

**Impulspapier des Arbeitskreises KI &
Verteidigung**

**Beitrag zu Grundlagen und
Grenzen Künstlicher Intelligenz
(KI) beim Einsatz in
Verteidigungstechnologien**



Juni 2023

Co-Autoren

Dr. Hans Christoph Atzpodien, BDSV

Dr. Jürgen Bestle, Hensoldt

Jörg Eschweiler, Eviden Deutschland

Klaus Kappen, Rheinmetall

Dr. Florian Keisinger, Airbus Defence & Space

Prof. Dr. Wolfgang Koch, Fraunhofer FKIE

Dr. Stefan Mück, IBM Deutschland

Wolfgang Niedermark, BDI

PD Dr. Frank Sauer, UniBw München

Dr. Tassilo V. Singer, Capgemini Deutschland

Stephan Ursuleac, Bitkom

Dr. Charlotte Weil von der Ahe, Helsing

Impressum: Beitrag zu Grundlagen und Grenzen Künstlicher Intelligenz (KI) beim Einsatz in Verteidigungstechnologien

Herausgeber: BDSV, BDI, BDLI, Bitkom

Titelbild: Bildmontage © BDSV e.V.: iStock® monsiti

Inhaltsverzeichnis

I.	Einleitung.....	2
II.	Grundlagen und Begriffsklärungen	3
III.	Rechtliche und ethische Determinanten militärischer KI-Nutzung.....	7
IV.	Aspekte eines ethisch verantwortbaren Umgangs mit militärischer KI	11
V.	„Verantwortbarkeit“ und „Meaningful Human Control“ im Konzept-Vergleich.....	15
VI.	Folgerungen.....	17
VII.	Fazit.....	18

I. Einleitung

Das vorliegende Impulspapier soll einen praxisbezogenen Beitrag zum Diskurs über Stellenwert und Grenzen des Einsatzes Künstlicher Intelligenz (KI, siehe II.2) im Bereich militärischer Technologien leisten. Es bewegt sich hierbei entlang der Kategorien rechtlicher Möglichkeiten, operativer Notwendigkeiten und ethischer Verantwortbarkeit, immer getragen von dem Blick auf den Soldatinnen und Soldaten, der sich angesichts aller denkbaren Angriffsszenarien stets bestmöglich verteidigen und den militärischen Auftrag ausführen können sollen. Autoren dieses Impulspapiers sind die hier aufgeführten Personen, die eine verbandliche, industrielle oder wissenschaftliche Basis haben und sich im Arbeitskreis KI & Verteidigung unter Moderation des Bundesverbands der Deutschen Industrie e.V. (BDI) sowie des Bundesverbands der Deutschen Sicherheits- und Verteidigungsindustrie e.V. (BDSV) zusammengefunden haben.¹ Adressaten des Impulspapiers sind Akteure aus Regierung, Parlament, Medien und Interessenvertreter von Nichtregierungsorganisationen. Als solches kann und soll das vorliegende Dokument Impulse und konkreten Beitrag zur Formulierung einer nationalen militärischen KI-Strategie liefern, die angesichts der aktuellen Technologieentwicklung einerseits und der Bedrohungslage für Landes- und Bündnisverteidigung andererseits unerlässlich erscheint.

¹ Der Arbeitskreis wurde im Jahr 2020 ins Leben gerufen, und zwar unter Beteiligung der Verbände BDI, BDSV, BDLI und Bitkom.

II. Grundlagen und Begriffsklärungen

1. Im Eckpunktepapier „Digitalpolitik der Bundesregierung: Neuordnung digitalpolitischer Zuständigkeiten“ vom 31.08.2022² kommt das Bundesministerium der Verteidigung (BMVg) nicht vor. Gleichwohl kann kein Zweifel daran bestehen, dass gerade im Bereich der militärischen Verteidigungstechnologien KI von höchster Bedeutung ist, wie sowohl praktische Erfahrungen aus dem Ukraine-Krieg³ als auch die weitreichende Literatur hierzu⁴ belegen.

Auf Deutschland bezogen ergibt sich der Verteidigungsauftrag der Bundeswehr aus Art. 87a Grundgesetz (GG), bezogen auf Deutschlands Rolle im NATO-Bündnis aus der Präambel des Nordatlantikvertrages vom 04.04.1949.⁵ Daraus leitet sich der Auftrag der Bundeswehr ab, der lautet:⁶ „Die Bundeswehr verteidigt Deutschlands Souveränität sowie das Staatsgebiet und schützt seine Bürger. Sie trägt außerdem dazu bei Staat und Gesellschaft widerstandsfähig gegen äußere Bedrohungen zu halten und sicherheitspolitische Handlungsfähigkeit zu sichern. Die Bundeswehr

² Abrufbar unter:

https://www.bundesregierung.de/resource/blob/975226/2080044/6831e8a530126ca480340e1ce807ae18/20_21-08-31-eckpunktepapier-digitalpolitik-data.pdf?download=1

³ Siehe <https://www.nzz.ch/international/ukraine-krieg-kuenstliche-intelligenz-am-schlachtfeld-ld.1695576>;

Sandra Lumetsberger, Künstliche Intelligenz auf dem Schlachtfeld: Eine US-Software revolutioniert den Artillerie-Krieg in der Ukraine, Tagesspiegel vom 18.01.2023, abrufbar unter:

<https://www.tagesspiegel.de/internationales/kuenstliche-intelligenz-an-der-front-kann-moderne-software-den-krieg-grundlegend-verandern-9160431.html>

⁴ Dazu beispielhaft mit weiteren Nachweisen die Beiträge in: Lammert, Norbert, Koch, Wolfgang (Hrsg.): Bundeswehr der Zukunft – