Information & Security

M. Lachova
vol. 55, no. 2 (2024): 213-221
https://doi.org/10.11610/isij.5551

Published since 1998 ISSN 0861-5160 (print), ISSN 1314-2119 (online)

Research Article

Data, Privacy and Human-Centered AI in Defense and Security Systems: Legal and Ethical Considerations

Maria Lachova

Legal expert, Law and Internet Foundation, Sofia, Bulgaria
https://www.netlaw.bg/en

ABSTRACT:

The increasing integration of data-driven technologies and Artificial Intelligence (AI) into defense and security systems offers both transformative opportunities and significant challenges. This paper examines how data management intersects with the General Data Protection Regulation, privacy concerns, and human-centered AI in the context of defense and security systems. AI-driven systems utilize large volumes of data for various purposes, including decision-making, surveillance, threat detection, and autonomous operations. While these technologies can enhance effectiveness and provide strategic advantages, they also raise critical issues related to data security, privacy violations, and the potential for biased or unethical outcomes. A human-centered approach to AI emphasizes the importance of transparency, accountability, and ethical alignment in these systems. This approach aims to ensure that technologies operate within legal and moral boundaries, minimizing the risk of unintended harm.

This article explores the delicate balance between technological advancements and ethical considerations, proposing frameworks to protect privacy and uphold human rights while maximizing the operational benefits of AI in defense and security.

ARTICLE INFO:

RECEIVED: 26 SEP 2024
REVISED: 25 OCT 2024
ONLINE: 01 NOV 2024

KEYWORDS:

artificial intelligence, AI, defense, privacy, information and communication technology, ICT, ethics, norms

Tel.: +359888766520 E-mail: maria.lachova@netlaw.bg

Introduction

The advancement of Artificial Intelligence (AI) and data-centric technologies is reshaping the domain of defense and security systems. As countries and defense entities progressively depend on AI for various functions such as decision-making, surveillance, threat identification, and autonomous operations, the significance of data becomes pivotal in attaining both strategic and operational advantages. The extensive gathering, processing, and analysis of large datasets facilitate more informed decision-making processes and offer real-time insights that can greatly enhance security responses.

However, the reliance on data and AI in defense introduces significant ethical, legal, and operational challenges. The sensitive nature of military data necessitates robust security protocols, ensuring data integrity and preventing unauthorized access, especially in an era of sophisticated cyber threats. Moreover, the growing use of AI for surveillance and predictive analysis raises pressing privacy concerns, as it has the potential to infringe on civil liberties by enabling mass data collection and intrusive monitoring.

To address these challenges, the development and deployment of AI in defense must adopt a human-centered approach, focusing on ethical principles such as transparency, fairness, accountability, and respect for human rights.[1] Human-centered AI seeks to integrate technology in ways that prioritize individuals' needs, values, and well-being, ensuring that defense and security systems operate within acceptable moral and legal frameworks.

This article explores the challenges of applying the concept of human-centered AI in defense and security systems. It examines the benefits and risks associated with data-driven AI technologies in military contexts and the need for robust legal and ethical frameworks to guide the responsible and lawful use of AI. By addressing both legal and ethical considerations, it aims to contribute to the ongoing dialogue on how AI can enhance defense capabilities while ensuring compliance with law regulations, safeguarding human rights, and maintaining ethical integrity.

Ethical Considerations in AI-Driven Defense Systems

Ethical considerations must serve as a framework for the deployment of AI in defense systems to ensure the protection of human rights and the integrity of military operations. Central to these ethical considerations is the imperative to minimize harm to civilians and non-combatants. This is also introduced in the subject matter of the AI Act, [2] which states that AI systems should be designed with a priority on protecting fundamental rights, health, and safety. In the context of military actions, this requirement emphasizes that the development and implementation of AI technologies must align with standards that prevent unnecessary suffering and uphold humanitarian values. Given the complexity of modern warfare, it is imperative that AI technologies are not only efficient in achieving military objectives but are also fundamentally ethical in their operations, ensuring that the impact on civilian populations is carefully considered and mitigated.

Furthermore, the potential for AI technologies to be weaponized raises profound ethical concerns regarding the use of Lethal Autonomous Weapon Systems (LAWS).[3] As these systems operate with a degree of autonomy, the ability to make life-and-death decisions without direct human intervention poses serious moral dilemmas. It is essential that ethical guidelines be established to govern the use of such technologies, ensuring that there are stringent checks in place to prevent misuse or accidental escalation of conflicts. The ethical implications extend beyond the battlefield; they encompass broader societal concerns, such as how military use of AI can affect public perceptions of safety and security, potentially normalizing the use of technology in contexts that may undermine civil liberties.

