supervision, with the option to intervene or override if necessary.

Fully autonomous systems, by contrast, are designed to select and engage targets without real-time human input once activated. These systems rely on complex algorithms and sensor data to interpret their environment, identify threats, and determine appropriate actions. While proponents argue that such systems could increase operational speed and reduce risks to military personnel, critics raise concerns regarding reliability, unpredictability, and the moral implications of delegating life-and-death decisions to machines.

From a legal perspective, the debate centres on compliance with International Humanitarian Law (IHL), particularly the principles of distinction, proportionality, and precaution. Questions arise about whether fully autonomous systems can reliably distinguish between combatants and civilians in complex, rapidly changing environments. Additionally, issues of accountability become more complex when harmful outcomes result from machine-driven decisions, potentially diffusing responsibility among programmers, commanders, and states.

As a result, many states and international organisations advocate for the preservation of meaningful human control over the use of force. Ongoing international discussions seek to define acceptable levels of autonomy in weapon systems while balancing military effectiveness with ethical responsibility and legal obligations.

## **4. Current Global Situation**

### **4.1 States Developing AI Military Capabilities**

#### **United States**

The **United States** is widely regarded as a global leader in the development and integration of artificial intelligence (AI) for military purposes. AI is a central component of U.S. defence modernisation efforts, particularly amid broader strategic competition with technologically advanced rivals. The U.S. Department of Defence (DoD) has identified AI as a critical capability for maintaining operational superiority across the land, air, sea, space, and cyber domains.

Institutionally, the U.S. established the Joint Artificial Intelligence Centre (JAIC) in 2018 to accelerate the adoption of AI across military functions. In 2022, the JAIC was integrated into the Chief Digital and Artificial Intelligence Office (CDAO), which now coordinates AI implementation, data governance, and digital transformation efforts within the DoD. These initiatives reflect a structured approach to scaling AI across defence operations.

Operationally, U.S. AI applications span several key areas:

- ***Intelligence, Surveillance, and Reconnaissance (ISR):*** AI assists in analysing large volumes of imagery and sensor data, improving target detection, threat assessment, and battlefield awareness. Programs such as Project Maven use machine learning to process drone footage and identify objects of military interest more efficiently.
- ***Autonomous and Semi-Autonomous Systems:*** The United States invests heavily in unmanned aerial systems, autonomous naval vessels, and robotic ground platforms. These systems are designed to reduce personnel risks and enhance operational reach. Swarming drone technologies are also under development to enable coordinated operations among multiple unmanned units.
- ***Logistics and Predictive Maintenance:*** AI is used to optimise supply chains, forecast equipment failures, and improve readiness by analysing maintenance data across military platforms.
- ***Cyber and Electronic Warfare:*** AI strengthens cyber defence capabilities by detecting anomalies and responding to network intrusions in real time, while also enhancing comprehensive cyber and electronic warfare strategies.

Ethically and legally, the United States has emphasised the importance of responsible AI development. In 2020, the Department of Defence adopted five ethical principles for AI use: responsibility, equity, traceability, reliability, and governability. These principles aim to ensure that AI systems remain under appropriate human oversight and comply with international humanitarian law.

Strategically, U.S. investment in military AI reflects concerns about maintaining technological leadership amid growing competition, particularly from China and Russia. As a result, AI development is closely tied to national security priorities, alliance interoperability, and deterrence strategies. However, debates continue domestically and internationally regarding the risks of escalation, accountability gaps, and the need for international regulatory frameworks governing AI-enabled military systems.

## **China**

**China** has emerged as one of the leading global actors in the development of artificial intelligence (AI) for military applications. AI is central to China's long-term military modernisation strategy and is closely linked to its objective of transforming the People's Liberation Army (PLA) into a "world-class military" by mid-century. Chinese leadership has repeatedly emphasised that future warfare will be increasingly "intelligentized," meaning that AI, big data, and autonomous systems will fundamentally shape the character of conflict.

A defining feature of China's approach is its **Military-Civil Fusion (MCF)** strategy, which seeks to integrate civilian technological innovation with military development. Through this model, advancements made by private technology firms, research institutions, and universities can be rapidly adapted for defence purposes. This structure accelerates AI innovation across machine learning, robotics, computer vision, and data analytics.

Operationally, China is investing in several key AI-enabled military domains:

- ***Autonomous Systems:*** Development of unmanned aerial vehicles (UAVs), autonomous underwater vehicles, and unmanned ground systems. China has demonstrated interest in drone swarming technologies capable of coordinated group operations.
- ***Intelligent Command and Control:*** AI systems are being designed to assist with battlefield data integration, predictive analysis, and decision-support functions for commanders.
- ***ISR and Surveillance:*** AI-enhanced surveillance technologies improve target recognition, maritime monitoring, and border security.
- ***Cyber and Information Warfare:*** China integrates AI into cyber capabilities and information operations, enhancing its ability to conduct network exploitation, electronic warfare, and strategic influence campaigns.

Chinese military writings often describe AI as a tool to offset traditional military disadvantages and to compete effectively with technologically advanced powers. There is also significant focus on leveraging AI for anti-access/area denial (A2/AD) strategies, particularly in maritime regions of strategic importance.

Unlike some Western states, China has been less transparent about specific ethical constraints governing military AI use, though it has publicly supported discussions within international forums on regulating autonomous weapons. At the same time, rapid technological development and strategic competition raise concerns about the acceleration of an AI arms race.

Overall, China's military AI development reflects a comprehensive national strategy that combines technological innovation, state coordination, and long-term strategic planning to reshape the future balance of military power.

## United Kingdom

The **United Kingdom** is one of Europe's leading actors in integrating artificial intelligence into defence systems. As a key NATO member, the UK views AI as essential for maintaining interoperability with allies and preserving technological competitiveness in an increasingly complex security environment.

The UK Ministry of Defence (MoD) has adopted a formal **Defence AI Strategy** that emphasises responsible innovation, operational effectiveness, and adherence to international law. The UK approach prioritises maintaining meaningful human involvement in decisions regarding the use of force while expanding AI's role in support functions.

Operationally, the United Kingdom employs AI across several domains:

- **Intelligence and Surveillance:** AI supports image recognition, threat analysis, and data fusion across ISR platforms.
- **Autonomous Systems:** Investment in unmanned aerial and maritime systems, including AI-enabled naval platforms for surveillance and mine countermeasures.
- **Cyber Defence:** AI enhances threat detection, network protection, and resilience of military communications systems.
- **Logistics and Predictive Maintenance:** Machine learning tools optimise equipment performance and readiness.

The UK places significant emphasis on **ethical governance** and transparency in the deployment of military AI. It has publicly supported international discussions on the regulation of autonomous weapons and promotes compliance with International Humanitarian Law (IHL). Through NATO and multilateral partnerships, the United Kingdom also contributes to collaborative AI research and interoperability initiatives.

Overall, while the UK's investment scale is smaller than those of the United States or China, its structured strategy, regulatory focus, and strong technological base position it as a major European contributor to military AI development.

## Israel

**Israel** is widely recognised for its rapid integration of AI into military systems, driven by its strong defence industry and security environment. AI is used extensively in **intelligence analysis, surveillance, missile defence systems, and autonomous drones**. Israel has also invested in AI-powered border-monitoring and predictive threat-analysis tools. Its defence sector emphasises real-time data processing and precision targeting, making AI a core component of operational effectiveness.

## South Korea

**South Korea** is actively developing AI capabilities to address regional security challenges. AI is incorporated into **unmanned systems, automated border defence technologies, cybersecurity tools, and command-support platforms**. South Korea has also explored AI-based sentry systems and advanced robotics. Given tensions on the Korean Peninsula, technological modernisation — including AI integration — is a significant element of its defence strategy.

## Japan

**Japan** integrates AI primarily into **defensive and support-oriented military applications**, consistent with its security policy framework. AI is used in robotics, maritime surveillance, missile defence analytics, and cyber defence systems. Japan places strong emphasis on responsible innovation and cooperation with allies, particularly in developing AI-enabled technologies for maritime security and regional stability in the Indo-Pacific.

## Turkiye

**Turkiye** has increasingly incorporated AI into its growing domestic defence industry. AI plays a key role in **unmanned aerial vehicles (UAVs), smart munitions, surveillance systems, and electronic warfare technologies**. Turkey's development of AI-enabled drones has drawn international attention, and the country continues to expand investments in autonomous systems and defence electronics as part of its broader military modernisation efforts.

