COT Legal
Rechtsanwältin Claudia Otto

COT Legal
Rechtsanwältin Claudia Otto
Rudolf-Diesel-Str. 11
69115 Heidelberg
Germany
e-mail: claudia.otto@cotlegal.de
pkey: https://keys.openpgp.org
internet: https://cotlegal.de
phone: +49 6221 648 4036

White Paper

Introducing the Framework for Trustworthy Military AI: A Timely Response to Evolving Security Needs

9 March 2025

About this White Paper

In light of the ongoing rearmament initiatives – particularly those of the European Union (EU)[1] – it is crucial to bridge the gap between these defence efforts and the rapid advancements in Artificial Intelligence (AI). As AI technologies evolve, they present both opportunities and risks that must be managed effectively.

The EU Artificial Intelligence Act (EU AI Act), which is designed for peacetime applications, explicitly excludes AI systems and their outputs that are intended for military, defence, or national security purposes. While this exclusion may be well-intentioned, it introduces significant practical challenges. Furthermore, the innovation incentives for AI technologies within this framework require urgent revision, as they appear to be more detrimental than beneficial. Last but not least, the EU Commission must address the inconsistencies that threaten to undermine the defence capabilities of its member states.

Notwithstanding the EU AI Act's shortcomings, there is a current but unmet global need for guiding principles focussing on AI in the military domain. This White Paper is the first to outline principles for Trustworthy Military AI, enabling immediate implementation in development and procurement processes, fostering dialogue between stakeholders and contributing to building a resilient national and collective defence.

Checklists outlining the key requirements are provided prior to the detailed explanations and are primarily designed to facilitate practical application. Spaces have been allocated for annotations.

☞ The White Paper does not answer your questions? Contact the author directly by sending an e-mail to claudia.otto@cotlegal.de. Consider encryption using the corresponding public key (https://keys.openpgp.org).

[1] Cf. ReArm Europe, Press statement by President von der Leyen on the defence package, 4 March 2025, https://ec.europa.eu/commission/presscorner/detail/sv/statement_25_673 (last accessed 9 March 2025).

About the author

Claudia Otto is the founder of COT Legal, a renowned law firm based in Germany, a distinguished lawyer and advisor to legislators on critical technologies, including artificial intelligence, digital technologies, and biotechnology. With a profound expertise in risk assessment and management, she plays a pivotal role in shaping the legal landscape of these dynamic fields.

As a thought leader, Claudia teaches Strategic Foresight, Corporate Governance, and Digital Innovation at Fresenius University of Applied Sciences in Germany. She has authored over 50 articles that contribute to the discourse surrounding legal and ethical implications of technology.

Leveraging her extensive consulting experience across high-risk industries and her in-depth research in safety and security management, Claudia is currently writing her Master's Thesis on AI risk assessment in alignment with the EU AI Act, further solidifying her position as an expert in the intersection of law and technology.

Content

A. Navigating the landscape of Trustworthy Military AI: key checklists

I. The 8 key requirements for Trustworthy Military AI

	Military AI should be:	*meaning:*
1)	evident	having a well-defined military purpose and being clearly distinguishable from civilian and humanitarian AI technologies.
2)	lawful	in compliance with all applicable legal requirements, in particular international humanitarian law (IHL) and international human rights law (IHRL).
3)	ethical	adhering to ethical principles and values, especially those that are internationally shared, such as - human autonomy, requiring transparency and explainability for effective oversight; - prevention of unnecessary harm to civilians, civilian objects, and the environment.
4)	resilient	robust, meaning safe and secure, as well as reliable and accurate from both a technical perspective and that of the deployer and individual user.
5)	sovereign	independent from changes in contractual partners' interests, values, and situation, that could impact the resilience and/or use of Military AI.
6)	interoperable	the capability of military equipment to function cohesively through shared standards and, for example, be replaceable without disrupting entire systems.
7)	resource-efficient	the consumption of resources required for operation is low, which is particularly important for users in situations without access to charging facilities.
8)	explained	approvers are provided the necessary information to make informed decisions, while users are provided with information to enable appropriate use.

II. The 6 key requirements for procurement of Trustworthy Military AI

To leverage the eight qualities of Trustworthy Military AI,

the procurement of Trustworthy Military AI considers

1) the (intended) military purpose
2) lawfulness
3) resilience to be prioritised over price and comfort
4) sovereignty to be prioritised over price and comfort
5) interoperability when resilience and sovereignty are ensured
6) resource-efficiency prioritised over price and comfort.

III. The 3 key principles for the responsible use of Trustworthy Military AI

The responsible use of Trustworthy Military AI means

the use of Trustworthy Military AI considering

1) lawfulness (e.g. distinction, proportionality and precautions in attack)

2) accountability (e.g. responsible human chain of command and control)

3) the user's agency and oversight.

B. Introducing the framework for the concept of Trustworthy Military AI

Amidst the global search for guiding principles for Military AI technologies[2] and the urgency to enhance defence capabilities in response to pressing global challenges, this White Paper aims to offer a solution. It is particularly important, because the prominent EU AI Act offers few answers and raises many questions. While its focus is primarily on individual fundamental rights, the EU AI Act does not pay much attention to the details of defence.

The White Paper's framework for the concept of Trustworthy Military AI is non-binding, yet it can assist procurement officials in making informed decisions and help to effectively communicate their needs to Military AI providers. Simultaneously, Military AI providers get valuable insights to drive the development of innovative defence solutions.

I. The EU AI Act's legal uncertainty

The EU AI Act explicitly excludes AI systems and their outputs that are intended solely for military, defence, or national security purposes, cf. Article 2 (3) of the EU AI Act. Unfortunately, it fails to define these three areas, leading to legal uncertainty. Article 4 (2) of the Treaty on European Union (TEU) and Chapter 2 of Title V of the TEU (cf. Recital 24 of the EU AI Act) also lack definitions. Nevertheless, it is possible to gain a clearer understanding of these terms through alternative interpretations and context:

1. National security

Article 4 (2) of the TEU only provides that national security remains the sole responsibility of each Member State. According to the Court of Justice of the European Union (CJEU), that responsibility corresponds to the primary interest in protecting the essential functions of the State and the fundamental interests of society and encompasses the

[2] Cf. UNIDIR, Governance of Artificial Intelligence in the Military Domain: A Multi-Stakeholder Perspective on Priority Areas, 5 September 2024, https://unidir.org/publication/governance-of-artificial-intelligence-in-the-military-domain-a-multi-stakeholder-perspective-on-priority-areas/; REAIM 2023 Call to Action, Responsible AI in the Military domain Summit, 15-16 February 2023, https://www.government.nl/ministries/ministry-of-foreign-affairs/documents/publications/2023/02/16/reaim-2023-call-to-action; US Department of State, Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, 9 November 2023, https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2/ (all last accessed on 9 March 2025).

prevention and punishment of activities capable of seriously destabilising the fundamental constitutional, political, economic or social structures of a country and, in particular, of directly threatening society, the population or the State itself, such as terrorist activities.[3]

The exemption for national security concerns in EU secondary law applies exclusively to state actors and does not extend to private entities, as highlighted by the CJEU.[4] However, the EU AI Act appears to overlook the jurisprudence of the CJEU, which is part of EU primary law and takes precedence.

2. Military vs. defence

Defence encompasses military and civil capabilities (cf. Article 42 (3) of the TEU). "Civil defence"[5] may equal or include civil protection, but also civilian support for the armed forces, i.e. the military. Those defence-related civilian purposes, however, are not excluded under the EU AI Act. Accordingly, Article 46 (2) of the EU AI Act grants civil protection authorities a right to derogate from the conformity assessment procedure. Consequently, civil protection, as an integral part of defence, remains within the framework of the EU AI Act, making the term "defence" and the exception somewhat misleading.

To enhance legal certainty, it seems beneficial to treat military and non-civil defence purposes as synonymous within the context of the EU AI Act.

II. When is the EU AI Act applicable?

1. The legal situation

The dual-use nature of AI technologies suggests that exclusive military purposes are likely to be relatively rare, particularly when considering the overlaps associated with

[3] CJEU, judgement dated 6 October 2020, "La Quadrature du Net (LQdN)", ECLI:EU:C:2020:791, para 135.
[4] Cf. CJEU, judgement dated 6 October 2020, "La Quadrature du Net (LQdN)", ECLI:EU:C:2020:791, paras 99-104.
[5] Cf. Article 61 of Protocol I Additional to the Geneva Conventions of 12 August 1949 and relating to the Protection of Victims of International Armed Conflicts (Protocol I) defines civil defence through the humanitarian tasks carried out for the protection of the civilian population against the dangers arising from hostilities or disasters, and to help it to recover from the immediate effects.

administrative tasks and workplace applications. This raises the question of how military and civilian purposes can be effectively delineated under the EU AI Act.

