DISEC STUDY GUIDE REGULATING THE DEVELOPMENT AND USE OF LETHAL AUTONOMOUS WEAPONS SYSTEMS

TABLE OF CONTENT

VII. Real-World Case Studies and Combat Deployments

- **A.**The STM Kargu-2 in Libya (2020)
- **B.** The Nagorno-Karabakh Conflict (2020) and Autonomous SEAD
- **C.** The Russo-Ukrainian War (2022-Present)
- **D.** Israel's Operation Guardian of the Walls (2021) and the "First AI War"
- **E.** Defensive Autonomy: The Aegis and Iron Dome Precedents
- **F.** Peace-time Threats

VIII. Proposed Solutions and Frameworks

- **A.** Legally Binding Instrument (LBI)
- **B.** The "Two-Tier" Regulatory Approach
- **C.** Political Declarations and Codes of Conduct
- **D.** Confidence-Building Measures (CMBs) and Transparency
- **E.** The Shared Human-Machine Decision-Making Model
- **F.** The Three-Phase Operational Test and Evaluation (OT&E) Standard

IX. Questions to be Addressed (QARMA)

X. Bibliography

Letter From Under Secretary General

Esteemed delegates,

I was the Under-Secretary-General of the DISEC committee at OAFLMUN 24, and the Under-Secretary-General of NATO at OAFLMUN 25 last year. Now, it is my distinct honor to serve once again as the Under-Secretary-General of the Disarmament and International Security Committee (DISEC) for this year's conference

The security problems are increasing day by day at the international level. As technology advances at an unprecedented pace, the nature of warfare and global defense is fundamentally transforming. This year, we are tackling one of the most complex and critical issues of our time: the proliferation and regulation of Lethal Autonomous Weapons Systems (LAWS).

What distinguishes DISEC from other committees is its profound responsibility to address threats before they irreparably damage global peace. As the delegates participating in the DISEC committee, you are expected to articulate your thoughts on these emerging technologies while securing the core mission of the United Nations: maintaining international peace and security.

The debate over machines making life-and-death decisions without human intervention has become more important and urgent than ever. Hence, the international community must start to take precautions for its own safety. Delegates must discuss these precautions and think of new, comprehensive regulatory frameworks to minimize the risks that autonomous weapons pose to humanity and International Humanitarian Law.

I hope that participating in this DISEC committee will make you understand better the intricate balance between technological advancement and ethical warfare. Furthermore, I expect delegates to assimilate a true idea about the ongoing problems and bring it to life through robust resolutions.

As I conclude my letter, I would like to thank all the members of the organization for their hard

work and you, the delegates, for participating here.

Sincere wishes,

Burak ŞAHAN

Under-Secretary-General of the Disarmament and International Security Committee (DISEC)

Deadline for Position Papers:20th April, 2026
e-mail:sahanburak234@gmail.com

Letter From Academic Assistant

Dear Delegates,

I am Ahmet Şevket Demirci ACAS of the DISEC comitee this year and ex PGA of this conferance. Before continuing with my writing, I would like to congratulate our honorable USG on his exceptionally brilliant use of AI in his letter.

Let's be honest — the world is changing faster than most of us can keep up with. And nowhere is that more obvious than in how modern warfare is evolving. We're no longer just talking about soldiers on a battlefield. We're talking about machines that can identify, target, and eliminate without a single human pulling the trigger.

That's exactly why this year's topic matters so much. In DISEC, we'll be diving into one of the most controversial issues on the global agenda right now: the proliferation and regulation of Lethal Autonomous Weapons Systems (LAWS). It's a topic that sits right at the intersection of technology, ethics, and international law — and there are no easy answers.

As your Academic Assistant, I'm here to make sure you're not navigating this alone. Whether it's breaking down the legal frameworks, understanding where current international law falls short, or helping you build a resolution that actually holds up — I've got you covered.

But here's what I'd ask from you: come in with an open mind. Don't just represent your country's position on paper — actually think about what it means to let a machine decide who lives and who dies. That's the kind of debate that makes MUN worth it.

Can't wait to see what you all bring to the table.

Best, Ahmet

Academic Assistant, DISEC OAFLMUN 2026

I. Introduction To The Committee: Disarmament and International Security Committee (DISEC)

The Disarmament and International Security Committee (DISEC) is the First Committee of the United Nations General Assembly (UNGA). Established in 1945 alongside the creation of the United Nations itself, DISEC was formed to address the most pressing global challenges regarding disarmament, non-proliferation, and the maintenance of international peace and security.

DISEC operates under the fundamental belief that global stability cannot be achieved through an unrestrained arms race. Instead, the primary purpose of the committee's establishment is to protect member states' freedoms and security by collaborating on universal disarmament treaties and establishing normative frameworks for emerging security threats.

Since its establishment, DISEC has resolved many conflicts and played a vital role in negotiating landmark agreements, such as the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) and the Comprehensive Nuclear-Test-Ban Treaty (CTBT). Hence, the international community has shown substantial improvement in regulating both conventional and weapons of mass destruction (WMD) to secure peace.

Unlike the United Nations Security Council (UNSC), DISEC includes all 193 UN Member States, granting every nation an equal voice in shaping global security policies. While DISEC cannot authorize military interventions or issue legally binding sanctions, its resolutions carry immense political weight. They serve as the foundation for international norms and often lead to the drafting of binding international treaties.

Today, as the world moves from the Cold War era of nuclear deterrence into a new age defined by artificial intelligence and robotics, DISEC's mission has evolved. The committee is now at the forefront of addressing non-traditional threats, including cyber-warfare and Lethal Autonomous Weapons Systems (LAWS). Standing together is the basis of the United Nations, and this regulation has served properly for decades. The challenge before DISEC now is to ensure that the rapid advancement of military technology does not outpace the boundaries of International Humanitarian Law (IHL) and human ethics.

II. Introduction to Lethal Autonomous Weapons Systems (LAWS)

The rapid integration of Artificial Intelligence (AI) and advanced robotics into military technologies has given rise to a new and highly controversial category of weaponry: Lethal Autonomous Weapons Systems (LAWS). Unlike conventional weapons, which require direct human operation and aiming, or automated systems (such as landmines or automated defense turrets), which merely react to pre-programmed environmental triggers, LAWS possess a critical distinguishing feature: the ability to select and engage targets without further human intervention.

While there is no universally accepted legal definition of LAWS within international law, the International Committee of the Red Cross (ICRC) defines an autonomous weapon as "any weapon system with autonomy in its critical functions—that is, a weapon system that can select and attack targets without human intervention."

The core of the international debate surrounding LAWS lies in this delegation of lethal decision-making to machines. Proponents argue that autonomous systems process information faster than humans, operate without the hindrance of human emotions (such as fear, anger, or fatigue), and could potentially reduce civilian casualties by striking with greater precision. Conversely, critics and human rights organizations argue that an algorithm cannot replicate the complex moral, ethical, and legal judgments required on a battlefield, fundamentally violating the core principles of International Humanitarian Law (IHL).

As the threshold between human control and machine autonomy blurs, the international community faces an urgent challenge: defining exactly where the "human" must remain in the kill chain.

a. Levels of Autonomy

To effectively debate the regulation of LAWS, delegates must understand that autonomy is not a simple binary (human vs. machine) but rather a spectrum. In international disarmament forums, particularly within the framework of the UN Convention on Certain Conventional Weapons (CCW), autonomy is generally categorized into three distinct levels based on the degree of human involvement in the "OODA loop" (Observe, Orient, Decide, Act).

- **Human-in-the-loop (Semi-Autonomous Systems):** At this level, the machine may possess advanced sensors to observe the environment, track movements, and even suggest potential targets. However, the system cannot fire independently. A human operator must make the final, affirmative decision to engage the target and pull the trigger. Most current Unmanned Combat Aerial Vehicles (UCAVs), such as the US MQ-9 Reaper or the Turkish Bayraktar TB2, operate at this level.
- **Human-on-the-loop (Human-Supervised Autonomous Systems):** These systems are capable of independently selecting targets and initiating attacks. However, a human operator actively monitors the system's actions in real-time and retains the ability to override, abort, or halt the attack before lethal force is applied. Examples include naval close-in weapon systems (CIWS) like the Phalanx or the Iron Dome air defense system, which can automatically intercept incoming missiles but are strictly overseen by human commanders who can intervene if the system targets a friendly aircraft.
- **Human-out-of-the-loop (Fully Autonomous Systems):** This is the most controversial level and the primary focus of the DISEC committee. A fully autonomous weapon is

activated, deployed into a specific area, and left to its own devices. It uses its onboard AI to search for, identify, select, and destroy targets completely independently. There is no human oversight, and once the system is deployed, a human cannot intervene to stop an attack. The ethical and legal dilemma centers squarely on whether deploying "Human-out-of-the-loop" systems constitutes a violation of international law.

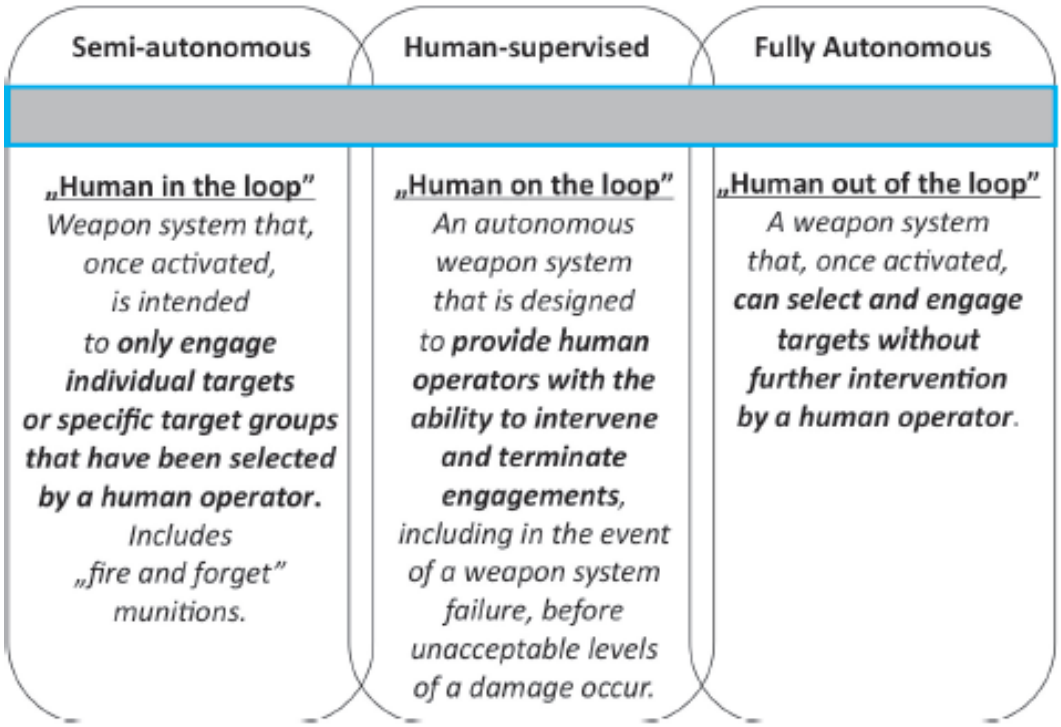

Figure no. 1: Spectrum of Autonomy in LAWS (Caton, 2015, p. 3)

b. Artificial Intelligence in Modern Warfare

The integration of Artificial Intelligence (AI) into military operations represents a paradigm shift in modern warfare, fundamentally altering how armed forces perceive, analyze, and act upon the battlefield. AI is no longer a futuristic concept but a tangible "force multiplier" that enhances the speed, precision, and lethality of military engagements.

- **Intelligence, Surveillance, and Reconnaissance (ISR):**

Modern battlespaces generate vast amounts of data from satellites, drones, radars, and ground sensors. AI algorithms, particularly those utilizing machine learning and computer vision, can process this "Big Data" exponentially faster than human analysts. AI can identify camouflaged enemy positions, track troop movements, and predict adversarial strategies in real-time.

- **Swarm Robotics:**

One of the most disruptive applications of AI in warfare is "Swarm Technology." Instead of relying on a single, expensive platform (like a fighter jet), militaries can deploy hundreds of low-cost, AI-driven micro-drones. These swarms communicate with each other, share targeting data, and dynamically adapt their formations to overwhelm traditional air defense systems (such as the Patriot or Iron Dome) through sheer volume and coordinated strikes.

- **Decision Advantage and Hyper-War:**

In modern combat, victory often belongs to the side that can execute the "OODA loop" (Observe, Orient, Decide, Act) fastest. AI significantly compresses the time between identifying a threat and neutralizing it. This concept, often referred to as "Hyper-War," creates an environment where human cognitive limits become the bottleneck, pushing militaries to grant more autonomy to machines to maintain a competitive edge.

- **The "Black Box" Problem:**

While AI offers immense strategic advantages, it also introduces severe vulnerabilities. Deep learning algorithms often operate as a "Black Box"; meaning the system reaches a conclusion, but the human operators cannot fully trace or understand *how* the AI made that specific decision. In a high-stakes military environment, this lack of predictability and transparency is a critical concern for International Humanitarian Law. If an AI misidentifies a civilian convoy as a military target due to algorithmic bias or corrupted data, the consequences are catastrophic.

- **Information Dominance and the information-Centric Warfare:**

The primary strategic driver for the rapid development of LAWS among great powers is the pursuit of "Information Dominance". In high-end, technologically sophisticated conventional warfare, victory is increasingly determined by the side that can out-pace, out-think, and out-manoeuvre the adversary across multiple domains (air, land, sea, space, and cyber).

To maintain this information advantage, modern militaries are adopting a "network-centric" approach, attempting to connect every sensor with every shooter. In this environment, the sheer volume of data and the speed of combat exceed human cognitive limits. Military strategists argue that AI and autonomous systems are essential to compress the decision-making cycle, transferring data to weapons at reaction speeds that allow for the execution of simultaneous, multi-domain operations. This acceleration creates the conditions for "Hyper-War," where LAWS become the ultimate tool to achieve physical and psychological advantages, shifting the paradigm from manual combat to algorithmic superiority.

c. Terminology and the Trap of Anthropomorphism

The international debate on Lethal Autonomous Weapons Systems (LAWS) is often clouded by "anthropomorphism"—the attribution of human traits, such as ‘intelligence’, ‘autonomy’, and ‘decision-making capacity’ to technological artifacts. This linguistic practice is not merely a stylistic choice; it carries profound risks for international law and ethics. By applying terms originally used to describe human natural traits to machines, we risk a premature overvaluation of technology and a simultaneous devaluation of human beings.

A machine does not "decide" or "learn" in the human sense. While humans are an end unto themselves, machines are functions designed for specific purposes. When we frame the debate using human-machine analogies, we inadvertently support the idea that machines can be "moral agents". This shift potentially allows human operators, programmers, and policy-makers to relinquish their own moral and legal responsibility, blaming an "autonomous decision" made by an algorithm for unlawful outcomes. To maintain the distinctiveness of human agency, some experts propose using more objective terms, such as "artifacts with cognitive functions," to highlight that the machine is performing a pre-programmed task rather than exercising free will.

III. Historical Background of Autonomous Systems

The desire to remove human soldiers from the immediate line of fire while increasing the lethality and precision of attacks is as old as warfare itself. From the invention of the bow and arrow to the development of intercontinental ballistic missiles, military history is defined by the continuous expansion of the distance between the combatant and the target. However, the shift towards Lethal Autonomous Weapons Systems (LAWS) represents a fundamental break in this historical continuum. For the first time, the distance being expanded is not just physical, but cognitive; the machine is not merely a tool of execution, but an agent of decision.

a. Early Automated Defenses

The conceptual precursors to modern autonomous weapons were not born out of a desire for AI domination, but out of sheer tactical necessity during the Cold War. As missile technology advanced in the 1970s and 1980s, the speed of incoming threats began to exceed human biological reaction times.