## 4.2 Technological Arms Race and Great Power Competition

Where earlier generations of major powers relied primarily on conventional or nuclear arms buildups, contemporary strategic competition has shifted toward the weaponisation of emerging technologies, including artificial intelligence (AI), robotics, cyber capabilities, directed energy systems, and space-based assets. Although the tools have evolved, the strategic objectives remain familiar: securing competitive advantage, deterring adversaries, and preventing strategic inferiority. However, analysts increasingly question whether this new technological arms race enhances stability or instead generates new risks. AI-enabled systems may compress decision-making timelines and incentivise preemptive action, while cyber advancements could undermine strategic stability by targeting nuclear command-and-control infrastructures. Additionally, the entanglement of advanced technologies with critical early-warning systems increases the possibility of inadvertent escalation.

Despite significant investment—particularly by the United States in pursuit of “overmatch” and technological superiority—historical precedent suggests that any advantage gained through arms racing is likely to be temporary. Technological breakthroughs are rapidly replicated, countered, or neutralised, especially in an era of advanced computing, globalised supply chains, and accelerating innovation cycles. Unlike during the Cold War, today's principal competitors are more economically comparable, limiting the feasibility of outspending rivals to secure lasting dominance.

Furthermore, deterrence depends not only on capability but also on political will and relative risk tolerance; in regional disputes, states with greater resolve may offset technological disadvantages. Finally, high-tech weapons programs frequently encounter engineering complexity, cost overruns, and schedule delays, diverting resources from core readiness functions such as maintenance, training, and logistics. Collectively, these concerns raise substantive doubts about whether sustained competition in emerging military technologies delivers durable strategic benefits or instead fuels instability, economic inefficiency, and heightened escalation risks.

### 4.3 UN Discussions on Autonomous Weapons (CCW Process)

On 2 December 2024, the **United Nations** General Assembly adopted a resolution on Lethal Autonomous Weapons Systems (LAWS) with overwhelming support—166 votes in favour, three against (Belarus, the Democratic People’s Republic of Korea, and the Russian Federation), and fifteen abstentions. The resolution reflects growing international concern regarding the development and deployment of LAWS in contemporary conflicts, including in Ukraine and Gaza, and proposes consideration of a two-tiered regulatory approach: prohibiting certain systems while regulating others under international law.

Parallel discussions have taken place within the **Convention on Certain Conventional Weapons** (CCW), where the Group of Governmental Experts (GGE) on LAWS has debated regulatory frameworks for over a decade. Although the CCW has produced guiding principles and draft regulatory elements, progress has been slowed by its consensus-based decision-making model. Nonetheless, international momentum toward clearer and more robust legal frameworks continues to grow.

LAWS have emerged through advances in artificial intelligence, robotics, sensor technologies, and real-time data processing. While no universally agreed legal definition exists, LAWS are generally understood as weapons systems that, once activated, can select targets and apply force without human intervention. These systems are commonly categorised along a spectrum of human involvement:

- **Human-in-the-loop** (*semi-autonomous*): *requiring human authorisation before applying force.*
- **Human-on-the-loop** (*supervised autonomous*): *capable of selecting and engaging targets without prior approval but subject to human supervision and override*
- **Human-out-of-the-loop** (*fully autonomous*): *operating without human authorisation, supervision, or intervention after activation.*

Examples frequently cited in international discussions include the Phalanx Weapon System (United States), the HARPY loitering munition (Israel), the Lancet-3 drone (Russian Federation), and the KARGU system (Türkiye). These illustrate varying levels of autonomy in targeting and engagement.

There is a broad international consensus that existing international law fully applies to LAWS. This includes the **United Nations Charter**, international humanitarian law (IHL), international criminal law, international human rights law, state responsibility doctrines, and specific weapons treaties. Under IHL, LAWS must comply with the principles of distinction, proportionality, and precautions in attack. States are obligated to conduct weapons reviews under Article 36 of Additional Protocol I to ensure that new weapons are lawful. Accountability remains with human operators, commanders, and states, even where autonomous systems are employed.

However, concerns persist about the predictability and reliability of AI systems, bias in AI systems, and the delegation of life-and-death decisions to machines.

**The International Committee of the Red Cross** and the UN Secretary-General have called for a new international treaty establishing explicit prohibitions and restrictions on LAWS. Proposed measures include ensuring meaningful human control, limiting operational scope, incorporating self-deactivation mechanisms, preventing deployment in densely populated civilian areas, and mitigating algorithmic bias.

