The protection of humanity from the weaponization of artificial intelligence and the need for an international convention on AI peace ethics

The protection of humanity from the weaponization of artificial intelligence and the need for an international convention on AI peace ethics

The automation of kill chains in the era of high-tech wars, the future of peace ethics and the role of religious actors as advocates for the sanctity of life

Dr. Dietrich Werner, Berlin/Geneva

Integration of AI into systems of warfare as stimulus for the changing face and brutalization of wars

One frightening prospect which has characterized recent wars and military tensions in several contexts in the past two decades is the increased use of sophisticated technologies of warfare and the deployment of AI supported software tools called AI DSS (Artificial Intelligence Decision Support Systems)[1] which are dramatically increasing the speed in which crucial decisions on life and death are made in military actions. Whether it is in the war between Russia and the Ukraine, in the war in Jemen or in DRC Kongo or the military threats of China against Taiwan or India – in all these contexts drones, new military robots and automated systems of mass destruction dominate modern war scenarios. Winning or losing a war becomes dependent on the speed of technological innovations and the race who is faster in military communication technologies. The rapid and often unnoticed integration of AI into armed systems is one of the fastest processes transforming the face of war – and for the most part not for the better. It is changing not only every aspect of warfare in the actual battlefield but also every aspect of military and political strategies which is coming under heavy pressures of new alliances between the military industrial complex, global tech companies and the armaments industry with its enormous economic gains. It is not just phantasy of Science Fiction Literature, but it is reality in many contexts that ideas and realized projects of fully autonomous military robots are circulating which are programmed to act solely by algorithms which have been fed by millions of pieces of data. China has created a first army of killer robots which are able to make decisions only informed about by pre fed data. The recent debate on China's possible deployment of robots[2], automated killer machines like Sharp Claw and Mule 200[3] along its border with India after the Covid-19 pandemic is only one example about the actual relevance of these technological advances. In addition recent messages suggest that China's new 6th-gen stealth fighter "J-36" has been constructed in a way to serve as a command centre for a swarm of combat drones which would allow to launch deadly airstrikes on multiple targets. It therefore needs to be seen as part of a "broader shift to network-centric warfare".[4]

Examples like this underline "how technology has changed from hand-operated tanks that were used in WWI to the present day's more autonomous weapon systems that are challenging the conventional distinction between combatants and non-combatants as well as between machines and soldiers. The military domain is undergoing rapid transformation due to the adoption of artificial intelligence and other technological advances. The use of this technology in military affairs has the capacity to enhance efficiency while also presenting (serious) ethical challenges."[5]

The diversity of AI supported military technologies and the trend towards more autonomous weapon systems

While robotics systems are constantly redesigned and becoming more sophisticated, this by far is not the only emerging technology which is substantially reshaping the face and characteristic features of warfare today. There are several

---

[1] See: Arthur Holland Michel: Decisions, Decisions, Decisions: computation and Artificial Intelligence in military decision-making . Report for the International Committee of the Red Cross (ICRC). 2024, in: https://shop.icrc.org/decisions-decisions-decisions-computation-and-artificial-intelligence-in-military-decision-making-pdf-en.html

[2] Deployment of Robot Soldiers by China along the Indian Border and Its Repercussions, in: https://www.cescube.com/vp-deployment-of-robot-soldiers-by-china-along-the-indian-border-and-its-repercussions; see also: https://www.youtube.com/watch?v=ZFSQ0GwHCqY

[3] Early versions of these automated killer machines exist already since 2014, see: China's Army Now Has Killer Robots: Meet the 'Sharp Claw', in: https://nationalinterest.org/blog/buzz/chinas-army-now-has-killer-robots-meet-sharp-claw-145302

[4] https://www.scmp.com/news/china/military/article/3293421/chinas-mystery-6th-gen-stealth-fighter-poised-be-command-centre-combat-drones; and: https://www.pressreader.com/china/south-china-morning-post-6150/20250106/281685440479448?srsltid=AfmBOor2pjUjKB5N7r_QvBXuIVGhwMDX1RE1Ec7-XKr4dbcFF9cZzqqc

[5] Khola Junaid: Emerging Technologies and their Impact on Warfare. Modern Diplomacy, June 11, 1024, in: https://moderndiplomacy.eu/2024/06/11/emerging-technologies-and-their-impact-on-warfare/