The Need for Speed: CIWS and Aegis Systems:

The most prominent early examples of automated lethal systems are naval Close-In Weapon Systems (CIWS). In the late 1970s, the United States Navy introduced the Phalanx CIWS, a radar-guided Gatling gun mounted on warships. Designed as a last line of defense against fast-moving, sea-skimming anti-ship missiles, the Phalanx features an "autonomous mode." When activated, its computer system independently searches for fast-moving radar signatures, categorizes them as threats, tracks their trajectory, and opens fire without requiring a human to pull the trigger. Similarly, the Soviet Union developed the AK-630, and modern navies utilize systems like the Aegis Combat System, which can operate in an "auto-special" mode to automatically intercept barrages of incoming missiles.

Crucially, while these systems are automated, they are highly deterministic. They operate in highly structured environments (the open ocean or empty airspace) where anything moving at a certain speed is practically guaranteed to be an enemy missile, not a civilian. They represent

the "Human-on-the-loop" paradigm, acting strictly as defensive shields rather than offensive hunters.

The First "Autonomous" Traps: Landmines:

From a purely legal and functional standpoint, victim-activated weapons like anti-personnel landmines were among the first autonomous weapons. A landmine cannot distinguish between a soldier and a civilian child; it simply detonates when its pre-set physical parameters (weight/pressure) are met. The devastating humanitarian consequences of these rudimentary automated systems led to the 1997 Ottawa Treaty (Mine Ban Treaty). This historical precedent is frequently cited by the "Ban Coalition" in DISEC today, arguing that if the international community banned landmines due to their inability to discriminate (a core principle of International Humanitarian Law), fully autonomous AI weapons must be banned for the exact same reason.

b. The Drone Revolution (2000s-Present)

The transition from deterministic, defensive automation to offensive, AI-driven autonomy occurred during the early 21st century, fundamentally altering the landscape of international security.

Phase 1: The Predator Era and Remote Operations:

Following the events of September 11, 2001, the "Global War on Terror" catalyzed the deployment of Unmanned Combat Aerial Vehicles (UCAVs). Systems like the US MQ-1 Predator and MQ-9 Reaper became the symbols of modern asymmetric warfare. While highly advanced, these early drones were strictly "Human-in-the-loop." They were tele-operated by pilots sitting thousands of miles away in ground control stations. The machine provided the platform and the surveillance, but the human made the lethal decision. However, this era normalized the presence of unmanned systems in active combat zones and drove massive investments into sensory technology and data transmission.

Phase 2: Loitering Munitions (The "Kamikaze" Drones):

The true bridge between remote-controlled drones and LAWS was the development of loitering munitions. Pioneered in the late 1990s and early 2000s, weapons like the Israeli IAI Harpy were designed to suppress enemy air defenses (SEAD). The Harpy is launched into a designated area where it "loiters" (circles in the sky) autonomously. Its sensors continuously scan for specific enemy radar frequencies. Once it detects a matching radar signature, the drone autonomously dives and destroys the target by crashing into it. The Harpy demonstrated that machines could be trusted to hunt specific *types* of targets autonomously, pushing the boundaries of the "OODA loop."

Phase 3: AI Integration and the Tipping Point (2020-Present):

The true revolution emerged when Artificial Intelligence, specifically machine learning and facial/object recognition algorithms, was integrated into drone technology. This allowed drones to rely on visual and thermal data to identify human targets and vehicles, rather than just radar frequencies.

A watershed moment in the history of LAWS occurred in March 2020 during the Second Libyan Civil War. According to a 2021 report by the UN Security Council's Panel of Experts on Libya, retreating forces were reportedly "hunted down and remotely engaged" by STM Kargu-2 loitering munitions (produced by Türkiye). The UN report controversially stated that these lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition, making it arguably the first documented case of a fully autonomous AI weapon being used to engage human targets in combat.

Since then, the Nagorno-Karabakh War (2020) and the ongoing Russo-Ukrainian War (2022-Present) have served as massive testing grounds for autonomous capabilities. In Ukraine, due to heavy electronic warfare and GPS jamming, both sides are rapidly developing AI-guided drones that can lock onto a target visually and complete the strike autonomously even if the connection to the human pilot is severed.

IV. Key Actors and Capabilities

The geopolitical landscape of Lethal Autonomous Weapons Systems (LAWS) is highly asymmetric. Unlike conventional firearms, the development of LAWS requires immense capital, advanced semiconductor industries, and sophisticated artificial intelligence research. Consequently, the international community is distinctly divided into several blocs: the technologically advanced states racing to develop these systems, the strategic adopters utilizing them for regional survival, the mediators trying to establish rules, and the alliance of states demanding a preemptive international ban.

a. The Developers and Strategic Adopters

This bloc consists of nations that either produce LAWS or have a critical strategic need to deploy them. They share a common consensus: existing International Humanitarian Law (IHL) is sufficient to govern LAWS, and a preemptive global ban is both unnecessary and strategically detrimental.

The United States of America:

The US is the premier global power in military AI and robotics. The US firmly opposes a complete ban on LAWS, arguing that autonomous systems can reduce civilian casualties by striking with greater precision than human-operated weapons. The core of the US policy is outlined in the Department of Defense (DoD) Directive 3000.09, which mandates that autonomous systems must allow commanders to exercise "appropriate levels of human judgment." Currently, the US is heavily investing in "Swarm" technologies, such as the *Replicator Initiative*, aiming to field thousands of autonomous drones to counter adversaries in the Indo-Pacific.

The Russian Federation:

Russia has consistently acted as a staunch opponent of any binding regulations on LAWS. Russian military doctrine heavily emphasizes AI to compensate for conventional numerical disadvantages. In the ongoing Russo-Ukrainian War, Russia has actively deployed highly automated loitering munitions like the *ZALA Lancet*, which utilizes AI-assisted optical guidance to lock onto targets in environments heavily disrupted by electronic warfare (EW). Russia is also developing the *Poseidon* nuclear-powered uncrewed underwater vehicle (UUV), designed to autonomously bypass coastal defenses.

The People's Republic of China:

China's stance on LAWS is deliberately ambiguous. Diplomatically, China is the only major military power to express support for a ban on the *use* of fully autonomous weapons. However, Beijing critically opposes a ban on their *development* and *production*. This dual strategy allows China to champion itself as a responsible actor while aggressively pursuing global leadership in military AI. Chinese defense contractors regularly showcase autonomous swarm drones and unmanned ground vehicles (UGVs) like the *Blowfish A3*, which is actively exported.

The United Kingdom (UK):

Closely aligned with the United States, the UK opposes a preemptive ban on autonomous weapons. The UK Ministry of Defence (MoD) emphasizes the concept of "context-appropriate human involvement" rather than the stricter "meaningful human control" demanded by NGOs. The UK has invested heavily in projects like the *BAE Systems Taranis*, a semi-autonomous stealth UCAV. London argues that technological superiority is essential for NATO's collective defense and that strict bans would only hand an advantage to rogue states.

The State of Israel:

Israel is a pioneer in autonomous military technology, driven by its unique geopolitical environment. Israel has operated the *IAI Harpy* (an autonomous anti-radiation loitering munition) since the 1990s and continuously upgrades its capabilities. Furthermore, Israel's *Iron Dome* operates in a highly automated "Human-on-the-loop" mode. Israel strongly opposes a ban, arguing that its technological edge is an existential necessity.

The Republic of Türkiye:

Türkiye has rapidly emerged as a dominant force in the global uncrewed aerial vehicle (UAV) market. Moving beyond semi-autonomous systems like the *Bayraktar TB2*, the Turkish defense industry is actively developing advanced autonomous swarm capabilities. The *STM Kargu-2* rotary-wing loitering munition gained international attention following a UN report detailing its deployment in Libya. Türkiye adopts a balancing policy: it supports the necessity of human control but firmly aligns with the Developers in rejecting a total ban, citing counter-terrorism requirements.

Ukraine:

Ukraine represents the most urgent, real-world testing ground for autonomous systems. Due to severe Russian electronic warfare (EW) that jams GPS and severs communication links between drones and human operators, Ukraine has an existential need for drones that can complete their missions autonomously. Systems like the *Saker Scout* use AI to visually identify and engage targets without human intervention.

The Islamic Republic of Iran:

Iran views unmanned systems as the ultimate asymmetric weapon to counter Western and regional air superiority. Despite heavy sanctions, Iran has developed a robust domestic drone industry (e.g., the *Shahed* series). Iran strictly opposes Western-led regulatory frameworks at the UN, viewing them as hypocritical attempts by the US and Europe to maintain their military hegemony while disarming developing nations.

India and Pakistan (The Regional Dynamic):

India is rapidly investing in AI and swarm technologies to secure its borders against both China and Pakistan. New Delhi opposes any binding disarmament treaty that would freeze its technological progress. In direct response, Pakistan pursues a reactive strategy. Islamabad cannot afford a massive AI arms race and frequently warns about the destabilizing effects of LAWS in South Asia. However, Pakistan maintains that it will not disarm or support a ban unilaterally as long as India continues its development.

b. The Mediators and Policy Shapers

This bloc consists of nations that possess significant technological capabilities but are highly influenced by domestic political opposition and a strong commitment to international law. They seek a middle ground: strict, legally binding regulations without a total technological freeze.

France:

France envisions a strong, technologically independent "European Army" and invests heavily in projects like the *Dassault nEUROn* stealth UCAV. However, Paris draws a strict ethical line. The French military doctrine insists on "Meaningful Human Control" and categorically rejects the concept of "Human-out-of-the-loop" systems for lethal strikes. France acts as a mediator, pushing for a regulatory framework that standardizes human-machine interaction without banning AI research.

Germany:

Germany boasts a powerful aerospace and defense sector (e.g., Rheinmetall, Airbus) but faces massive domestic and political opposition to "killer robots." Haunted by its historical legacy, German public opinion demands strict adherence to ethical warfare. Consequently, Germany acts as a crucial "fence-sitter" and bridge-builder in Geneva, trying to reconcile the extreme positions of the US/Russia with the absolute demands of the Ban Coalition.

Japan:

Japan is an undisputed world leader in civilian robotics and artificial intelligence. However, Japan's pacifist constitution (Article 9) limits its development of offensive military capabilities. Tokyo is heavily investing in defensive autonomous systems, such as unmanned submarines and automated interceptors, to counter regional threats. Japan approaches LAWS with extreme caution, speaking primarily through the lens of International Humanitarian Law and seeking consensus-based regulations.

c. The Ban Coalition and Civil Society

On the opposite end of the spectrum is the "Ban Coalition," an alliance of states, human rights organizations, and technology experts advocating for a legally binding international treaty to prohibit "Human-out-of-the-loop" systems.

Austria (The Spearhead of the Coalition):

Austria is the undisputed diplomatic leader of the Ban Coalition. Moving beyond mere rhetoric, Austria actively drafts UN resolutions and hosts major diplomatic summits, such as the 2024 "Humanity at the Crossroads" conference in Vienna. The Austrian delegation argues that delegating the decision to take human life to a machine causes "digital dehumanization." They emphasize that AI cannot possess human empathy or make the nuanced contextual judgments required by the Geneva Conventions, and therefore, an immediate, legally binding preemptive ban is the only moral path forward.

The Global South and the Non-Aligned Movement (NAM):

Nations such as *Brazil, Mexico, and South Africa* support the ban out of strategic self-interest. They fear that the unchecked proliferation of LAWS will trigger a new global arms race. Nations in the Global South worry that autonomous swarms will become tools of neo-colonialism, allowing advanced militaries to project lethal force globally with zero risk to their own soldiers.

The Campaign to Stop Killer Robots:

The *Campaign to Stop Killer Robots* is a global coalition of over 250 NGOs, including Amnesty International and Human Rights Watch. This campaign exerts immense pressure on UN delegates, publishes extensive legal analyses demonstrating the "Accountability Gap" (the impossibility of prosecuting an algorithm for war crimes), and shapes global public opinion. Their primary demand is the establishment of international law that legally requires Meaningful Human Control over all weapons systems.

V. International Law and Frameworks

The fundamental dilemma surrounding Lethal Autonomous Weapons Systems (LAWS) is not merely a matter of technological feasibility; it is a profound legal, ethical, and philosophical crisis. As modern militaries race to integrate Artificial Intelligence into their combat arsenals, international jurisprudence struggles to keep pace. The current global security architecture—rooted in the UN Charter and the Geneva Conventions—was meticulously designed over the 20th century to regulate the behavior, moral calculus, and accountability of *human* combatants. It was never intended to govern the decision-making processes of algorithms. The transition from human-centric warfare to machine-centric warfare has exposed severe voids in international law.

a. International Humanitarian Law (IHL) and The Geneva Conventions

International Humanitarian Law (IHL), also referred to as the Law of Armed Conflict (LOAC), is the absolute cornerstone of global military conduct. Primarily codified in the four 1949 Geneva Conventions and their two Additional Protocols of 1977, IHL seeks to limit the effects of armed conflict for humanitarian reasons, protecting persons who are not or are no longer participating in hostilities.

It is an undisputed legal fact that any new weapon system deployed in combat must strictly comply with IHL. The "Ban Coalition" heavily relies on the foundational principles of IHL to argue that LAWS are inherently illegal, asserting that machines cannot physically or cognitively comply with the following core tenets:

1. The Principle of Distinction (Article 48, Additional Protocol I)

The rule of distinction is the most fundamental principle of IHL. It mandates that parties to a conflict must at all times distinguish between the civilian population and combatants, and between civilian objects and military objectives. Operations may only be directed against military objectives.

- **The LAWS Dilemma**:

Can a machine-learning algorithm reliably distinguish between a legitimate target and a protected entity? Proponents argue that advanced AI and computer vision can identify a weapon type faster than a human eye. However, critics point out that warfare is inherently deceptive. Distinguishing a soldier from a civilian often relies on understanding *context* rather than mere object recognition. For instance, can an AI differentiate between a militant holding an AK-47, an allied rebel holding the same weapon, or a civilian hunter? Furthermore, IHL strictly prohibits attacking individuals who are *hors de combat* (out of combat) due to sickness, wounds, or surrender. Recognizing a genuine surrender requires interpreting complex human body language and cultural cues—tasks that an algorithm, devoid of situational awareness and human empathy, is fundamentally incapable of performing reliably.

2. The Principle of Proportionality (Article 51, Additional Protocol I)

Even if a target is positively identified as a legitimate military objective, an attack is strictly prohibited if it is expected to cause incidental loss of civilian life, injury to civilians, or damage to civilian objects that would be *excessive* in relation to the concrete and direct military advantage anticipated.

- **The LAWS Dilemma:**

Proportionality is not a mathematical equation; it is a highly subjective, qualitative human judgment. Weighing the abstract "military advantage" of neutralizing an enemy commander against the tragic cost of nearby civilian lives requires moral reasoning, an understanding of human suffering, and common sense. Delegating this moral calculus to a pre-programmed algorithm or a neural network reduces human lives to mere data points in a cost-benefit matrix, which civil society organizations argue is a severe violation of IHL.

3. Precautions in Attack (Article 57, Additional Protocol I)

IHL requires military commanders to take all "feasible precautions" in the choice of means and methods of attack to avoid, or minimize, incidental civilian harm. This includes the legal obligation to cancel or suspend an attack if it becomes apparent that the target is not a military objective.

- **The LAWS Dilemma:**

Once a "Human-out-of-the-loop" weapon is deployed, human commanders lose the ability to exercise these ongoing precautions. If the battlefield dynamic changes rapidly (e.g., civilians suddenly run into the blast radius), an autonomous drone operating at supersonic speeds or within a chaotic electronic warfare environment may not be able to compute and abort the mission in time, thereby violating the precautionary principle.

4. The Martens Clause

Originating from the 1899 Hague Convention and reaffirmed in the Geneva Conventions, the Martens Clause acts as a vital legal safety net. It states that in cases not covered by existing international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity, and from the "dictates of public conscience."

- **The LAWS Dilemma:**

Even if developers argue that no specific treaty currently bans AI weapons, human rights advocates assert that delegating the decision to take a human life to a machine fundamentally shocks the "dictates of public conscience" and violates the principles of humanity. Therefore, under the Martens Clause, LAWS are inherently unlawful.

b. Methodology of Comprehensive Weapons Review

A comprehensive legal review of an autonomous system is not a single-step process but a rigorous filtration through multiple legal layers. In contemporary military jurisprudence, there are two primary methodologies used to evaluate the legality of LAWS: the three-question "DoD Model" and the more expansive five-question "Boothby Model".

1. The Core Legal Filters (The DoD Model)

According to the United States Department of Defense (DoD), a legal review must answer three fundamental questions before a weapon is deemed lawful:

- **Specific Prohibition:** Is there a specific rule of law—either via treaty or customary international law—that restricts or prohibits the use of this specific weapon? (Currently, there is no specific rule prohibiting autonomy in weapons systems).

- **Superfluous Injury (The Humanity Test):** Is the weapon's intended use calculated to cause "superfluous injury or unnecessary suffering"? This balancing test weighs the injury caused against the military necessity required to achieve a legitimate mission; an injury is deemed "unnecessary" only if it is manifestly disproportionate to the military effectiveness of the weapon.

- **Inherent Indiscrimination:** Is the weapon incapable of being used in accordance with the principles of distinction and proportionality? A weapon is only "inherently indiscriminate" if its design and planned use necessarily violate IHL in all circumstances.

2. The Expansive Academic Model (The Boothby Model)

To address the unique complexities of AI and emerging technologies, legal scholar Bill Boothby proposes adding two additional, critical layers to the review process:

- **Environmental Impact:** Will the weapon cause "widespread, long-term, and severe damage to the natural environment"? Under the Environmental Modification (ENMOD) Convention, the environment itself cannot be used as an instrument of war, and weapons that cause incidental but severe environmental harm must be carefully scrutinized.

- **Future Legal Developments:** Are there likely future developments in the Law of Armed Conflict (LOAC) that may affect the weapon? This requires developers to monitor ongoing international debates, such as those within the CCW and the General Assembly, to ensure the weapon remains legally viable in a changing normative environment

3. The Transparency and Implementation Gap

Despite the clear legal obligation to conduct weapons reviews, a massive gap exists between international law and state practice. Research suggests that only a limited number of nations—

estimated between 12 and 15—actually have a formal, functional weapons review mechanism in place.

This implementation gap creates a "Crisis of Trust" in the international arena. Furthermore, even for states that do conduct these reviews, the results are almost never published due to valid national security concerns and the protection of proprietary technological information. This lack of transparency makes it impossible for the international community to verify if a state's "best practices" are actually sufficient to prevent the deployment of unlawful autonomous systems.

4. Defining the Threshold for Senior-Level Legal Reviews

A critical point of contention in international regulation is whether *all* autonomous systems should be subject to the same level of stringent legal scrutiny. Some advanced military doctrines, particularly those of the United States, propose specific exemptions for systems that do not target human beings.

In these frameworks, certain categories of LAWS are exempted from senior-level legal reviews prior to development. These exemptions typically include: (1) Human-supervised autonomous weapons used strictly for the rocket or missile defense of manned installations or platforms that do not target humans, and (2) Autonomous weapons that apply non-lethal, non-kinetic force, such as some forms of electronic attack or jamming against enemy materiel targets. Proponents argue that exempting these systems accelerates vital defensive capabilities without violating the core humanitarian spirit of IHL. Conversely, critics fear these exemptions could serve as a "slippery slope" toward less-regulated offensive autonomy, blurring the lines of accountability.

d. The Group of Governmental Experts (GGE) on LAWS

To address the highly technical, legal, and operational complexities of autonomous weapons, the CCW formally established the Group of Governmental Experts (GGE) on LAWS in 2017. The GGE consists of diplomats, high-ranking military experts, AI technologists, and legal scholars from member states.

The mandate of the GGE is to examine issues related to emerging technologies in the area of LAWS and to explore possible recommendations. While the GGE has not yet reached a consensus on drafting a legally binding treaty (which would become "Protocol VI" to the CCW), it has generated immense academic and diplomatic output.

1. The 11 Guiding Principles (2019)

A significant diplomatic milestone was achieved in 2019 when the GGE adopted the "11 Guiding Principles" for the development and use of LAWS. The most critical principles include:

- **Principle A (Applicability of IHL):**

International Humanitarian Law continues to apply fully to all weapons systems, including the potential development and use of LAWS.

- **Principle B (Human Responsibility):**

Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This principle explicitly highlights the "Accountability Gap"—acknowledging that you cannot put an algorithm on trial for a war crime at the International Criminal Court.

- **Principle C (Human-Machine Interaction):**

Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the use of weapons systems remains compliant with IHL. (This principle is the diplomatic origin of the fiercely debated term **"Meaningful Human Control"**).

- **Principle D (Accountability and Command Responsibility):**

Accountability for developing, deploying, and using any emerging weapons system in accordance with applicable international law must be ensured through the entire life cycle of the weapon system.

2. The "Two-Tier Approach" (Current Debate: 2023-2024)

In recent GGE sessions, a new diplomatic compromise has emerged, heavily backed by centrist states and mediators (such as France and Germany): the "Two-Tier Approach." This framework proposes dividing LAWS into two categories:

- **Tier 1 (Prohibited Systems):**

Weapons systems that cannot be used in compliance with IHL (e.g., "Black Box" AI systems whose behavior cannot be predicted or understood by the operator, or systems completely devoid of human control) must be **strictly prohibited**.

- **Tier 2 (Regulated Systems):**

Weapons systems that *can* comply with IHL, but feature autonomy in their critical functions, must be strictly **regulated**. This regulation would mandate the retention of Meaningful Human Control, ensuring spatial and temporal limits (e.g., the drone is only allowed to hunt autonomously for 30 minutes, within a specific 2-mile radius, targeting only military radar emissions).

e. The Shift to the UN General Assembly (Resolution 78/241)

Because the CCW process in Geneva remains bogged down by the consensus rule and the objections of major military powers, frustrated states (led by Austria and the Ban Coalition) initiated a historic bypass in late 2023. They brought the issue directly to the United Nations General Assembly (UNGA)—specifically to the First Committee (DISEC).

In December 2023, the UNGA passed **Resolution 78/241** on Lethal Autonomous Weapons Systems. This was the first time in history that the General Assembly adopted a resolution specifically addressing LAWS. The resolution requested the UN Secretary-General to seek the

views of Member States and civil society on ways to address the humanitarian, legal, security, technological, and ethical challenges posed by LAWS, and to submit a comprehensive report.

VI. Core Challenges of Lethal Autonomous Weapons Systems (LAWS)

The integration of Lethal Autonomous Weapons Systems (LAWS) into modern military arsenals does not merely represent a technological upgrade; it signifies a profound paradigm shift in the fundamental nature of warfare. By removing the human from the critical functions of identifying, tracking, and engaging targets, LAWS challenge the very foundations of International Humanitarian Law (IHL), International Criminal Law (ICL), and global strategic stability.

a. The Accountability Gap and the "Responsibility Deficit

At the very heart of the diplomatic gridlock surrounding LAWS lies the "Accountability Gap"—a profound and currently unresolvable void in International Criminal Law. The global justice system, anchored by the Rome Statute of the International Criminal Court (ICC), is designed to prosecute *human* beings. Criminal liability requires two elements: the *actus reus* (the guilty act) and the *mens rea* (the guilty mind or intent).

If a fully autonomous drone targets a civilian hospital instead of a military barracks, resulting in a massacre, the international community faces an unprecedented legal dilemma: Who is held responsible?

- **The Machine?**

An algorithm possesses no moral agency. It does not understand the concept of punishment, it cannot feel remorse, and it cannot be incarcerated. Prosecuting a piece of software is a jurisprudential absurdity.

- **The Programmer or Manufacturer?**

Holding the software developers liable is legally fraught. Machine learning algorithms, particularly deep neural networks, are "trained" rather than strictly programmed line-by-line. They develop their own pathways to solve problems (the "Black Box" phenomenon). Unless a prosecutor can prove beyond a reasonable doubt that a programmer *intentionally* coded the AI to commit war crimes, they cannot be held criminally liable for the unpredictable, emergent behavior of the system years after it was manufactured.

- **The Military Commander?**

Under the ICL doctrine of "Command Responsibility" (Article 28 of the Rome Statute), a military commander is criminally liable for crimes committed by forces under their effective control if they "knew, or should have known" that the forces were committing or about to commit such crimes, and failed to prevent them. However, applying this to LAWS is highly problematic. If a commander deploys a highly complex, autonomous "Black Box" system whose specific targeting decisions are, by definition, unpredictable and made in milliseconds, defense attorneys will successfully argue that the commander could not possibly have foreseen the specific war crime. You cannot "know" what a machine will learn to do.