Although the CCW process has generated draft regulatory proposals, critics argue that consensus-based negotiations may hinder timely progress. As a result, the 2024 UN resolution initiates informal consultations among member states in 2025, signalling increasing political momentum toward the negotiation of a dedicated treaty by 2026.

In conclusion, while existing legal frameworks provide foundational obligations governing LAWS, gaps in specificity and enforcement underscore the growing demand for a comprehensive international instrument. A dedicated treaty could harmonise interpretations, clarify prohibitions and restrictions, and strengthen accountability mechanisms in response to the rapid advancement of autonomous military technologies.

## **5. Security Implications**

### **5.1 Risks of Escalation and Rapid Conflict**

#### **1. Unpredictability**

These systems rely on complex machine-learning algorithms interacting with dynamic battlefield environments. Their behaviour is difficult to predict in real-world settings, and in some cases, they are intentionally designed to behave unpredictably to outmanoeuvre adversaries. This creates serious concerns about reliability and control.

#### **2. Escalation Risks**

Due to their speed and automation, autonomous weapons may unintentionally escalate conflicts. Research by RAND and UNIDIR indicates that AI-driven systems can contribute to crisis instability and inadvertent escalation. The U.S. National Security Commission on AI also warns that faster, automated warfare reduces opportunities for human-led de-escalation.

#### **3. Proliferation**

Unlike conventional weapons, so-called “Slaughterbots” are cheap to produce, easy to transport, and difficult to detect. Once developed by major powers, they are likely to spread globally, potentially reaching terrorists, dictators, or non-state actors.

#### **4. Lowered Barriers to War**

Because autonomous weapons reduce financial and human costs, they may weaken traditional deterrents to war, thereby increasing the likelihood of armed conflict.

#### **5. Mass Destruction Potential**

Autonomous weapons are highly scalable. A single individual could deploy large swarms capable of causing massive casualties, leading some to classify them as potential weapons of mass destruction.

## 6. Selective Targeting and Bias

Facial recognition and biometric targeting increase risks of discrimination, ethnic cleansing, and biased violence, especially given documented racial and gender inaccuracies in AI systems.

## 7. AI Arms Race

The development of autonomous weapons risks triggering a global AI arms race, prioritising speed and dominance over safety and ethical considerations.

### 5.2 Lowering the Threshold for War

The integration of artificial intelligence into military systems risks significantly lowering the threshold for war. Traditionally, decisions to use force were constrained by high financial costs, logistical complexity, and—most importantly—the risk to soldiers' lives. AI-enabled and autonomous weapons systems reduce these constraints.

First, autonomous systems decrease human exposure on the battlefield. When fewer soldiers are at risk, political leaders may face less domestic opposition to military action. The human cost—historically a powerful deterrent—becomes less immediate. This shift can make the initiation of force appear more politically manageable.

Second, AI systems operate at high speed and scale. They can identify, track, and engage targets faster than human operators, compressing decision-making time.

In crises, this speed may incentivise preemptive or automated responses, leaving little room for diplomacy or de-escalation.

Third, many AI-enabled weapons—such as loitering munitions and autonomous drones—are relatively inexpensive and scalable compared to conventional platforms like fighter jets or armoured divisions. Lower production and deployment costs reduce economic barriers to sustained or repeated military engagement.

Finally, the AI arms race intensifies the problem. If rival states fear falling behind technologically, they may deploy increasingly autonomous systems prematurely, prioritising strategic advantage over safety and stability.

Taken together, reduced human risk, lower financial costs, increased operational speed, and competitive military pressures may normalise the use of force. Rather than acting as a last resort, military action in an AI-driven era risks becoming a faster, easier, and more frequent policy option.

### 5.3 Cyber Vulnerabilities and AI Manipulation

Rapid advances in artificial intelligence (AI) promise to transform military operations by improving speed, efficiency, and decision-making. However, current AI systems remain insufficiently robust and reliable for the high-stakes environment of warfare. Their vulnerabilities—particularly to cyberattacks—pose serious strategic risks.

AI-enabled military platforms introduce new entry points for adversaries. Unlike many traditional systems, they rely heavily on large datasets and complex, manipulable models. Integrity attacks, such as data poisoning, can corrupt training data and cause AI systems to misidentify targets, even mistaking friendly forces for hostile ones. Evasion attacks exploit model weaknesses to deceive detection systems—for example, by altering drone imagery to conceal military assets.