Recital 24 of the EU AI Act addresses this issue as follows:[6]

> If an AI system developed, placed on the market, put into service or used for military purposes is used temporarily or permanently for non-military purposes, the *deployer* must comply with the EU AI Act (cf. Recital 24 sentences 4 and 5 of the EU AI Act).
>
> If an AI system is placed on the market or put into service for military *and* civilian purposes, the *provider* must comply with the EU AI Act (cf. Recital 24 sentence 6 of the EU AI Act). However, due to the exemption of military use, the military *deployer* does not have an obligation under the EU AI Act.
>
> If an AI system is placed on the market for civilian purposes but used for military purposes, the EU AI Act does not apply for the user, i.e. *deployer* (cf. Recital 24 sentence 8 of the EU AI Act).

However, many questions remain unanswered, particularly in the context of defence, where military and civilian purposes often overlap.

2. Implications of this legal situation

Given that AI innovation primarily originates from the private sector and gradually transitions to the military, while state-initiated developments often face challenges related to limited funding, bureaucratic processes, and slower adaptation to technological advancements, there is a pressing need to streamline cooperation to foster innovation and ensure modern military capabilities.

The peacetime and individual-focused EU AI Act, however, poses challenges for such cooperation. Providers may find it more strategic to develop and market their AI

[6] For improved readability, the text has been condensed to refer specifically to military purposes. Italics for emphasis are made by the author.

technologies exclusively for military purposes, potentially outpacing state-initiated efforts while shifting responsibility and liability to the deployers.

This creates a dilemma for the often underfunded civil defence actors, which need to adapt AI technologies originally developed for military purposes to civilian applications, such as saving lives, ultimately at the expense of the affected civilians. Furthermore, the separation of civilian and military purposes is likely to undermine civilian support for the armed forces and thus the defence of a state, as civil defence actors often do not have the financial means to comply with the EU AI Act to the same extent as private entities.

Lastly, there is a risk that innovation for civilian users will decelerate, as providers may focus on capturing a portion of the massive defence budgets. The EU AI Act appears to function more as a barrier or funnel than an incentive, prompting the question of whether the disadvantages imposed on civil defence – potentially even harming the armed forces – are justified. Simultaneously, the civilian pool of ideas and developments that could also benefit the armed forces may diminish or even dry up.

3. The need for reassessment and amendments to the EU AI Act

The EU Commission should reassess its innovation incentives for AI technologies in light of the emerging war economy, which appear to be more detrimental than beneficial. Furthermore, it is crucial to resolve inconsistencies that undermine the defence capabilities of EU member states. For instance, the conflict between Article 46 (2) and Recital 24, sentences 4 and 5 of the EU AI Act highlights that the current disadvantageous regulations are unintended and merit re-evaluation. It is perplexing that a civil protection authority may derogate from the conformity assessment procedure for immediate putting into service (and use) while being required to comply with the EU AI Act when temporarily using an AI system with a military purpose for civilian purposes.

These are discrepancies that national legislators cannot rectify. However, Article 113 of the EU AI Act provides some relief and time for the EU Commission to deliver crucial amendments: The relevant provisions will not come into effect until 2 August 2026, while those pertaining to so-called high-risk AI systems according to Article 6 (1) of the EU AI Act will not apply until 2 August 2027.

III. What is Military AI?

Military AI refers to AI technologies intended for use in the military (defence) domain, including AI (software), high-performance computing, cloud and edge computing, and data analytics.[7]

Military AI is a broad term that encompasses a wide range of applications, including more commonplace functions such as extracting valuable information from extensive volumes of unstructured data. Military AI should therefore not be equated with Lethal Autonomous Weapon Systems (LAWS), even though they can fall under its umbrella.[8]

Building on the above discussion, its (intended)[9] military (defence) purpose should be clearly distinguishable from civilian purposes. It is the first of the eight requirements for Trustworthy Military AI.

IV. What is Trustworthy Military AI?

Given the lack of legal and non-legal frameworks, principles, or other shared understandings regarding what is necessary to trust Military AI,[10] a guiding list of requirements could help bridge the gap in the urgent need for innovations in the military domain.

Trustworthy Military AI shares some similarities with civilian Trustworthy AI, outlined in the Ethics Guidelines for Trustworthy AI by the High-Level Expert Group on Artificial Intelligence (AI HLEG),[11] but it is only partially aligned:

---

[7] Cf. Commission Recommendation of 3 October 2023 (C(2023) 6689 final) (last accessed on 9 March 2025).
[8] Cf. United Nations Office for Disarmament Affairs, Lethal Autonomous Weapon Systems (LAWS), https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/ (last accessed 9 March 2025).
[9] Cf. Art. 3 (12) of the EU AI Act.
[10] Cf. UNIDIR, Governance of Artificial Intelligence in the Military Domain: A Multi-Stakeholder Perspective on Priority Areas, 5 September 2024, https://unidir.org/publication/governance-of-artificial-intelligence-in-the-military-domain-a-multi-stakeholder-perspective-on-priority-areas/; REAIM 2023 Call to Action, Responsible AI in the Military domain Summit, 15-16 February 2023, https://www.government.nl/ministries/ministry-of-foreign-affairs/documents/publications/2023/02/16/reaim-2023-call-to-action; US Department of State, Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, 9 November 2023, https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2/ (both last accessed on 9 March 2025).
[11] Cf. Ethics Guidelines for Trustworthy AI, 8 April 2019, https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1 (last accessed on 9 March 2025).

Providers should also gain a competitive advantage by integrating Trustworthy Military AI into their products and services, maximising the benefits of this technology while effectively mitigating associated risks.

States, militaries, and military users, however, have interests in procuring and utilizing Military AI that differ significantly from those of civilian deployers or consumers. Moreover, it is particularly important to acknowledge that Military AI comes with its own unique set of risks, including potential dependencies that could undermine operational effectiveness and compromise the very essence of national or collective defence. Therefore, the concept of Trustworthy Military AI must specifically address both the (military) defence-related risks and benefits involved.

To qualify as trustworthy in general, Military AI must be:

1. *evident*, i.e. having a well-defined military purpose and being clearly distinguishable from civilian and humanitarian AI technologies.

2. *lawful*, i.e. in compliance with all applicable legal requirements, in particular international humanitarian law (IHL) and international human rights law (IHRL).

3. *ethical*, i.e. adherent to ethical principles and values, especially those that are internationally shared, such as human autonomy, which requires transparency and explainability for effective oversight, and the prevention of unnecessary harm to civilians, civilian objects, and the environment.

4. *resilient*, i.e. robust, meaning safe and secure, as well as reliable and accurate from both a technical perspective and that of the deployer and individual user.

5. *sovereign*, i.e. independent from changes in contractual partners' interests, values, and situation, that could impact the resilience and/or use of Military AI.

6. *interoperable*, i.e. the capability of military equipment to function cohesively through shared standards and, for example, be replaceable without disrupting entire systems.

7. *resource-efficient*, i.e. the consumption of resources required for operation is low, which is particularly important for users in situations without access to charging facilities.

8. *explained*, i.e. approvers are provided the necessary information to make informed decisions, while users are provided with information to enable appropriate use.

For instance, a Military AI-based system integrated into a fighter jet, which relies on continuous updates, upgrades, and data exchanges with a contractual partner whose interests and values may evolve, may fall short of the standards required for Trustworthy Military AI.

Under certain circumstances, there may be intersections between some of these listed requirements, particularly when legislation has been implemented that mandates criteria following the lawfulness requirement.

To leverage the qualities of Trustworthy Military AI, the procurement of Trustworthy Military AI should consider the following six key requirements:

1. the (intended) military purpose
2. lawfulness
3. resilience to be prioritised over price and comfort
4. sovereignty to be prioritised over price and comfort
5. interoperability when resilience and sovereignty are ensured
6. resource-efficiency prioritised over price and comfort.

V. Military AI system

Considering that the development and application of AI technologies often begin in the private sector before being integrated into military systems, the Trustworthy Military AI framework aims to ensure compatibility with the EU AI Act and other relevant legislation.

The definition of an AI system[12] in the EU AI Act has raised considerable debate, making it potentially less suitable for the military domain, where clarity and prompt legal decision-making are essential. However, a definition of a Military AI system must align with the EU AI Act's framework, especially in cases of military and civilian use. A possible conciliatory solution could be this definition:

> *A machine-based system that fully or partially autonomously processes and analyses diverse input data to generate output based on input-derived information that facilitates or relieves the user of decision-making or action.*

This definition takes into account that the demand for AI solutions in the military domain is primarily driven by the need to organize and leverage unstructured data, develop enhanced capabilities, and facilitate faster decision-making, rather than being centred on sophisticated AI tools alone. Yet, the definition also includes these.

VI. Responsibility

It is important to note that while decision-making and actions may be delegated to a machine-based system, the obligations and accountability under (international) law cannot be transferred to such systems. The following three key principles for the responsible use of Trustworthy Military AI illustrate this point, demanding:

1. lawfulness (e.g. distinction, proportionality and precautions in attack)
2. accountability (e.g. responsible human chain of command and control)
3. the user's agency and oversight.

☞ The White Paper does not answer your questions? Contact the author directly by sending an e-mail to claudia.otto@cotlegal.de. Consider encryption using the corresponding public key (https://keys.openpgp.org).

[12] Article 3 (1) of the EU AI Act.