cutting-edge and inventive technologies that are currently developing or are being introduced and many of them also tested in actual battlefields. “Some of the new technologies include artificial intelligence, 3D printing, the Internet of Things, blockchain, and others. These technologies have transformed the methods by which military strategies are formulated. The use of unmanned aerial vehicles (UAVs) that include automated drones allows for real time data gathering, surveillance and reconnaissance, as well as Internet of Things (IoT) devices that allow for effective and efficient communication on the battlefield. Also, biotechnology, nanotechnology, augmented reality and precision guided hypersonic missiles. All of these are the products of these emerging technologies. These systems, such as sophisticated unmanned aerial vehicles, possess the capability to autonomously identify and neutralise possible dangers with a pretty high level of precision. The emerging technologies used are spearheading an entirely new generation of warfare. These instruments have significance in several domains of human existence; however, their adaptable uses also make them potentially perilous and destructive weapons.”[6] The continued protractedness of the war of Russia against the Ukraine - despite all feeling of immense exhaustiveness on both sides - is not only due to the adamant nationalistic and imperial Russian ideology of Russkij Mir and the religious justification of war, it is also due to the fact that this war by many is seen as a war and race for the better military digital technologies. There are illusionary perceptions that the more digitization of warfare is accelerated the sooner the end of the war is going to come - most probably a tragic misunderstanding and miscalculation of contemporary military strategists: The war in the Ukraine has become very much a technology war, analysts have dubbed it the first commercial space war, the first full-scale drone war, the first AI war. Research has “examine(d) the role of drones, cyber warfare, software-defined warfare and AI, and space technologies in the war in Ukraine. It reveals how the conflict has become a testing ground for new military systems. Innovation is happening at high speed. Thousands of drones, so many of which as never used before, have been deployed in military confrontation. Cloud services and cyber defences have provided existential support. Software, often AI-enabled, is used to improve legacy systems. Without support from satellites, Ukraine would not be able to defend its territory.”[7]

People not interested so much in militaria and weaponry seldom look into new commercial brochures about modern technologies of warfare. More people should do - as the intensity of the militarization of technological and digitization progresses is indeed shocking and only the iceberg of a broader movement of one-sided militarization of the whole concept of security in which more and more tech giants, militarists and right-wing extremists are joining hands. Costly brochures and coloured pictures praise the capabilities of modern AI integrated systems of “defence” technologies and play with the fascination of clean technological systems where you do not see the deadly impact of your own weapons systems except on a clean and colourful screen.[8] No doubt: The new era of high-tech wars has begun.[9]

Misleading promises of the introduction of AI supported software into military applications

The almost religious promises and expectations combined and associated with an ever-greater acceleration of the integration of modern AI technologies into military applications is obvious in many brochures from the armaments industry and from military circles:

- Crunching big data and using AI systems for harvesting, combining and communicating them will enable faster decision making in the battlefield;
- Military software tools of AI are capable of animating sensors which is important for soldiers and people in the command chain;
- Application of military AI tools will enable more “rapid-fire coordination of fleets of autonomous military platforms”[10];

---

[6] Khola Junaid: Emerging Technologies and their Impact on Warfare. Modern Diplomacy, June 11, 1024, in: https://moderndiplomacy.eu/2024/06/11/emerging-technologies-and-their-impact-on-warfare/

[7] Ulrike Franke, Jenny Söderström: Star tech enterprise: Emerging technologies in Russia’s war on Ukraine, in: European Council on Foreign Relations, 5 September 2023, in: https://ecfr.eu/publication/star-tech-enterprise-emerging-technologies-in-russias-war-on-ukraine/

[8] See a brochure of The European Defence Agency in Brussels: https://eda.europa.eu/docs/default-source/eda-magazine/edm22singleweb.pdf; see also from the US context: https://www.armyupress.army.mil/Journals/NCO-Journal/Archives/2023/July/Future-Weapons-Technology-of-2040/

[9] https://www.economist.com/leaders/2023/07/06/a-new-era-of-high-tech-war-has-begun

[10] See a brochure of The European Defence Agency in Brussels, page 7: https://eda.europa.eu/docs/default-source/eda-magazine/edm22singleweb.pdf

- Broad based introduction of AI supported target selection systems will increase the number of precision targets for bombings thereby increasing the military pressure on any aggressor;
- Often also the expectation is added that military AI supported systems will allow for a more rapid and clear distinction between combatants and non-combatants.
- Therefore, at the end of the ideological rhetoric supporting the weaponization of AI and sophistication of related technologies there always is the combination of the promises of immunity (over against attacks of the adversary in the military battlefield) and the promises of impunity (over against any attacks on the home-based ethical accusations due to the anonymization of final moral responsibility lines).

Israels war in the Gaza strip and the role of "The Gospel" and "Lavender" targeting software as example of horrific consequences of the weaponization of AI