Confidentiality attacks allow hackers to extract sensitive information about AI models or their training data, potentially exposing classified material and enabling further exploitation. Availability attacks, including denial-of-service (DoS) or ransomware, can disrupt logistics, supply chains, or operational coordination at critical moments.

These threats are especially concerning because exploiting AI vulnerabilities often requires fewer resources than developing the systems themselves. Moreover, efforts to improve AI resilience frequently involve trade-offs that may reduce performance or create new weaknesses.

As AI becomes integrated into conventional and even nuclear-related systems, the risks of malfunction, manipulation, and escalation increase. States must therefore adopt a cautious, risk-based approach, carefully assessing vulnerabilities and their potential consequences before expanding AI deployment in military contexts.

#### **5.4 Proliferation to Non-State Actors**

The proliferation of artificial intelligence (AI) poses growing risks to international peace and security, particularly if accessed by malicious non-State actors. Concerns about terrorist acquisition of lethal autonomous weapons have been raised in multilateral forums, including discussions within the UN Group of Governmental Experts and recent UN General Assembly resolutions. However, policy debates often oversimplify AI proliferation as merely the diversion of autonomous weapons, overlooking the broader ecosystem through which AI technologies spread.

This research project seeks to address that gap by examining what AI proliferation truly entails, how it occurs, and how it can be mitigated. The first phase maps pathways through which AI can be repurposed, accessed, misused, or weaponised. Because AI depends on a decentralised value chain—encompassing software, hardware, data, compute infrastructure, and talent—there are multiple entry points for diversion and misuse. Risks vary by application: the proliferation dynamics of autonomous drones differ from those of large language models (LLMs), with distinct implications for governance.

Two case studies will explore (1) the repurposing of commercial unmanned systems for military use, and (2) the malicious use and spread of open-source LLMs.

The second phase evaluates existing non-proliferation and export control frameworks, identifying implementation and enforcement gaps. The project ultimately aims to develop practical policy recommendations for governments, multilateral institutions, and industry to strengthen counter-proliferation efforts and safeguard peace in the age of AI.

## 6. Challenges in Regulation

### 6.1 Defining “Autonomous Weapon”

Defining an “Autonomous Weapon” is one of the most contested issues in contemporary arms control debates. As highlighted by the United Nations Institute for Disarmament Research, definitions are rarely neutral; they often reflect political, strategic, and legal interests. States seeking flexibility in military innovation may favour narrow definitions, while those advocating regulation or prohibition often favour broader ones.

At its core, an autonomous weapon system (AWS) is a weapon that can select and engage targets without further human intervention once activated. This distinguishes it from automated or remotely operated systems, where humans remain directly involved in target identification and engagement decisions.

However, disagreement arises over key elements:

- *Level of Human Control: Should autonomy be defined by the absence of “meaningful human control,” or merely by reduced human input?*
- *Scope of Functions: Does autonomy refer only to target engagement, or also to navigation, surveillance, and decision-support?*
- *Operational Context: Should definitions vary depending on the environment (e.g., defensive systems vs offensive battlefield use)?*

The United Nations Office for Disarmament Affairs has documented significant variation in national definitions, reflecting divergent legal and security priorities.

A workable definition must therefore balance technical precision with legal clarity. It should distinguish autonomous weapons from automated systems, clarify the role of human oversight, and remain adaptable to evolving technologies. Ultimately, defining “autonomous weapon” is not merely semantic—it shapes accountability, compliance with International Humanitarian Law, and the future regulation of AI in warfare.

### 6.2 Verification and Enforcement Mechanisms

Effective regulation of Autonomous Weapon Systems (AWS) depends not only on definitions and norms, but also on credible verification and enforcement mechanisms. Without these, legal commitments risk remaining purely declaratory.

At the international level, discussions within the framework of the United Nations Office for Disarmament Affairs and expert processes supported by the United Nations Institute for Disarmament Research highlight the difficulty of verifying compliance in the AI domain. Unlike nuclear or chemical weapons, AI systems rely on dual-use software, commercial hardware, and widely distributed research communities. This makes traditional arms control verification tools—such as on-site inspections—harder to apply.