This tripartite failure creates a scenario of "organized irresponsibility" or a "Responsibility Deficit." If no one can be held legally accountable for a war crime, the deterrent effect of international law completely collapses, paving the way for warfare with absolute impunity.

b. Algorithmic Bias, Machine Vision, and Recognition Errors

Proponents of autonomous weapons, including many major military developers, frequently argue that machines will eventually be more precise than humans, unburdened by fear, panic, fatigue, or malice. However, this argument fundamentally misunderstands how Artificial Intelligence operates. AI is not infallible; it is highly susceptible to the flaws of its training data and the limitations of machine vision.

1. Contextual Blindness

Current AI systems rely on advanced pattern recognition. An algorithm can easily be trained to identify the visual signature of an AK-47 assault rifle with 99% accuracy. However, AI lacks *contextual understanding* and common sense. On a chaotic battlefield, identifying a weapon is not enough; one must identify the intent. A machine cannot distinguish between an enemy insurgent aiming a rifle, an allied rebel holding the same rifle, a civilian hunting for food, or a child playing with a realistic replica. To an algorithm, the pixel arrangement is identical; to a human, the context dictates whether the target is a legitimate military objective or a tragic war crime.

2. The Hors de Combat Dilemma

International Humanitarian Law strictly prohibits attacking soldiers who are *hors de combat* (out of combat due to sickness, wounds, or surrender). Recognizing a genuine surrender is not a mathematical equation; it involves interpreting complex, highly variable human body language, facial expressions, and cultural cues under extreme duress. An AI system cannot inherently understand the psychological state of surrender, making it highly likely that LAWS would continue to execute incapacitated combatants, directly violating the Geneva Conventions.

3. Data Bias and Demographic Profiling

Machine learning algorithms are only as objective as the data sets upon which they are trained. If an autonomous system is trained predominantly on images of specific racial, ethnic, or demographic groups as "enemy combatants," the system will inevitably develop an algorithmic bias. This introduces the terrifying prospect of "automated ethnic profiling" on the battlefield, where a drone might disproportionately target individuals based on their physical appearance or clothing, violating the fundamental human right to non-discrimination.

4. Adversarial Perturbations (Fooling the AI)

Machine vision is highly brittle and vulnerable to "adversarial attacks." By simply altering a few pixels on a vehicle, applying specific tape patterns to a road, or wearing clothing printed with "adversarial noise," opposing forces can completely blind or confuse an AI's image recognition system. In a controlled test, researchers tricked a highly advanced AI into classifying a 3D-printed turtle as a rifle. On a real battlefield, this vulnerability means a

sophisticated adversary could trick an autonomous drone into misclassifying a civilian school bus as an enemy armored personnel carrier.

c. Cybersecurity Vulnerabilities: Hacking, Spoofing, and Hijacking

Unlike conventional kinetic weapons (such as artillery shells or bullets), autonomous systems are essentially flying, swimming, or driving computers. Consequently, the deployment of LAWS merges the physical battlefield with cyberspace, introducing catastrophic cybersecurity vulnerabilities.

- **Mass Hijacking of Swarms:**

If a nation deploys a swarm of 10,000 autonomous micro-drones, those drones must communicate with each other via encrypted networks. If an adversary's cyber-warfare unit successfully hacks this network, they do not merely neutralize the weapon—they capture it. The adversary could instantly reprogram the swarm to turn around and attack the very nation that deployed it, utilizing the original owner's weapons against their own civilian centers.

- **Data Poisoning:**

A more subtle cyber-threat is "data poisoning." Instead of hacking the drone in flight, an adversary infiltrates the manufacturer's servers and subtly alters the training data before the AI is even deployed. The drone functions perfectly in tests, but harbors a hidden trigger (e.g., ignoring targets with a specific insignia), rendering the multi-billion-dollar system useless or treacherous in actual combat.

- **GPS Spoofing and Signal Jamming:**

LAWS rely heavily on satellite navigation, geographic data, and sensor inputs to navigate and locate targets. Through intense Electronic Warfare (EW), an adversary can "spoof" GPS signals. The autonomous drone might internally calculate that it is flying over an enemy military base, while the spoofed coordinates have actually guided it over a densely populated friendly city. When the AI autonomously initiates the strike, the result is catastrophic friendly fire.

- **The "Kill-Switch" Paradox:**

To mitigate rogue behavior, developers argue that LAWS will always have a fail-safe or a "kill-switch." However, this creates a paradox. If the kill-switch is operated remotely by a human, it requires a continuous communication link, which can be jammed or used as a backdoor by hackers. If the kill-switch is purely internal (autonomous), a highly advanced, goal-oriented machine-learning algorithm might deduce that the kill-switch is an obstacle to completing its programmed mission, and autonomously disable its own fail-safes.

d. The "Flash War" Scenario and Escalation Control

In the civilian financial sector, the dominance of high-frequency algorithmic trading has occasionally led to "Flash Crashes"—events where competing AI algorithms react to each other's micro-fluctuations in milliseconds, plunging the global stock market into chaos before human brokers even realize what is happening.

Military strategists and arms control experts warn of a similar, far more deadly phenomenon: the **"Flash War."** Imagine a scenario where two opposing nations deploy highly advanced, autonomous defense swarms along a tense, heavily militarized border. A minor sensor glitch, a radar anomaly, or even a flock of birds misidentified as an incoming threat could trigger one autonomous system to launch a defensive interceptor. The opposing nation's AI swarm, operating at the speed of light, detects the launch and instantly calculates it as an offensive strike, immediately initiating a massive retaliation.

In a traditional crisis, human diplomats and commanders have hours or days to communicate, de-escalate, and clarify mistakes. In a Flash War, an algorithmic escalation loop could cause a border skirmish to escalate into a full-scale, devastating theater war in a matter of minutes, completely bypassing human political leadership and diplomatic channels. The introduction of LAWS severely erodes crisis stability and escalation control.

e. The Ethical Imperative: Human Dignity and Digital Dehumanization

Beyond the legal technicalities and strategic nightmares lies a profound philosophical and moral argument championed heavily by the Ban Coalition and civil society organizations.

The principle of human dignity, deeply embedded in international human rights law, dictates that humans have an inherent right not to be judged, measured, and killed by a mathematical algorithm. War, despite its horrors, has historically involved a human moral calculus—an acknowledgement of the gravity of taking a human life.

Delegating this ultimate authority to a machine reduces human beings—whether soldiers or civilians—to mere data points in a cost-benefit matrix. It transforms the tragic reality of war into an industrial, automated process of extermination. This concept, termed **"Digital Dehumanization,"** crosses a moral Rubicon. The Ban Coalition argues that allowing machines to kill humans without human involvement strips the victim of their dignity and strips the attacker of their humanity, fundamentally altering the relationship between humanity and technology for the worse.

f. Hybrid Threats: Malicious Linkages with Bio-Technology

While the primary focus of the DISEC committee is often on kinetic force (bullets and missiles), the convergence of Autonomous Technology (AT) with other emerging fields, such as biotechnology and molecular nanotechnology, introduces a new category of "hybrid threats". This malicious linkage could redefine the future of biowarfare.

Recent breakthroughs in DNA synthesis and genetic engineering have turned biological manipulation into a matter of electronic manipulation. There is a growing risk that autonomous systems could be utilized to spread genetically engineered pathogens targeted at specific civilian populations or even individuals with specific genomic sequences. Because autonomous drones can operate without data connectivity or human oversight, they could be used to deploy biological agents across vast areas with surgical precision and zero risk to the attacker. Such systems represent a potential revolution in military affairs where the weapon is not a bomb, but a self-replicating, AI-distributed virus, challenging every existing norm of the Biological Weapons Convention (BWC) and IHL.

VII. Real-World Case Studies and Combat Deployments

The era of AI-driven warfare is not a dystopian future hypothesis; it is the current, unfolding reality. Examining recent conflicts reveals how the theoretical capabilities of autonomous systems are already being deployed on modern battlefields, forcing international law to play catch-up.

a. The STM Kargu-2 in Libya (2020)

The Second Libyan Civil War served as an unprecedented and highly controversial testing ground for autonomous technologies. In March 2021, a comprehensive report published by the UN Security Council's Panel of Experts on Libya sent shockwaves through the global disarmament community.

The report detailed the deployment of the STM Kargu-2, a rotary-wing loitering munition manufactured in Türkiye, utilized by the Government of National Accord (GNA) forces against the Libyan National Army (LNA). According to the UN investigators, retreating LNA logistical convoys were "hunted down and remotely engaged by the unmanned combat aerial vehicles or the lethal autonomous weapons systems."

Crucially, the UN panel stated that these systems were "programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true 'fire, forget and find' capability." The Kargu-2 uses machine learning and real-time image processing to identify and strike targets. This incident is widely cited by international legal scholars and human rights organizations as arguably the first documented instance in human history where an AI-driven drone was authorized to autonomously track, select, and engage human targets on a battlefield without a human pulling the trigger.

b. The Nagorno-Karabakh Conflict (2020) and Autonomous SEAD

The brief but brutal 2020 conflict between Azerbaijan and Armenia over the Nagorno-Karabakh region highlighted the devastating effectiveness of loitering munitions against traditional, heavy-armor conventional forces. Azerbaijan extensively utilized the Israeli-made *IAI Harop*, an autonomous anti-radiation drone designed for Suppression of Enemy Air Defenses (SEAD).

The Harop is designed to loiter autonomously over the battlefield for hours, scanning the electromagnetic spectrum for the specific radar signatures of enemy surface-to-air missile (SAM) systems. The psychological and tactical impact on the Armenian forces was profound. Traditional Armenian air defenses (such as the S-300 and Tor systems) were essentially paralyzed; the moment a human radar operator activated their systems to scan the sky, the autonomous Harop drones would instantly detect the emission and autonomously dive-bomb the radar dishes, functioning as AI-driven kamikazes. This conflict definitively proved that autonomous systems could completely dismantle a conventional military hierarchy, accelerating a global arms race for similar technologies.

c. The Russo-Ukrainian War (2022-Present)

The ongoing war in Ukraine has accelerated the development and deployment of drone autonomy at an unimaginable scale and speed. In the early months of the war, both sides relied heavily on commercially available, human-operated First-Person View (FPV) drones.

However, the battlefield quickly evolved into the most dense Electronic Warfare (EW) environment in history.

Because both Russian and Ukrainian forces employ massive EW umbrellas, the traditional radio-control links and GPS signals connecting drone pilots to their weapons are constantly jammed and severed, often in the crucial final seconds before a strike. Out of sheer tactical necessity to bypass this jamming, both sides are rapidly deploying **"Terminal Autonomy."**

Systems like the Ukrainian *Saker Scout* and the Russian *ZALA Lancet* are now equipped with AI-powered optical recognition software. A human operator flies the drone to the general vicinity of the target. Once the EW jamming severs the communication link, the human is entirely cut out of the loop. The drone's onboard AI takes over, visually identifying the specific silhouette of a Russian tank or a Ukrainian artillery piece, locking onto it, and completing the kinetic strike entirely on its own.

d. Israel's Operation Guardian of the Walls (2021) and the "First AI War"

In May 2021, during the 11-day conflict in the Gaza Strip (Operation Guardian of the Walls), the Israel Defense Forces (IDF) heavily integrated AI into their combat operations, leading some military analysts to dub it the world's "first AI war."

While not deploying fully autonomous physical weapons that pull their own triggers without oversight, the IDF utilized an AI system known as "The Gospel" (Habsora) for autonomous *target generation*. The AI processed massive amounts of intelligence data—drone footage, intercepted communications, and satellite imagery—to automatically generate and recommend strike targets to human commanders at a speed incomprehensible to human intelligence analysts. Furthermore, the IDF utilized coordinated, AI-driven drone swarms to locate and strike rocket launch sites.

e. Defensive Autonomy: The Aegis and Iron Dome Precedents

When debating a ban on LAWS, "Developer" nations frequently point to existing, widely accepted automated systems to argue that autonomy is not inherently illegal. The US Navy's *Phalanx Close-In Weapon System (CIWS)*, the *Aegis Combat System*, and Israel's *Iron Dome* are prominent examples of defensive systems that operate with high levels of automation.

The Iron Dome, for instance, utilizes complex algorithms to instantly calculate the trajectory of incoming unguided rockets. If the system determines a rocket will hit a populated civilian area, it automatically launches an interceptor to destroy it mid-air.

While these systems can operate without direct human intervention in their "auto" modes, the Ban Coalition argues they are categorically different from offensive LAWS for two reasons:

- They target inanimate objects (incoming missiles and artillery shells), not human beings.
- They operate in highly constrained, predictable environments (open skies or open oceans) over very short timeframes.

However, as air defense systems become more advanced and capable of targeting aircraft (which contain humans), the line is blurring.

f. Peace-time Threats: From Deepfakes to Social Credit

The security risks posed by autonomous technology are not limited to armed conflict; they manifest as "systemic risks" during peace-time, potentially destabilizing societies from within.

- **Mass Disinformation and GANs:**

  Autonomous intelligent agents and social bots can generate and spread individualized fake news at an unprecedented scale. The rise of Generative Adversarial Networks (GANs) allows for the creation of "Deepfakes"—highly realistic but entirely fabricated videos and audio recordings. In a political crisis, these tools can be used to manipulate public opinion or incite violence, making reality itself indistinguishable from artificial creation.

- **Autonomous Criminal Profiling:**

  Deep learning software is increasingly used for facial recognition and criminal profiling in public surveillance networks. This risks turning the "presumption of innocence" upside down, as algorithms categorize individuals as potential threats based on biased data or correlation recognition uncontrollable by humans.

- **Algorithmic Governance (The Social Credit Model):**

  Rating systems for citizens, such as the "Citizen Score Card" tested in some jurisdictions, represent the ultimate peace-time application of autonomous profiling. These systems weigh the "value" of an individual citizen to the state, potentially informing decisions on who receives medical treatment or whose freedoms are restricted based on utilitarian calculations set by an algorithm.

VIII. Proposed Solutions and Frameworks

The ultimate objective of the DISEC committee is to draft a comprehensive, actionable, and widely supported resolution that addresses the proliferation and regulation of Lethal Autonomous Weapons Systems (LAWS). Because the international community is deeply divided, there is no single "correct" solution. Instead, delegates must negotiate and amalgamate various frameworks to achieve a consensus.

a. A Legally Binding Instrument (LBI): The Preemptive Ban

Championed by the Ban Coalition (e.g., Austria, New Zealand, Mexico) and civil society groups like the *Campaign to Stop Killer Robots*, this framework demands the creation of a new, legally binding international treaty (similar to the Ottawa Treaty on Landmines or the Chemical Weapons Convention).

Key Components of an LBI:

- **Categorical Prohibition of "Out-of-the-Loop" Systems:**

The treaty would establish a red line, explicitly banning any weapon system that selects and engages human targets without Meaningful Human Control (MHC).

- **Ban on "Black Box" Algorithms:**

Any AI system whose decision-making process cannot be fully understood, predicted, or traced by its human operators would be deemed inherently illegal under International Humanitarian Law.

- **Positive Obligations for MHC:**

Rather than just listing bans, the treaty would legally mandate strict "positive obligations." This means international law would dictate the minimum level of human-machine interaction required (e.g., a human must always have the visual feed and a 10-second window to abort an attack).

- **The Diplomatic Challenge:**

Verification. How do UN inspectors verify a software code? Unlike nuclear centrifuges, which are massive physical facilities, an illegal AI algorithm can be hidden on a standard hard drive.

b. The "Two-Tier" Regulatory Approach

Favored by mediator states (e.g., France, Germany) and increasingly popular in recent GGE sessions, this framework seeks a middle ground between a total ban and total deregulation. It acknowledges that autonomy has legitimate defensive uses but poses unacceptable risks in offensive, anti-personnel contexts.

Key Components of the Two-Tier Approach:

- **Tier 1 (Strict Prohibitions):**

Systems that are completely unpredictable, systems that use biometrics/facial recognition to hunt specific ethnic groups, or systems that target civilians indiscriminately would be absolutely banned.