Whether the military salvation promise of the integration of AI into warfare technology holds true or whether the promise of a more "clean", "effective", "resource-saving" and less costly human warfare is one of the greatest self-deceits in our time can be tested by one major case study which is from the Middle East: looking at the "technological innovations" of the Israeli "war of defence", the 'War of Iron Swords' (Hebrew: מלחמת חרבות ברזל), in the Gaza strip. This war has been going on since October 2023 after the brutal terrorist attacks of Hamas forces in Israel which has killed more than 1200 Israelis and in turn led to the deadliest military campaign against the Palestinians. The Gaza war has killed more than 44.700 Palestinians (70% of which women and children) while harming and wounding thousands more.[11] It is disturbing to note in this context that while the war still has not yet ended in late December 2024 a first celebrative "Israeli DefenseTechSummit" was brought together in Tel Aviv University to celebrate achievements in new technological innovations of warfare. The conference brought together IDF military specialists, particularly from the Lotem Unit (Digital and Data Unit of IDF), and American representatives of high tech companies, future government officials from the Trump administration and other international IT specialists: The advertisement on the website reads: "Defense Tech Summit 2024 unites top-tier attendees, including CEOs, CISOs, senior executives, venture capitalists, investors, government officials, and startup leaders. Engage directly with key decision-makers, top military figures, and professionals from defense, intelligence, and security sectors across the region."[12] The two-day event featured panels like on "The Future of Global Conflict," "Challenges of Iron Swords" (the IDF's name for the war in Gaza) and "Exploring Innovation in Drone Technology." The DefenseTech Summit was meant to showcase "Israel's cutting edge technologies and strategies for addressing global security," but observers stated that it felt more like a celebration of a new and unrestrained era of techno-militarization inaugurated by Donald Trump's re-election.[13]

What is behind this renaissance of a major technological advance in military arms is a revolution implied in the technological systems to define military targets for bombings – an AI enabled software development which is known since around 2021:

"In 2019, the Israeli government announced the creation of a 'targeting directorate' to produce targets for the Israeli Defense Force (IDF), especially the Israeli Air Force (IAF). In previous conflicts such as in the 2014 war and before, the IAF would run out of targets after just a few weeks of fighting, having hit all the targets of which they knew. The targeting directorate was created to mitigate this shortage by preemptively creating a 'bank' of militant targets prior to any conflict, thereby ensuring enough targets when hostilities began. The directorate, consisting of hundreds of soldiers and analysts, creates targets by aggregating data from a variety of sources – drone footage, intercepted communications, surveillance data, open source information, and data from monitoring the movements and behaviour of both individuals and large groups."[14]

---

[11] See on statistics: https://en.wikipedia.org/wiki/Casualties_of_the_Israel%E2%80%93Hamas_war#:~:text=As%20of%2010%20December%202024,includes%20179%20employees%20of%20UNRWA.

[12] https://deftech-summit.com/

[13] See: Sophia Goodfriend: With Gaza war and Trump's return, Silicon Valley embraces a military renaissance. At Israel's first DefenseTech Summit, corporate leaders and army officials openly touted their partnership in AI-driven warfare and surveillance. In: https://www.972mag.com/gaza-war-trump-silicon-valley-military/, 31 December 2024

[14] Noah Sylvia: Israel's Targeting AI: How Capable is It?, 8 February 2024, in: https://www.rusi.org/explore-our-research/publications/commentary/israels-targeting-ai-how-capable-it

The result of the new AI supported military target identification software allowed for a rapid move from identifying 50 military targets per year to a situation in which 250 targets could be identified per day. Thus, an immense acceleration of constant attacks became technologically possible. What previously had been perceived as totally impossible by relying on man-powered observation and secret service information became possible even beyond expectations, moving from human to machine based intelligence so that the IAF could already proudly state in November 2023 that it had bombed the unimaginable number of more than 22.000 targets in the Gaza strip.[15] Cynically enough the name given to this software which was now used regularly to assist with the identification of military and human targets in all the days of the Gaza war was called "Habsura", which is a Hebrew term standing for "the Gospel". Research by specialists using classified information and interviews with members of specialists units within IDF has provided more information on this system and its naming: "The "Gospel" (Habsora) AI system produces bombing targets for specific buildings and infrastructure in Gaza, working in conjunction with other AI tools. Notably, the specific usage of the term "Gospel" implies a biblical connotation of infallibility and ultimate authority potentially attributed to the Israeli system, reflecting its trusted and authoritative status within the IDF. Thereby, the connotation underscores the system's critical role in justifying and executing military strategies, much like the unquestioned truth of the religious gospel."[16]

While there is a lot of internal controversy in assessing how far the precision and infallibility aura of the "Gospel" software technically have proven as a fact in reality[17] and to what extend the claim can be substantiated that final decisions on actual bombings are still controlled by human minds and in transparent manner the doubts are overwhelming. Many support the conviction that the "Gospel" AI system used heavily by IDF forces has in reality not helped to humanize warfare but instead has significantly contributed to increase atrocities against civilian populations and was a major component resulting in a de facto genocide against the Palestinian population.

Demythologizing the role of AI supported technologies of warfare and the dangers of shortening and automatization of the "kill chain"

Thus, in contrast to its military salvation promises (mentioned above) the critical literature and research review on the impact of the "Gospel" AI program which could be done here can be summarized at this stage by the following six observations:

- that the AI systems "Gospel" and "Lavender" which were introduced by the Israeli army (originally designed for urban and regional warfare against a big country like Iran and not for a densely populated, very small area like in Gaza) have to large extend – together with the use of so-called "dumb bombs" which lead to indiscriminate killings in the neighbouring area - contributed to a massive and disproportionate rise of the number of non-militant victims in the civilian population; they have made the war more brutal and not more human and limited in in terms of causing casualties in the civilian population;[18]
- that these AI systems have heavily changed the ratio between killed combatants and non-combatants as they "led to the loosening of constraints regarding expected civilian casualties. The use of an artificial intelligence system to generate more potential targets than ever before, appears to have contributed to the destructive nature of the initial stages of Israel's current war on the Gaza Strip… ." The numbers increased from dozens of civilian deaths [permitted] as collateral damage as part of an attack on a senior official in previous operations, to hundreds of civilian deaths as (accepted) collateral damage.[19]
- that these military AI systems for combat situations have led to a dramatic acceleration of decisions on death and life for thousands of persons in a very short time-span and therefore fall short of their promise to keep the last decision in human hands in a responsible manner. Few seconds remaining for decision making which has been made dependent on data which include a lot of ambiguities which would need further investigation and

[15] See: https://en.wikipedia.org/wiki/AI-assisted_targeting_in_the_Gaza_Strip

[16] See: Julius Kurek, Björn Laurin Kühn: Habsora (הבשורה) and Lavender (לָאוֹבִיוֹן) Artificial Intelligence Systems – The Missing Piece Towards a Fully Algorithmically Automated F2T2EA Kill Chain?, in: https://www.epis-thinktank.de/post/habsora-lavender-ai-automated-kill-chain

[17] See: https://www.rusi.org/explore-our-research/publications/commentary/israels-targeting-ai-how-capable-it

[18] https://www.spiegel.de/ausland/israel-hat-ki-den-gazakrieg-brutaler-gemacht-a-f7713750-5888-452b-a5bb-2d979c6ddd11

[19] Yuval Abraham: 'A mass assassination factory': Inside Israel's calculated bombing of Gaza
Permissive airstrikes on non-military targets and the use of an artificial intelligence system have enabled the Israeli army to carry out its deadliest war on Gaza, in: ttps://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/

human checking are simply not enough to make responsible decisions. If hundred or 500 targets are identified and delivered by machines within few days the human ability to check and prove that the final target recommendations are based on solid and reliable information are minimal and simply not feasible.[20] How to maintain human responsibility and agency with AI DSS systems therefore remains a key concern.[21]

- that military software tools like "The Gospel" or "Lavender" need to be seen as one step in the technological attempt to create a totally automated "kill chain" where every single step in the different stages of the decision making leading to a military action and kill command can be done by the machines themselves without being slowed down, hindered or complicated by human intervention. Since a few years military research has been fascinated by the shortening and automation of "kill chains": "The "kill chain" serves to conceptually capture the process of combating an enemy entity. It begins with finding the target and encompasses every subsequent step up to its eventual destruction. One model to structure the kill chain internally is the so-called F2T2EA model, which is divided into six steps. *Finding* the target is a matter of intelligence, which may come in the form of surveillance or reconnaissance. Once the target is identified, its precise location needs to be determined (*fix*) and kept *track* of as the appropriate weapon is selected (*target*). Afterwards, the target can be *engaged* and, once the attack has been carried out, its effectiveness might be *assessed*."[22] The vision of acceleration or even automatization of the "kill chain" (Find, Fix, Track, Target, Engage, Assess) has been a long dream of the American Airforces.[23] The use of targeting software like "Gospel" and Lavender clearly is a potential step towards a realization of a fully automated "kill chain" presented by this F2T2EA model, as with current LAWS (Lethal Automated Weapon Systems) from six steps and stages already five can function to a large extend already automatically and autonomous.[24]
- Thus, the application of such AI DSS Systems also leads to quite essential changes on the roles of soldiers and their individual responsibility: "Comprehensive AI-based DSSs could foster a form of virtual remote command and control, reducing soldiers to executing orders displayed on their devices without critically engaging with these systems' outputs. This scenario risks soldiers not questioning orders, even if they have insights suggesting alternative actions. If soldiers receive commands about enemy positions via AI-based DSS, they might act without verifying the situation. This challenges the *military self-perception of conscious decision-making in the spirit of the "Auftragstaktik"* and in the worst case, result in soldiers "only following orders".[25]
- that the application of AI systems in warfare technologies and target selection according to what we know has not undergone any previous ethical investigation and clearance or admission system in which the ethical ambivalences and shortcomings of these kinds of technologies could have been critically reflected. Thus, there is no agreed national or even international standard for ethical clearance for AI assisted war technologies.
- Finally, a comment needs to be shared on quasi-religious authority claim of these AI supported military software tools and their naming: The fact that the naming of these AI systems is done by using deeply religious terms like "The Gospel" casts some serious doubts on their rationality and the ideologies behind. The term "Habsora" (or: The Gospel) in Christian (and Jewish) tradition stands for a life-giving power and liberating message of liberation, a Gospel of liberation and peace, not of destruction and death. In other words: a message of live and hope is encapsulated in the Gospel as properly understood, but not a death sentence provided by a machine governed by algorithms producing 250 killing commands a day. The language used symbolizes the deeply inhumane nature of such technologies and the cynical spirit and attitude of contempt which is behind which constitutes an act of

---

[20] See also: https://www.tagesschau.de/ausland/asien/israel-gazastreifen-ki-100.html

[21] See: Wen Zhou and Anna Rosalie Greipl: Artificial intelligence in military decision-making: supporting humans, not replacing them, in: https://blogs.icrc.org/law-and-policy/2024/08/29/artificial-intelligence-in-military-decision-making-supporting-humans-not-replacing-them/

[22] Julius Kurek, Björn Laurin Kühn: Habsora (הבשורה) and Lavender (אֲזוֹבִיוֹן) Artificial Intelligence Systems – The Missing Piece Towards a Fully Algorithmically Automated F2T2EA Kill Chain?, in: https://www.epis-thinktank.de/post/habsora-lavender-ai-automated-kill-chain

[23] John A. Tirpak: Find, Fix, Track, Target, Engage, Assess, 1 June 2000, in: Air Force Magazine, https://www.airandspaceforces.com/article/0700find/

[24] See: Unnamed author of blog entry: How Ultra I&C's solutions are improving the F2T2EA kill chain model, August 2, 2023, in: https://www.ultra-ic.com/blog/how-ultra-ic-s-solutions-are-improving-the-f2t2ea-kill-chain-model/; also US Airforce Doctrine on Targeting 12 November 2021, in: https://www.doctrine.af.mil/Portals/61/documents/AFDP_3-60/3-60-AFDP-TARGETING.pdf

[25] Matthias Klaus: Transcending weapon systems: the ethical challenges of AI in military decision support systems, in: https://blogs.icrc.org/law-and-policy/2024/09/24/transcending-weapon-systems-the-ethical-challenges-of-ai-in-military-decision-support-systems/

blasphemy and misuse of core biblical terms which are precious for many people. The more religious language is used indiscriminately and perversely to legitimize or distortingly presenting antagonist values like in lethal and automated AI supported weapon systems the more rigorous and outspoken should be the response of those religious actors which are the custodians of religious values and key terms which embody a centuries old commitment to human dignity, human rights and the sanctity of life in all its forms.

The need for a more intense and public discourse on ethics of AI in warfare

Researchers on the intersection between technological transformations in military weaponry and AI systems in summarizing their observations have voiced clearly the urgent need for more ethical frameworks and compatibility checks with international humanitarian laws (IHL): "The rise of these emerging technologies in warfare have raised crucial questions about the future of warfare. While emerging technologies have increased the efficiency and precision of weaponry, the potential for dehumanization and ethical dilemmas in war are of grave concern and demand immediate attention. The technological advancement worldwide, the autonomous nature of these AI weapons, the ease of access to 3D-printed weaponry, and all other emerging technologies necessitates the establishment of international regulations and ethical frameworks for regulating such technologies."[26]

There are some initial investigations into ethical implications od AI in warfare, like from Elke Schwartz from Queen Mary University in London which highlight serious ethical shortcomings and dilemmas and endeavour to safeguard the principles of humanity and accountability amidst the evolving dynamics of modern warfare:

- "The integration of AI-enabled weapon systems facilitates the objectification of human targets, leading to heightened tolerance for collateral damage
- Automation bias and technological mediation weaken moral agency among operators of AI-enabled targeting systems, diminishing their capacity for ethical decision-making
- Industry dynamics, particularly venture capital funding, shape discourses surrounding military AI, influencing perceptions of responsible AI use in warfare."[27]

Other contributions like from the Waterloo Institute for Artificial Intelligence have observed and analysed the "legal void in which AI weapons operate", the "false promises on civilian protection", the "Autonomy in Weapons Systems and the Struggle for Regulation" and the "Actual IHL Accountability Gap of AI".[28] Contributions clearly underline that ethics of AI in military warfare is more than just dealing with weapon systems, but concerns both the weaponry systems as such, the military decision making structures, issues of military research policies and international regulations concerning political legitimation and ethical assessment as well as legal admission of such systems.[29] Thus the whole range of issues of AI supported warfare and the IHL regulatory systems is at stake.[30] And this is not just an issue in the Middle East or in Eastern Europe, it is a global issue which needs attention within the whole of the UN system.[31] A leading role in this discourse demanding more ethical standards is played by the International Committee of the Red Cross which in

---

[26] Khola Junaid: Emerging Technologies and their Impact on Warfare. Modern Diplomacy, June 11, 1024, in: https://moderndiplomacy.eu/2024/06/11/emerging-technologies-and-their-impact-on-warfare/

[27] Elke Schwarz: The ethical implications of AI in warfare, in: https://www.qmul.ac.uk/research/featured-research/the-ethical-implications-of-ai-in-warfare/; see from the same author from Queen Mary University London: The (im)possibility of responsible military AI governance, Dec 2024, in: https://blogs.icrc.org/law-and-policy/2024/12/12/the-im-possibility-of-responsible-military-ai-governance/

[28] The Ethics of Automated Warfare and Artificial Intelligence. Contributions particularly from Rebecca Croothof, Branka Marijan and Frank Sauer, in: https://www.cigionline.org/the-ethics-of-automated-warfare-and-artificial-intelligence/

[29] Matthias Klaus: Transcending weapon systems: the ethical challenges of AI in military decision support systems, in: https://blogs.icrc.org/law-and-policy/2024/09/24/transcending-weapon-systems-the-ethical-challenges-of-ai-in-military-decision-support-systems/;

[30] Hartwig von Schubert: Addressing ethical questions of modern AI warfare. The use of Artificial Intelligence in war is only a question of time. But it must be regulated in terms of international humanitarian law, in: https://www.ips-journal.eu/topics/foreign-and-security-policy/addressing-ethical-questions-of-modern-ai-warfare-6587/

[31] Sgt. Maj. Ian M. Shaughnessey: The Ethics of Robots in War, in: https://www.armyupress.army.mil/Journals/NCO-Journal/Archives/2024/February/The-Ethics-of-Robots-in-War/

collaboration with the Geneva Academy of International Humanitarian Law and Human Rights has published a number of outstanding studies[32] and blog papers on the related issues.[33]

First steps and possible institutional partners to consider international regulations on challenges posed by AI in warfare

But few authors only proceed to the then imminent question of how to suggest clearcut and binding international regulations, how to initiate international decision making and how to operationalize regulations and compliance with them in order to stop the unregulated rapid advancement of AI supported sophisticated targeting and recognition systems. Mehmet Akif Uzer in an article from September 2024 has suggested four essential components for an international regulation, based on ethical principles:

"To address the challenges posed by AI in warfare, we propose the following recommendations:

1) Prohibit Autonomous Weapon Systems Without Human Control: Weapon systems that do not allow for sufficient human oversight in target selection and engagement should be prohibited.
2) Establish Positive Obligations for Human Control: For systems that are not prohibited, establish clear obligations for human control over weapon parameters (e.g., type of target), the environment of use, and human-machine interaction during use.
3) Ensure Human Command and Control: Any use of weapon systems with autonomous functionalities must be guided and overseen by a responsible chain of human command and control.
4) Preserve Human Judgment in the Use of Force: Actions that may result in the loss of human life through the use of force should remain under human intent and judgment. Once a human initiates a sequence of actions intended to end with lethal force, autonomous systems may complete the sequence only with ongoing human oversight."[34]

There are several attempts to formulate internationally binding ethical and political principles for the responsible use of AI systems, such as from the European Commission on "Ethics Guidelines for Trustworthy AI " from 2019.

There is a list formulated of seven key principles of trustworthy AI ((1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability) which are important in this regard, but there is no distinct chapter on ethical implications of AI usages in military warfare, except that a concern at the end was raised under "Examples of critical concerns raised by AI" to investigate this further as the some critical ethical questions have begun to be articulated:

"Currently, an unknown number of countries and industries are researching and developing lethal autonomous weapon systems, ranging from missiles capable of selective targeting to learning machines with cognitive skills to decide whom, when and where to fight without human intervention. This raises fundamental ethical concerns, such as the fact that it could lead to an uncontrollable arms race on a historically unprecedented level, and create military contexts in which human control is almost entirely relinquished and the risks of malfunction are not addressed. The European Parliament has called for the urgent development of a common, legally binding position addressing ethical and legal questions of human control, oversight, accountability and implementation of international human rights law, international humanitarian law and military strategies. Recalling the European Union's aim to promote peace as enshrined in Article 3

[32] See: Expert Consultation report – Artificial intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed conflicts: Current Developments and Potential Implications 2024, Geneva 2022, in: https://shop.icrc.org/expert-consultation-report-artificial-intelligence-and-related-technologies-in-military-decision-making-on-the-use-of-force-in-armed-conflicts-current-developments-and-potential-implications-pdf-en.html

[33] See the series of blog papers of ICRC on „Artificial intelligence in military decision making", in: https://blogs.icrc.org/law-and-policy/category/special-themes/artificial-intelligence-in-military-decision-making/

[34] Mehmet Akif Uzer : The Integration of AI in Modern Warfare: Ethical, Legal, and Practical Implications, 24 September 2024, in: https://www.cyis.org/post/the-integration-of-ai-in-modern-warfare-ethical-legal-and-practical-implications#:~:text=From%20a%20humanitarian%20perspective%2C%20the,to%20IHL%20rules%20governing%20hostilities.

of the Treaty of the European Union, we stand with, and look to support, the Parliament's resolution of 12 September 2018 and all related efforts on LAWS."[35]

Equally in the key 44 pages document of UNESCO on ethics of artificial intelligence from 2022 "Recommendation on the Ethics of Artificial Intelligence"[36] amongst the 11 key priority areas for policy action for member states one cannot find any section dealing with military use of AI programs. However, in the "Pact for the Future" which was decided by the UN assembly in 2024 in the section dealing with "new and emerging technologies" (Action 28) in the chapter on "International Peace and Security" a key-recommendation is given which needs to be closely monitored in terms of its follow up and implementation:

Para 46 b) reads: "Advance with urgency discussions on lethal autonomous weapons systems through the existing intergovernmental process to develop an instrument, and other possible measures, including to address the risks posed by lethal autonomous weapons systems that select targets and apply force without human control or oversight and cannot be used in compliance with international humanitarian law."[37]

One of more promising examples of how principles could be formulated or need to be explored in extension of what already exists comes from the United Nations Office for Disarmament Affairs (UNODA) which since some years had appointed a working group following up the 2019 Convention on Certain Conventional Weapons on (CCW – Group of Governmental Experts on Lethal Autonomous Weapons Systems) which has followed up on issues related to Lethal Autonomous Weapons Systems (LAWS).[38] LAWS are "weapons which have the potential to identify, engage and neutralize a target without any human intervention. While such systems do not exist, the complete autonomy of these weapons may pose a multitude of issues from moral, legal and operational perspectives".[39] The UN declaration on LAWS from 2019[40], which contains some general clauses of caution and prohibition on lethal autonomous weapons, is part of and in continuation of the principles on certain conventional weapons (CCW) which dates back already to 1983, then amended in 2001, and deals with substantive prohibitions and restrictions on certain types of weapons, such as non-detectable fragments, mines booby traps and other devices, incendiary weapons, blinding laser weapons, explosive remnants of war.[41]

While this can indicate some of the institutional spaces and affiliated UN organizations to work, the conclusion of this brief first survey on recent developments can only repeat what transpired and has been demanded for by other specialists already earlier: **"We Need Hard Laws on the Military Uses of AI — and Soon".**[42] Voices abound that committed agencies together with EU should spearhead an initiative for international regulation and standards for using AI in warfare as the lack of an international governance framework for military AI poses risks to global security: It should be listened to more widely what leading experts from Carnegie Europe's research project on 'The EU's Techno-Politics of AI' have stated quite succinctly:

"The absence of a comprehensive global governance framework for military artificial intelligence (AI) presents a perilous regulatory void. This gap leaves a powerful technology category unchecked, heightening risks to international peace and security, escalating arms proliferation, and challenging international law. Governments worldwide are competing for leadership in emerging and disruptive technologies (EDTs) and grappling with the profound and transformative implications of AI. Meanwhile, corporate tech players have joined a trillion-dollar arms race in generative AI, jockeying for venture capital investment in foundation models. In the battle for economic supremacy and the competition over

---

[35] High-Level Expert Group on Artificial Intelligence, European Commission 2019: ETHICS GUIDELINES FOR TRUSTWORTHY AI https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf, page 34

[36] Recommendation on the Ethics of Artificial Intelligence, UNESCO, Adopted on 23 November 2021, in: https://unesdoc.unesco.org/ark:/48223/pf0000381137

[37] Action 28, No b), page 16, in: https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf

[38] See: https://libraryresources.unog.ch/c.php?g=462684&p=5054227

[39] https://www.diplomatie.gouv.fr/en/french-foreign-policy/france-and-the-united-nations/multilateralism-a-principle-of-action-for-france/alliance-for-multilateralism/article/11-principles-on-lethal-autonomous-weapons-systems-laws#:~:text=LAWS%20are%20weapons%20which%20have,moral%2C%20legal%20and%20operational%20perspectives.

[40] Annex III: Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, 2019, page 10. in: https://documents.un.org/doc/undoc/gen/g19/343/64/pdf/g1934364.pdf

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

[42] Branka Marijan: We Need Hard Laws on the Military Uses of AI — and Soon. Notably missing from the discussion are Russia and China. In: Centre for International Governance Innovation November 20,2023. In: https://www.cigionline.org/articles/we-need-hard-laws-on-the-military-uses-of-ai-and-soon/

ethical standards, the global balance of power is precarious and the stakes are high… There is therefore a sense of urgency among international organizations, scientists, and researchers, prompted by the potential of runaway AI developments, including disruptive applications in the military domain. If indeed AI poses an extinction-level existential threat to the future of humankind akin to the atomic bomb, as many in the field claim, the absence of a universally accepted global governance framework for military AI is a crucial concern."[43]

The Centre for AI Safety (CAIS) in San Francisco also clearly had stated the major risk of a Military AI Global Arms Race:

"The rapid advancement of AI in military technology could trigger a "third revolution in warfare," potentially leading to more destructive conflicts, accidental use, and misuse by malicious actors. This shift in warfare, where AI assumes command and control roles, could escalate conflicts to an existential scale and impact global security. Lethal autonomous weapons are AI-driven systems capable of identifying and executing targets without human intervention. These are not science fiction. In 2020, a Kargu 2 drone in Libya marked the first reported use of a lethal autonomous weapon. The following year, Israel used the first reported swarm of drones to locate, identify and attack militants. Lethal autonomous weapons could make war more likely. Leaders usually hesitate before sending troops into battle, but autonomous weapons allow for aggression without risking the lives of soldiers, thus facing less political backlash. Furthermore, these weapons can be mass-manufactured and deployed at scale… As AI accelerates the pace of war, it makes AI even more necessary to navigate the rapidly changing battlefield. This raises concerns over automated retaliation, which could escalate minor accidents into major wars. AI can also enable "flash wars," with rapid escalations driven by unexpected behaviour of automated systems, akin to the 2010 financial flash crash."[44]

It is a positive sign that under the leadership of Netherlands recently some governments (31 nations, unfortunately without Russia and China) have signalled some readiness and interest to push towards more international regulations on responsible military use of artificial intelligence and signed the "**Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy**". [45] This was announced at the summit on Responsible Artificial Intelligence in the Military Domain (REAIM) held in February 2023 in Den Haag brought together by the Foreign Ministry of the Netherlands,[46] which also produced its own public declaration – the first with governmental and military experts together.[47] In addition the UN conference on AI, Security and Ethics" in Geneva in March 2025[48] promises to become a major factor in bringing crucial people together interested in the governance of artificial intelligence (AI) in security and defence. This is under the leadership of UNIDIR (United Nations Institute for Disarmament Research), which has already published a first draft paper with key directions on a possible common statement and framework agreement[49] and should elaborate further what had been suggested as a first resolution on "Artificial intelligence in the military domain and its implications for international peace and security" in the UNGA in October 2024.[50] These initiative partly answers the passionate appeal of the General Secretary of the United Nations from 19 December 2024: "AI's expansion into security systems raises fundamental concerns about human right, dignity and the rule of law – from autonomous border surveillance to predictive policing and beyond".[51] In addition the United Nations Office for Disarmament Affairs (UNODA) in New York has published end of 2024 (at the date of Christmas 24 Dec!) the result of an essay competition

---

[43] Raluca Csernatoni: Governing Military AI Amid a Geopolitical Minefield. Carnegie Europe 17 July 2024, in: https://carnegieendowment.org/research/2024/07/governing-military-ai-amid-a-geopolitical-minefield?lang=en&center=europe

[44] See website of Centre for AI Safety, in: https://www.safe.ai/ai-risk; Dan Hendrycks, Mantas Mazeika, Thomas Woodside, 9 Oct 2023: An Overview of Catastrophic AI Risks, in: https://arxiv.org/pdf/2306.12001

[45] https://www.state.gov/wp-content/uploads/2023/10/Latest-Version-Political-Declaration-on-Responsible-Military-Use-of-AI-and-Autonomy.pdf

[46] Government of the Netherlands: About the REAIM conference February 2023: https://www.government.nl/ministries/ministry-of-foreign-affairs/activiteiten/reaim/about-reaim-2023#:~:text=REAIM%202023%20took%20place%20at,vulnerabilities%20associated%20with%20military%20AI.

[47] REAIM Call to Action 2023: https://www.government.nl/ministries/ministry-of-foreign-affairs/documents/publications/2023/02/16/reaim-2023-call-to-action

[48] https://unidir.org/event/global-conference-on-ai-security-and-ethics-2025/

[49] Policy Policy Paper on: „Governance of Artificial Intelligence in the Military Domain: A Multi-Stakeholder Perspective on Priority Areas", in: https://unidir.org/publication/governance-of-artificial-intelligence-in-the-military-domain-a-multi-stakeholder-perspective-on-priority-areas/

[50] See: UNGA First Committee Agenda item 98. Resolution: Artificial intelligence in the military domain and its implications for international peace and security in: https://docs.un.org/en/A/C.1/79/L.43

[51] See: General Secretary's remarks to the Security Council on Artificial Intelligence, 19 Dec 2024, in: https://www.un.org/sg/en/content/sg/statement/2024-12-19/secretary-generals-remarks-the-security-council-artificial-intelligence-bilingual-delivered

on AI military ethics with young people concerned about the rapid speed of the weaponization of AI. This was the result of a "Republic of Korea-United Nations Sci-fAI Futures Youth Challenge" program in the same year which was designed to engage young minds in imagining the future role of artificial intelligence (AI) in international peace and security.[52] The stories of the young authors which were produced in a setting of a fictional world in the year 2145 are quite substantial reflections about how AI could be used in or to prevent armed conflicts, its potential interactions with other emerging technologies, potential risks and positive scenarios in which AI is leveraged for good in the military domain.[53] The meaningful collection is one promising example to engage also the younger generations in AI ethics which have to live and cope with the consequences of the rapid integration of AI software into the military arms race – and this already much before 2145!

[52] See in: https://disarmament.unoda.org/unoda-occasional-papers-no-43-december-2024
[53] See UNODA publication: "Youth and Peace in the Age of AI" in: https://publications.unoda.org/documents/4/OP43-web.pdf