Key verification challenges include:

- *Dual-use nature: The same AI models can serve civilian and military purposes.*
- *Intangible components: Algorithms, datasets, and model weights are easily transferable and difficult to monitor.*
- *Rapid iteration: AI systems evolve quickly, complicating long-term compliance assessments.*

### ***Possible mechanisms include:***

- 1. Transparency and Reporting Measures – States could disclose doctrine, testing standards, and Article 36 weapons review procedures under International Humanitarian Law.*
- 2. Confidence-Building Measures (CBMs) – Information exchanges, joint exercises, or peer reviews to reduce mistrust.*
- 3. Export Controls – Regulation of high-performance computing chips, specialised hardware, or military-grade autonomous platforms.*
- 4. Technical Auditing Standards – Independent verification of human oversight mechanisms, fail-safes, and cybersecurity protections.*
- 5. National Implementation Laws – Clear domestic legislation criminalising prohibited development or misuse.*

Enforcement remains particularly complex. Sanctions, trade restrictions, and diplomatic consequences may deter violations, but attribution in cyber and AI-related breaches is often difficult.

Ultimately, effective governance of AWS will likely require a hybrid model that combines international legal frameworks, national oversight, industry responsibility, and technical safeguards. Without robust verification and enforcement, efforts to regulate AI in warfare risk being undermined by mistrust, secrecy, and rapid technological advancement.

## **6.3 Dual-Use Technology Issues**

### **Dual-Use Technology in Military and Civil Contexts**

Dual-use technology refers to innovations that have both civilian and military applications, reflecting the increasingly blurred boundary between commercial research and defence capability development. Historically, programs such as the U.S. **Technology Reinvestment Project (TRP)** sought to bridge this gap by encouraging cooperation between commercial industry and military research to produce technologies that could serve both economic growth and defence needs. TRP aimed to leverage commercial innovation for military use while avoiding the inefficiencies of a separate defence industrial base. However, political and institutional pressures often shaped dual-use programs in favour of military priorities, leading to technological trajectories that served defence objectives more than broader commercial development. Such initiatives illustrate the inherent tension in dual-use strategies. While they can accelerate technological adaptation and enhance military capability, they also risk distorting market forces and prioritising military requirements, potentially limiting the broader benefits of scientific innovation. The Technology Reinvestment Project (TRP) was created to avoid the pitfalls of earlier dual-use programs, such as VHSIC and the Strategic Computing Program, which often forced separate military production lines and diverged commercial and defence technology development. TRP aimed to advance general technology without distorting market-driven commercial trajectories, accelerating military access to commercial innovations rather than boosting commercial competitiveness.

A council of civilian agencies evaluated TRP grants to ensure technical expertise and awareness of commercial trends. Still, the program was managed by ARPA, emphasising military value over broader goals.

From the start, TRP faced pressure from diverse political groups: U.S. businesses threatened by foreign competition, defence researchers worried about reliance on foreign tech, neoliberal Democrats favouring investment over consumption, labour and peace movements seeking domestic jobs, and officials seeking to prove that dual-use R&D could succeed. The Clinton administration positioned TRP as part of a national defence conversion and economic strategy. However, the program's true focus was on long-term technological advancement, not short-term job creation. Public messaging, including President Clinton's "Operation Restore Jobs" statement, created the impression that TRP was a jobs program.

This mismatch between expectations and purpose, combined with TRP's careful design to avoid past failures, made it politically vulnerable. Despite genuine achievements, the program disappointed some constituencies and struggled to gain broader political support.

#### **6.4 Balancing Innovation and Security**

Balancing innovation and security in AI-driven weapons is one of the most complex and urgent challenges facing modern military strategy, technological development, and international security. The rapid evolution of artificial intelligence offers unprecedented opportunities to enhance defence capabilities, ranging from advanced battlefield decision-making and automated targeting systems to predictive analytics, logistical optimisation, and autonomous reconnaissance operations. These technologies have the potential to improve operational efficiency, reduce human casualties, and provide strategic advantages by processing vast amounts of data far faster than human operators can. At the same time, the integration of AI into weapons systems introduces a wide range of risks unique to conventional military technologies. Unlike traditional weapons, AI systems can exhibit unpredictable behaviour in dynamic environments, make errors based on incomplete or biased data, and unintentionally escalate conflicts through automated responses or misinterpretation of enemy actions. Moreover, AI-driven weapons are highly susceptible to cyberattacks, manipulation, or adversarial exploitation, which could compromise both national security and global stability.