- **Tier 2 (Strict Regulations):**

Systems that feature autonomy but *can* be used in compliance with IHL would be heavily regulated through temporal and spatial limitations.

- **Operational Constraints (The "Geofence"):**

To legally deploy a Tier 2 autonomous weapon, a commander must restrict its operation to a specific, narrow geographic area (a "geofence") and a strict time limit (e.g., the drone may only hunt autonomously for 45 minutes within a 5-square-kilometer uninhabited combat zone). If the drone breaches these limits, it must automatically self-destruct or return to base.

c. Political Declarations and Codes of Conduct

Strongly advocated by the major "Developers" (e.g., United States, Russia, Israel), this framework rejects the creation of a new, legally binding treaty. These states argue that a new

treaty would take decades to negotiate, would likely be ignored by rogue states, and would stifle beneficial AI research.

Key Components of a Code of Conduct:

- **Voluntary Guidelines:**

Instead of a treaty, the UN would adopt a non-binding "Code of Conduct" based on the existing 11 Guiding Principles of the GGE. States would voluntarily pledge to adhere to these best practices.

- **National Article 36 Reviews:**

The framework would heavily emphasize Article 36 of Additional Protocol I to the Geneva Conventions. It would mandate that states establish highly rigorous, transparent national review boards to test their own AI weapons for IHL compliance *before* deploying them, keeping the regulatory power within the sovereign state rather than delegating it to the UN.

- **Standardized Testing Environments:**

Establishing international, standardized virtual testing environments (sandboxes) where militaries can safely test their algorithms against simulated IHL dilemmas to prove their reliability.

d. The Shared Human-Machine Decision-Making Model

A growing number of experts suggest that the international community should focus not on a total ban of the technology, but on developing rules for human participation in its functioning. This leads to the "Human-Machine Shared Decision-Making" model, which seeks to reconcile military efficiency with IHL compliance.

In this framework, the machine is not left to independently assess qualitative and multi-context concepts like "proportionality" or "humanity," which current AI cannot comprehend. Instead, the human operator determines the rigid boundaries—the spatial and temporal "fences"—within which the autonomous system is permitted to act. For example, a human makes the high-level decision to apply the weapon in a specific area after evaluating potential civilian casualties. The autonomous system then operates strictly within those established boundaries. This approach ensures that the "digital brain" handles the speed and accuracy of the engagement while the human retains the moral and legal responsibility for the core principles of humanitarian law.

e. The Three-Phase Operational Test and Evaluation (OT&E) Standard

As a practical regulatory framework, the DISEC committee could mandate a universal standardization for the testing of autonomous systems. To ensure that LAWS operate predictably and allow commanders to exercise appropriate levels of human judgment, systems must formally complete a rigorous Operational Test and Evaluation (OT&E) process before being released to the field. A proposed internationally recognized OT&E standard would encompass three mandatory phases:

1. **The Requirements and Development Phase:**

Designing the system's architecture to include robust safeties, anti-tamper devices, and strict boundaries for target engagement.

2. **Test and Evaluation:**

Subjecting the algorithm to rigorous testing in complex, contested environments (including heavy electronic warfare and anti-access/area denial scenarios) to verify that it will terminate engagements or seek operator input if it cannot fulfill its parameters.

3. Transition to Operational Deployment:

Ensuring that the technology is affordable, realistic, and securely integrated into the command-and-control network before final deployment.

IX. Questions to be Addressed

To guide the debate and assist in the drafting of Position Papers and Working Papers, delegates must consider the following critical questions. A comprehensive resolution should provide actionable answers to these dilemmas:

1. **Defining the Threshold:**

How should the international community legally and technically define "Meaningful Human Control" (MHC) or "Appropriate Levels of Human Judgment"? At what specific point does human supervision become insufficient?

2. **The Accountability Gap:**

In the event that a fully autonomous weapon commits a violation of International Humanitarian Law, what specific legal mechanisms should be established to assign criminal liability? Should the doctrine of Command Responsibility be updated for the AI era?

3. **Verification and Compliance:**

If a legally binding ban or regulation is adopted, how can the United Nations practically verify compliance? How do international inspectors monitor software algorithms and machine-learning training data without violating state sovereignty and military secrecy?

4. **Asymmetric Warfare and the Global South:**

How can the United Nations address the legitimate security concerns of developing nations who fear that LAWS will widen the technological gap and be used as tools of neo-colonial coercion?

5. **Cybersecurity and Hijacking:**

What universal cybersecurity standards must be mandated for autonomous systems to prevent mass hacking, GPS spoofing, or the unauthorized acquisition of swarm technology by non-state actors?

6. **The "Flash War" Threat:**

What specific Confidence-Building Measures (CBMs) and crisis-communication protocols should be established to prevent autonomous defense systems from triggering an accidental, algorithmic escalation?

7. **The Ethical Red Line:**

Is the "Digital Dehumanization" argument legally sound? Does delegating the decision to kill to a machine inherently violate the Martens Clause and the fundamental principles of human dignity, regardless of the weapon's accuracy?

X. Bibliography

Barbé, E., & Badell, D. (2020). The European Union and lethal autonomous weapons systems: United in diversity?. *Norm Research in International Relations*, 141-158.

Cernat, R. (2021). Lethal autonomous weapon systems – emerging and potentially disruptive technology. *Romanian Military Thinking*, (4), 156-175.

Christie, E. H., Ertan, A., Adomaitis, L., & Klaus, M. (2023). Regulating lethal autonomous weapon systems: Exploring the challenges of explainability and traceability. *AI and Ethics*, *4*, 229-245. doi:10.1007/s43681-023-00261-0

Dremliuga, R. (2020). General legal limits of the application of the lethal autonomous weapons systems within the purview of international humanitarian law. *Journal of Politics and Law, 13*(2), 115-121. doi:10.5539/jpl.v13n2p115

International Committee of the Red Cross (ICRC). (1977). *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I)*. Retrieved from https://ihl-databases.icrc.org/en/ihl-treaties/api-1977

International Committee of the Red Cross (ICRC). (2014). *Autonomous weapon systems: Technical, military, legal and humanitarian aspects*. Retrieved from https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014

Lewis, D. A., Blum, G., & Modirzadeh, N. K. (2016). *War-algorithm accountability*. Harvard Law School Program on International Law and Armed Conflict (PILAC). Retrieved from https://pilac.law.harvard.edu/war-algorithm-accountability-report

Longpre, S., Storm, M., & Shah, R. (2021). Lethal autonomous weapons systems & artificial intelligence: Trends, challenges, and policies. *MIT Science Policy Review.*

Meier, M. W. (2016). Lethal autonomous weapons systems (LAWS): Conducting a comprehensive weapons review. *Temple International & Comparative Law Journal*, *30*(1), 119-132.

Pedron, S. M., & da Cruz, J. A. (2020). The future of wars: Artificial Intelligence (AI) and Lethal Autonomous Weapon Systems (LAWS). *International Journal of Security Studies*, *2*(1), Article 2.

Righetti, L., Pham, Q. C., Madhavan, R., & Chatila, R. (2018). Lethal autonomous weapon systems. *IEEE Robotics & Automation Magazine*, *25*(1), 123-126.

Surber, R. (2018). *Artificial Intelligence: Autonomous Technology (AT), Lethal Autonomous Weapons Systems (LAWS) and Peace Time Threats*. Zurich, Switzerland: ICT4Peace Foundation.

United Nations Convention on Certain Conventional Weapons (CCW). (2019). *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems* (CCW/GGE.1/2019/3). Retrieved from https://documents.unoda.org/wp-content/uploads/2020/09/CCW_GGE.1_2019_3_E.pdf

United Nations General Assembly. (2023). *Resolution 78/241: Lethal autonomous weapons systems* (A/RES/78/241). Retrieved from https://undocs.org/en/A/RES/78/241

Watts, T. F. A., & Bode, I. (2023). Machine guardians: The Terminator, AI narratives and US regulatory discourse on lethal autonomous weapons systems. *Cooperation and Conflict*, *59*(1), 107-128. doi:10.1177/00108367231198155