COS30019 – Introduction to AI

Assignment 1 : Research into AI Ethics & Responsible AI

Name: Thenura Dulnath Kuruppuarachchi

ID: 103512993

Class: Tuesday 2.30pm to 4.30pm

Contents

Executive Summary

The integration of Artificial Intelligence (AI) into military operations has led to the development of Lethal Autonomous Weapons Systems (LAWS), capable of selecting and engaging targets without human intervention. This advancement raises critical ethical, legal, and technological questions regarding the delegation of life-and-death decisions to machines. Key issues include the balance between autonomy and accountability, adherence to international humanitarian law, ethical implications, technological reliability, and the adequacy of existing policy and regulatory frameworks.

This report examines these issues by analyzing academic literature, policy papers, and case studies to assess the viability and morality of permitting AI systems to kill. The findings reveal a complex landscape where technological advancements have outpaced regulatory measures, leading to concerns about accountability, ethical boundaries, and compliance with international laws. The report concludes that while AI offers significant strategic advantages, the deployment of LAWS without stringent oversight and clear ethical guidelines poses substantial risks. Recommendations emphasize the necessity for international cooperation, the establishment of comprehensive regulatory frameworks, and the maintenance of meaningful human control over lethal decision-making processes.

1. Introduction

The integration of Artificial Intelligence (AI) into military operations has revolutionized modern warfare, bringing capabilities that once belonged to the realm of science fiction. Among these, Lethal Autonomous Weapons Systems (LAWS) have become one of the most controversial and urgent developments in defense technology. These systems can independently select and engage targets once activated, potentially changing the very nature of warfare.

Recent global conflicts, rapid advances in AI decision-making, and the absence of cohesive international regulation have made the debate over LAWS increasingly urgent. The Russian-Ukrainian conflict, for example, has demonstrated the growing use of autonomous drones, fueling international concern about the future of war and the need for strong legal frameworks. The discussion around LAWS is not just theoreticality unfolding in real-time, with many governments investing in or experimenting with such technologies.

Historically, weapon innovation—from longbows to nuclear arms—has consistently challenged the ethical and legal norms of warfare. LAWS represent the latest, and perhaps most complex, iteration of this trend, as they remove human agency from decisions about life and death. The implications are enormous—not just for military strategy, but for international law, human rights, and the ethical underpinnings of conflict.

The central question explored in this report is: **Should AI systems be allowed to kill?** This question is examined through five key issues:

- **Autonomy vs. Accountability**: Who is responsible when AI makes lethal mistakes?
- **Compliance with International Humanitarian Law (IHL)**: Can AI reliably uphold the principles of distinction and proportionality?
- **Ethical Considerations**: Is it morally acceptable to delegate lethal decisions to machines?
- **Technological Limitations**: Can AI function reliably in chaotic, real-world combat?
- **Policy and Regulation**: Do current frameworks sufficiently address the risks of LAWS?

This report uses a qualitative research methodology, drawing on academic literature, policy proposals, legal analyses, case studies, and expert commentary. The goal is to provide a well-rounded, critically engaged exploration of whether LAWS can—or should—be ethically and legally integrated into modern warfare.

While the scope of this report centers on LAWS in formal military settings, it also touches on adjacent implications such as future use by law enforcement, state surveillance, and non-state actors. The report deliberately focuses on international perspectives to reflect the global implications of deploying autonomous weapons systems in cross-border conflicts.

Methodology

This report employs a qualitative research approach, analyzing academic literature, policy papers, case studies, and expert opinions. Primary sources include "Ethical Considerations in Artificial Intelligence Courses" by Burton et al. (2017) and Australia's government proposal paper on "Introducing Mandatory Guardrails for AI in High-Risk Settings" (2024). Additionally, insights from international organizations, legal analyses, and recent developments in AI military applications are incorporated to provide a comprehensive understanding of the topic.

Scope

The scope of this research encompasses the ethical, legal, and technological dimensions of LAWS, with a focus on international perspectives and implications. While the report acknowledges the broader context of AI in military use, it concentrates specifically on autonomous systems capable of lethal action without human intervention.

3. Findings

The deployment of Lethal Autonomous Weapons Systems (LAWS) introduces a complex array of ethical, legal, technological, and policy challenges. This section delves into these critical issues, analyzing perspectives from various stakeholders and synthesizing information from reliable sources to support the findings and conclusions.

3.1 Autonomy vs. Accountability

The introduction of autonomous systems capable of making life-and-death decisions raises significant concerns about accountability. Traditional military operations involve human decision-makers who can be held responsible for their actions. In contrast, autonomous systems operate based on algorithms and pre-programmed parameters, complicating the assignment of liability when such systems cause unintended harm or violate rules of engagement. Ethicist Robert Sparrow highlights this issue, noting that the fundamental condition of international humanitarian law requires that some person must be held responsible for civilian deaths. Army University Press

The lack of a clear accountability framework for LAWS undermines the principles of justice and responsibility in warfare. Determining who is culpable—the programmer, the commander, or the manufacturer—becomes a contentious issue, potentially leading to a responsibility gap where no party is held accountable for wrongful acts committed by autonomous systems.

In addition to legal responsibility, **military ethics and command structures** face unprecedented disruption from LAWS. Traditional military chains of command rely on human judgment and accountability at every level, from the soldier to the general. Introducing autonomy breaks that structure, making it unclear how to enforce codes of conduct, military tribunals, or court-martials for wrongful acts. Some argue that "human in the loop" oversight may preserve accountability, but others question whether this is functionally possible when split-second decisions are delegated to algorithms. Without a clear accountability chain, justice becomes more difficult to achieve, especially for victims of wrongful killings.

3.2 Compliance with International Humanitarian Law (IHL)

International Humanitarian Law governs the conduct of armed conflict, emphasizing principles such as distinction, proportionality, and necessity. The use of LAWS challenges these principles, particularly the requirement to distinguish between combatants and civilians. Ensuring that autonomous systems can make such distinctions reliably is a significant concern, as failure to do so could lead to unlawful killings and increased civilian casualties. The International Committee of the Red Cross (ICRC) has been actively involved in discussions surrounding the ethical and legal implications of autonomous weapon systems, emphasizing the need for compliance with IHL. ICRC Blogs

Moreover, the principle of proportionality, which mandates that the harm caused to civilians must not be excessive in relation to the anticipated military advantage, is difficult to program into autonomous systems. The subjective nature of proportionality assessments poses a challenge for LAWS, as they may lack the nuanced judgment required to make such determinations.Wikipedia+1Wikipedia, la enciclopedia libre+1

Further complicating the issue is the **interpretive flexibility of IHL principles**. For instance, the principle of distinction assumes a capacity for contextual awareness—something even human soldiers struggle with. While some developers argue that AI can outperform humans in target identification due to sensor accuracy, others stress that machines lack the judgment needed for ambiguous situations, such as distinguishing between armed civilians and insurgents. The challenge isn't just programming legal rules

into AI—it's about replicating human situational awareness, empathy, and moral discretion, which are essential in adhering to IHL.

3.3 Ethical Implications

Delegating the decision to take human life to machines raises profound ethical questions. Critics argue that allowing AI to make lethal decisions undermines human dignity and moral agency. The concept of meaningful human control is central to this debate, emphasizing that critical decisions in warfare should not be outsourced to machines lacking consciousness and moral reasoning. The Campaign to Stop Killer Robots highlights that autonomy in weapons systems is a profoundly human problem, as it changes the relationship between people and technology by handing over life-and-death decision-making to machines, thereby dehumanizing the act of killing.

Proponents of LAWS argue that AI could potentially reduce battlefield casualties by making faster, more calculated decisions. However, critics counter that **removing humans from the kill chain** risks a broader devaluation of life. Philosophers and ethics warn of a moral "desensitization," where governments may become more willing to engage in conflict if the political cost in human soldiers is minimized. There's also concern about desensitizing society itself, a world in which machines decide who lives and dies could alter fundamental norms of justice, responsibility, and human rights across generations.

Furthermore, the potential for LAWS to be used in ways that violate ethical norms, such as targeted assassinations or oppressive control over populations, raises concerns about their impact on human rights and societal values.

3.4 Technological Limitations and Predictability

The reliability and predictability of AI systems in complex and dynamic combat environments are critical concerns. AI systems may exhibit unpredictable behaviors due to unforeseen interactions with the environment or adversarial manipulation. Ensuring that LAWS operate as intended and can adapt appropriately to unforeseen circumstances remains a significant technological challenge. The IEEE highlights that advancements in AI have raised questions at national and international levels about what is technologically possible and what is legally and morally acceptable regarding autonomy and AI in weapon systems.

The problem isn't just unpredictability, it's the **opacity of AI decision-making**. Many modern AI systems operate as "black boxes," making decisions that are not fully explainable, even to their creators. In the fog of war, the inability to audit or understand how a decision was made can be catastrophic. Imagine an autonomous system misidentifying a

refugee convoy as a military target—without transparency, correcting or learning from such failures become almost impossible. Moreover, adversarial attacks, where opponents feed misleading data into the system to provoke a misfire, are an emerging cyberwarfare threat specific to AI weapons.

Additionally, the potential for technical malfunctions or hacking poses risks of unintended engagements, leading to unintended casualties and escalation of conflicts.

3.5 Global Policy and Regulation

The rapid development of LAWS has outpaced the establishment of comprehensive regulatory frameworks. International efforts to regulate or ban LAWS have faced challenges due to differing national interests and the strategic advantages that such systems may confer. The United Nations has been debating the legal, social, ethical, and security implications of LAWS for over a decade, with discussions highlighting the complexities and differences within and beyond the UN Group of Governmental Experts on LAWS.

In December 2023, the United Nations General Assembly adopted a resolution to support international discussion regarding concerns about LAWS, reflecting the global apprehension surrounding these systems. However, the lack of binding agreements raises concerns about an arms race and the proliferation of autonomous weapons without adequate oversight.

Nationally, some countries have begun to address these issues. For instance, Australia's government proposal paper on "Introducing Mandatory Guardrails for AI in High-Risk Settings" (2024) outlines measures to regulate AI applications, including those in military contexts. However, the effectiveness of such national policies is limited without international consensus and cooperation.

Beyond national interests, **technological asymmetry between nations** also hinders global consensus. Wealthier countries may have more advanced capabilities and thus be reluctant to accept binding constraints. Developing nations, on the other hand, fear being left behind or made vulnerable. This dynamic mirrors historical patterns in arms control debates—such as nuclear non-proliferation—and highlights the need for equitable, enforceable frameworks. Think tanks and NGOs continue to propose models for international cooperation, including LAWS-specific treaties modeled after the Chemical Weapons Convention, which could serve as a precedent for effective governance.

In conclusion, the deployment of LAWS presents multifaceted challenges that encompass ethical dilemmas, legal uncertainties, technological limitations, and regulatory gaps. Addressing these issues requires a concerted effort from the international community to

establish comprehensive frameworks that ensure accountability, uphold ethical standards, and regulate the development and use of autonomous weapons systems.

3.6 Public Perception and Societal Impact

Public perception plays a critical role in shaping policy and ethical discourse around LAWS. In democratic societies, public opposition to autonomous weapons can influence defense spending, regulatory decisions, and political accountability. Surveys conducted in multiple countries, including the U.S., UK, and Germany, show most civilians oppose fully autonomous weapons, citing concerns over ethics, misuse, and trust in technology.

Pop culture and media also shape how society understands AI in war. Films like *Terminator* and *Eye in the Sky* dramatize the consequences of AI decision-making, contributing to fear and skepticism. While fictional, these portrayals tap into very real anxieties about losing control over life-and-death decisions. Advocacy groups leverage this public concern to push for bans or moratoriums on LAWS development.

Conversely, in more authoritarian regimes, the lack of transparency can limit public engagement on these issues. As such, the societal impact is deeply uneven globally. Nonetheless, global norms often emerge from public pressure, making public perception a potentially powerful force in future international regulation.

4. Conclusion and Recommendations

Conclusion

The advent of Lethal Autonomous Weapons Systems (LAWS) has introduced profound ethical, legal, technological, and policy challenges. The delegation of life-and-death decisions to machines raises significant concerns regarding accountability, as traditional frameworks struggle to assign responsibility for autonomous actions. Furthermore, ensuring compliance with International Humanitarian Law (IHL) becomes complex when machines must distinguish between combatants and civilians and assess proportionality in attacks. Ethically, the use of LAWS questions the preservation of human dignity and moral agency in warfare. Technologically, the unpredictability and potential unreliability of AI systems in dynamic combat environments pose risks of unintended engagements and escalation. On the policy front, the rapid development of LAWS has outpaced the establishment of comprehensive regulatory frameworks, leading to international debates and resolutions aimed at addressing these emerging challenges.

Recommendations

1. **Establish Clear Accountability Frameworks:** Develop international agreements that delineate responsibility for the actions of autonomous systems, ensuring that accountability remains with human operators and commanders.
2. **Ensure Compliance with International Humanitarian Law:** Mandate that any deployment of LAWS includes mechanisms to guarantee adherence to IHL principles, particularly distinction and proportionality, with rigorous testing and validation protocols.
3. **Maintain Meaningful Human Control:** Implement policies that require human oversight in critical functions of weapon systems, especially decisions involving the use of lethal force, to preserve ethical standards and moral responsibility.
4. **Invest in Robust Technological Safeguards:** Prioritize research into the reliability and predictability of AI systems, developing fail-safes and override mechanisms to prevent unintended actions and ensure operational safety.
5. **Develop Comprehensive Regulatory Frameworks:** Encourage international collaboration to create binding agreements that regulate the development, proliferation, and use of LAWS, preventing an arms race and ensuring global security.
6. **Promote Transparency and Public Discourse:** Foster open discussions among governments, industry leaders, ethicists, and the public to address concerns, share information, and build consensus on the acceptable use of autonomous weapons.

By implementing these recommendations, the international community can work towards mitigating the risks associated with LAWS, ensuring that the integration of AI into military operations aligns with ethical principles, legal standards, and the preservation of human dignity.

5. Acknowledgements/Resources

I would like to express my sincere gratitude to the following individuals and organizations for their invaluable contributions to this report:

- **Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T.**: Their article, "Ethical Considerations in Artificial Intelligence Courses," provided

foundational insights into the ethical frameworks surrounding AI, which were instrumental in shaping the ethical analysis presented in this report.

- **Australian Government Department of Industry, Science, Energy and Resources**: The proposal paper on "Introducing Mandatory Guardrails for AI in High-Risk Settings" offered a comprehensive overview of regulatory approaches to AI, informing the policy discussions within this report.
- **International Committee of the Red Cross (ICRC)**: Their publications on autonomous weapon systems and International Humanitarian Law were crucial in understanding the legal implications of LAWS, contributing significantly to the legal analysis section.
- **Campaign to Stop Killer Robots**: The resources and advocacy materials provided by this organization offered diverse perspectives on the ethical and societal impacts of autonomous weapons, enriching the ethical considerations discussed herein.
- **IEEE Standards Association**: Their reports on the ethical and technical challenges of autonomous weapons systems provided a technological perspective that was essential for evaluating the feasibility and risks associated with LAWS.
- **Various Academic Journals and Publications**: A multitude of peer-reviewed articles and case studies were consulted to ensure a well-rounded and evidence-based analysis throughout this report.

Additionally, I acknowledge the use of AI tools for research assistance, which facilitated the efficient gathering and organization of information pertinent to this study.

6. References

- Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T. (2017). Ethical Considerations in Artificial Intelligence Courses. *AI Magazine*, 38(2), 22-34. Link
- Australian Government Department of Industry, Science, Energy and Resources. (2024). Introducing Mandatory Guardrails for AI in High-Risk Settings: Proposals Paper. Link
- International Committee of the Red Cross. (2014). Autonomous Weapon Systems under International Humanitarian Law. Link
- Campaign to Stop Killer Robots. (n.d.). Facts about Autonomous Weapons. Link

- IEEE Standards Association. (2017). Ethical and Technical Challenges in the Development of Autonomous Weapon Systems. Link
- United Nations General Assembly. (2023). Resolution on Lethal Autonomous Weapons Systems. Link
- Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy*, 24(1), 62-77.
- Docherty, B. (2012). Losing Humanity: The Case against Killer Robots. *Human Rights Watch*.
- Scharre, P. (2018). Army of None: Autonomous Weapons and the Future of War. *W.W. Norton & Company*.