Beyond technical and operational risks, AI in weaponry raises profound ethical and legal questions. Delegating life-and-death decisions to algorithms challenges fundamental principles of accountability, proportionality, and human judgment in warfare. Determining responsibility when autonomous systems fail or cause unintended harm is a problem that existing military and legal frameworks are ill-equipped to address. Furthermore, the development and deployment of AI-driven weapons have strategic implications for global power dynamics, potentially sparking an arms race among technologically advanced states, incentivising rapid adoption over careful evaluation, and increasing the likelihood of destabilising military competition. These factors make it clear that balancing innovation and security is not just a matter of technical design but also requires robust governance, ethical oversight, and international coordination.

Achieving this balance demands a multi-layered approach that integrates technological, strategic, and policy considerations. On the technological side, rigorous testing protocols, validation frameworks, and fail-safe mechanisms must be developed to ensure that AI systems operate reliably in unpredictable conditions. On the policy side, clear regulations and oversight structures are needed to guide research, procurement, and operational deployment, while maintaining accountability for human decision-makers. Ethical frameworks must be established to define acceptable limits of autonomy, determine the role of human control in lethal decisions, and ensure compliance with international humanitarian law. Equally important is international cooperation, as unilateral development of AI weapons by individual nations risks undermining collective security and escalating tensions. Transparency, dialogue, and agreements on norms for the design, testing, and deployment of AI-driven systems are critical to prevent misunderstandings, miscalculations, and misuse.

Ultimately, balancing innovation and security in AI-driven weapons is not about slowing technological progress but about directing it responsibly. The goal is to harness AI's potential to strengthen defence capabilities, enhance precision, and protect human lives while minimising risks to both national and global security. Success requires a holistic approach that combines cutting-edge research, ethical responsibility, strategic foresight, and international collaboration. Only by acknowledging the dual imperatives of innovation and safety, and by embedding them into the very design, regulation, and deployment of AI systems, can societies ensure that these powerful technologies contribute to defence and deterrence without compromising human oversight, stability, or the moral and legal frameworks that govern warfare.

## 7. Questions to Ponder

- How can the international community ensure that AI weapons are used responsibly without stifling technological innovation?
- At what point does a weapon become too autonomous for humans to remain ethically accountable?
- Can AI-driven weapons ever fully comply with international humanitarian law?
- Who should bear responsibility when an AI weapon makes a lethal mistake?
- How might AI lower the threshold for entering armed conflicts, and how can this risk be mitigated?
- Could AI in warfare unintentionally escalate conflicts faster than humans can respond?
- How do dual-use technologies complicate efforts to regulate autonomous weapons?
- What measures could prevent AI weapons from being acquired by non-state actors or rogue regimes?
- How can transparency, trust, and verification be maintained in an AI arms race?
- Should there be global limits or bans on fully autonomous weapons, or is human oversight enough?
- How might AI in military systems affect the balance of power between nations?
- Can ethical programming of AI ever account for the unpredictable nature of real-world conflict?
- How should states balance national security interests with international legal and ethical obligations?
- What role should multinational institutions, like the UN, play in regulating AI weapons?
- How can military innovation be encouraged while minimising risks to civilians and global stability?

## 8. Bibliography

<https://futureoflife.org/>

<https://www.ejiltalk.org/>

<https://logisticsviewpoints.com/>  
<https://disarmament.unoda.org>

[https://disarmament.unoda.org/the-convention-on-certain-conventional weapons/](https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/)

<https://www.un.org/en/ga/first>

<https://www.un.org/sg>

<https://www.techrepublic.com/>

<https://www.flysight.it/>

<https://www.cevians.com/>

<https://stanleycenter.org/>

<https://stanleycenter.org/wp-content/uploads/2020/05/MilitaryApplicationsofArtificialIntelligence-US.pdf>

[https://disarmament.unoda.org/the-convention-on-certain-conventional weapons/](https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/)

<https://www.icrc.org/en/what-we-do/weapons/autonomous-weapons>

<https://www.rand.org/topics/artificial-intelligence-and-defense.html>

<https://www.defence-industries.com/>

<https://justiceinconflict.org/2025/08/06/algorithms-automation-and-accountability-imagining-responsibility-for-the-crimes-of-machines/>

<https://guide-humanitarian-law.org/content/article/3/military-necessity/#:~:text=%E2%80%9CThe%20question%20of%20military%20necessity%20of%20force%20is%20therefore%20prohibited.%E2%80%9D>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC12529059/>

<https://europeanleadershipnetwork.org/>

<https://www.asil.org/>

<https://privacyinternational.org/>

<https://issues.org/>

<https://yalelawjournal.org/>