Transparency in AI[4] is another consideration that cannot be overlooked in the deployment of AI systems within defense contexts. In that respect, the AI algorithms must be designed to be transparent and understandable, allowing operators to discern the rationale behind AI-generated decisions. This transparency not only fosters trust in technology but also holds the systems accountable to oversight mechanisms. When military operators are equipped with clear insights into how AI reaches its conclusions, they can make informed decisions, especially in high-stakes scenarios. Without this level of transparency, there is a risk of AI systems making decisions that are not aligned with ethical principles or legal standards.

In addition to transparency, ensuring fairness in AI systems is paramount. Ethical AI systems must be designed to actively minimize bias and ensure equitable treatment of all individuals, regardless of their background or status. Bias in AI algorithms can lead to disproportionate targeting of specific populations or individuals, which raises serious ethical concerns about discrimination and injustice. That is why developers and military organizations must implement rigorous testing protocols to identify and mitigate any biases that may inadvertently influence AI decisions. This responsibility extends to ensuring that training data used for AI systems is diverse and representative, thereby reducing the likelihood of perpetuating existing societal inequities.

A human-centered approach to AI in defense systems necessitates robust accountability and human oversight. Establishing mechanisms that ensure human operators retain control over critical decisions is crucial. This principle is particularly vital in life-or-death situations, where reliance on automated systems without human intervention could lead to catastrophic outcomes. By embedding human judgment into the operational framework of AI systems, military organizations can better navigate ethical dilemmas and maintain the moral integrity of their actions. Clear accountability mechanisms, including regular audits and oversight committees, are essential for ensuring that the deployment of AI technologies adheres to established legal and ethical standards, thereby safeguarding human rights in military operations.

To support these ethical considerations, several guidelines provided by the European Commission for trustworthy AI[5] should be mentioned. First, AI systems should be designed with the principle of robustness and safety in mind, ensuring that they operate reliably under various conditions and can withstand adversarial attacks. Second, the principle of privacy by design should be applied, meaning that data protection measures are integrated into AI systems from the outset, safeguarding individuals' personal information against misuse. Third, fairness and inclusivity should be prioritized in both the design and deployment phases of AI technologies, ensuring that diverse perspectives are represented and that algorithms are tested for equitable outcomes across different demographic groups.

Moreover, there should be a commitment to accountability and traceability, meaning that every AI system must be designed in a way that allows for the identification of the data and decision-making processes behind its outputs. This can facilitate transparency and help establish trust among users and stakeholders. Finally, a culture of continuous learning and adaptation must be fostered, where organizations actively seek feedback on AI systems and are open to revising practices in light of new ethical insights and technological advancements.

By following the ethical guidelines for trustworthy AI, defense organizations can ensure that their AI technology implementation meets moral and legal standards. This commitment fosters a more responsible and ethical approach to military operations, aligning technological advancements with essential ethical principles and compliance obligations.

Legal Considerations in AI-Driven Defense Systems

AI-driven defense systems introduce a range of complex legal and ethical challenges as they significantly alter the landscape of military operations, law enforcement, and national security strategies. As AI technology becomes more integrated into defense, the legal frameworks governing its use must evolve to address these challenges, ensuring accountability, compliance with international laws, transparency, data privacy, and the protection of human rights. In particular, the *EU Artificial Intelligence Act (AI Act)* [6] and the *General Data Protection Regulation (GDPR)* [7] are central to shaping the governance of AI technologies in defense, balancing innovation with legal safeguards to prevent abuse and misuse.

One of the most pressing legal challenges is accountability when AI systems make autonomous decisions that result in harm or legal violations. Traditional legal frameworks assume human decision-makers, but AI systems—especially those involved in autonomous weapons or surveillance—can act independently and beyond direct human control. This raises a critical question: who is responsible when an AI-driven system causes harm? Is it the developer, the military commander, the manufacturer, or the state that deployed the system? Establishing a clear chain of responsibility is essential, especially when AI systems operate without direct human oversight or when their decision-making processes are too complex or opaque for human operators to understand fully. The AI Act seeks to address this by enforcing transparency and accountability, but the challenge remains profound. While the Act provides important guidelines for high-risk AI systems,[8] it stops short of offering comprehensive solutions for accountability in defense contexts where life-and-death decisions are involved. This is a key area where further legal development is, in my opinion, urgently needed, especially as AI becomes more advanced and autonomous.

Beyond accountability, AI-driven defense systems must also comply with International Humanitarian Law (IHL),[9] which governs the conduct of warfare. The key principles of IHL—such as distinction (between combatants and civilians), proportionality (ensuring that harm to civilians is proportional to military objectives), and necessity (the use of force only when absolutely required)—are fundamental to the legal and ethical use of force in conflicts. However, ensuring that AI systems are designed to adhere to these principles presents a unique set of legal and technical challenges. Autonomous AI systems [10] may not have the nuanced judgment required to reliably comply with IHL, particularly in complex and rapidly changing combat environments. In that respect, it is important to mention that the AI Act places defense-related AI systems under its high-risk category, mandating rigorous oversight, but there is still significant uncertainty about whether AI can ever truly be programmed to make ethical decisions in war. This raises profound moral questions about whether the use of AI in lethal operations can ever be truly justified under current legal and ethical standards.

Another critical legal concern is transparency. Many AI systems, particularly those that rely on machine learning and deep learning, are often described as "black boxes" [11] because their decision-making processes are difficult to interpret or explain. This lack of transparency is problematic in any context but becomes especially dangerous in defense, where human lives are at stake. Without clear insight into how AI systems arrive at their decisions, it is difficult to ensure accountability, prevent bias, or investigate incidents when something goes wrong. The GDPR, which emphasizes transparency and individual rights in data processing, requires organizations to ensure that people understand how their data is being used and processed. However, the GDPR's principles of transparency are difficult to apply to AI-driven military systems, where operational secrecy and national security concerns often limit transparency. The AI Act, in turn, seeks to enforce explainability for high-risk AI systems, yet the reality remains that many AI systems used in defense are inherently complex and not easily interpretable. This is a significant legal gap that must be addressed—ensuring that military AI systems, no matter how advanced, remain accountable and open to oversight.

Data privacy is another pressing legal issue when it comes to AI-driven defense systems. These systems often rely on vast amounts of data to function—whether for surveillance, intelligence gathering, or predictive analysis—raising serious concerns about mass surveillance and violations of individual privacy. AI systems often collect data without consent, infringing on personal privacy rights, which could lead to significant human rights violations. The GDPR plays a crucial role in regulating data protection across the EU, requiring strict rules for data collection, processing, and storage. However, even under the GDPR, balancing national security with privacy rights is a delicate and challenging task.

While the GDPR ensures that personal data is protected, the question remains whether military AI systems can or should be held to the same privacy standards, especially in scenarios where national security is prioritized. In the case of defense systems that utilize facial recognition or predictive policing, this challenge is even more pronounced, as these technologies are prone to overreach and can lead to unlawful surveillance and profiling if not carefully regulated.

Another important consideration is the dual-use [12] nature of AI technologies, where systems designed for civilian applications can easily be repurposed for military use. This flexibility, while beneficial in many contexts, poses significant challenges for arms control and export regulation, as existing frameworks often fail to address the unique risks created by AI. Traditional armaments can usually be categorized and controlled, but the adaptability of AI technology means that tools initially developed for benign purposes can swiftly transform into powerful military assets.

This reality raises urgent questions about regulation and oversight regarding the potential misuse of AI technologies. As discussed in a Forbes article, AI systems originally designed to enhance healthcare efficiency or improve transportation logistics can be modified for military applications, such as surveillance, autonomous weaponry, and information warfare. This dual-use aspect complicates monitoring and governance efforts, making it difficult to establish clear guidelines for responsible usage.

While some existing regulations attempt to mitigate harm by providing guidelines for AI applications, the international legal community must act much more swiftly to revise treaties and agreements to address the nuanced risks associated with AI. In that respect, the *AI Liability Directive*,[13] proposed by the European Commission, is a critical step toward establishing accountability for damage caused by AI systems. This directive aims to clarify liability issues, ensuring that victims can seek redress when harmed by AI technologies, particularly in high-risk sectors like defense.

The AI Liability Directive is particularly relevant in the context of military applications, where the potential for misuse and unintended consequences is significant. It emphasizes that developers and deployers of AI systems can be held liable for harm caused, thereby fostering a culture of responsibility and diligence. However, the pace of technological development is outstripping existing legal frameworks, creating gaps that could be exploited for harmful purposes. Current treaties and agreements often do not sufficiently cover the complexities of AI technologies, leaving significant legal ambiguities.

The introduction of the AI Liability Directive highlights the urgent need for a coordinated international effort to update legal standards and frameworks to effectively address these challenges. Experts from various sectors emphasize that strong legal mechanisms are essential to ensure that AI technologies are developed and deployed responsibly, thereby minimizing risks and safeguarding human rights.

Finally, the AI Act emphasizes the need for AI systems to be tested for fairness and non-discrimination, particularly in high-risk sectors like defense. However, the reality is that many AI systems used in defense are developed in secrecy, often without the same level of ethical scrutiny as civilian technologies. Moreover, while the GDPR provides a strong foundation for protecting personal data and preventing unlawful profiling, its principles should be explicitly integrated into military AI systems. Establishing a framework that mandates adherence to these principles will reinforce the commitment to human rights in all AI applications.

As AI becomes more prevalent in defense, maintaining transparency and accountability is vital. Defense agencies must subject AI systems to rigorous testing, validation, and independent oversight. Implementing audit trails for AI decision-making ensures that each action can be traced and reviewed, preventing misuse and providing accountability when AI systems fail or make incorrect predictions. In that respect international organizations, such as NATO and the United Nations, are increasingly involved in setting standards for AI use in defense. The future of AI-driven technologies, especially in military context, must focus on creation of AI systems' ability to make explainable, transparent decisions that support both security objectives and human dignity.

Conclusions

In conclusion, AI-driven defense systems present significant legal and ethical challenges that must be thoroughly addressed to ensure their responsible and accountable use. The growing autonomy of AI technologies complicates traditional legal frameworks surrounding accountability, raising pressing questions about who is liable when AI systems cause harm. This complexity is particularly pronounced in military contexts, where decisions made by autonomous systems can have life-and-death consequences. The existing legal acts often struggle to keep pace with the rapid evolution of AI, highlighting the urgent need for clearer regulations that delineate responsibility among programmers, military commanders, and states deploying these technologies.

While the AI Act and General Data Protection Regulation provide essential regulatory foundations, they are insufficient to fully address the complexities and unique risks posed by AI in defense. There is a pressing need for regulations that not only promote accountability and transparency but also ensure that AI systems are designed to operate within established legal frameworks governing warfare.

Finally, the ethical implications surrounding AI in defense extend beyond legal compliance. The development and deployment of AI algorithms must prioritize fairness, non-discrimination, and respect for human rights. This requires the establishment of robust governance frameworks that promote ethical standards throughout the lifecycle of AI systems used in defense.

Ultimately, while the AI Act and GDPR lay important foundations for regulating high-risk AI systems, more comprehensive legal frameworks are needed that specifically address the challenges posed by AI in the defense sector. The rapid pace of AI development necessitates urgent action to prevent the misuse of these technologies in military contexts and to ensure they align with ethical principles and international legal standards. Fostering a responsible and ethical approach to AI in defense is not just a legal imperative; it is crucial for maintaining public trust and safeguarding human rights in an increasingly complex technological landscape.

References

1. UNESCO, "Recommendation on the Ethics of Artificial Intelligence," HS/BIO/REC-AI-ETHICS/2021, 2021, https://unesdoc.unesco.org/ark:/48223/pf0000380455.

2. European Parliament and Council, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonized Rules for Artificial Intelligence (Artificial Intelligence Act)," *Official Journal of the European Union*, July 12, 2024, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689.

3. United Nations Office for Disarmament Affairs, "Lethal Autonomous Weapon Systems (LAWS)," United Nations, CCW/GGE.1/2023/CRP.1, 2023.

4. Stefan Larsson and Fredrik Heintz, "Transparency in Artificial intelligence," *Internet Policy Review* 9, no. 2 (2020), https://doi.org/10.14763/2020.2.1469.

5. European Commission, "Ethics guidelines for trustworthy AI," High-Level Expert Group on Artificial Intelligence, April 08, 2019.

6. European Parliament and Council, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13 2024 Laying Down Harmonized Rules for Artificial Intelligence (Artificial Intelligence Act)," *Official Journal of the European Union*, July 12, 2024.

7. European Parliament and Council of the European Union, "Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)," *Official Journal of the European Union*, L 119, May 04, 2016.

8. European Parliament and Council, "Regulation (EU) 2024/1689," Article 6, 2024.

9. International Committee of the Red Cross (ICRC), "What is International Humanitarian Law?" July 05, 2022.

10. IEEE Standards Association, "Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems," 2019, https://standards.ieee.org/initiatives/ethical-ai.html.

11. Pablo G. Bejerano, "AI's Black Box Problem: Why Is It Still Indecipherable to Researchers?" *El País*, October 15, 2024, https://english.elpais.com/technology/2024-10-15/ais-black-box-problem-why-is-it-still-indecipherable-to-researchers.html.

12. Jayshree Pandya, "The Dual-Use Dilemma of Artificial Intelligence," *Forbes*, January 07, 2019, https://www.forbes.com/sites/cognitiveworld/2019/01/07/the-dual-use-dilemma-of-artificial-intelligence.

[13] European Commission, "Liability Rules for Artificial Intelligence," February 19, 2020, https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en.

About the Author

Maria Lachova graduated from Sofia University, Faculty of Law in 2022. In 2016-2017, she studied in the Bachelor's Program in International and European Law at the University of Groningen, the Netherlands. She has undergone a number of certification trainings at the European Human Rights Training Programme for Lawyers (HELP Programme) of the Council of Europe. Her professional experience includes working in civil enforcement, and she subsequently gained expertise in digital law, data protection, and intellectual property rights. Ms Lachova's interests focus on patent and trademark registration, telecommunications and media law, and information and communications technology law.
https://orcid.org/0009-0007-5203-4121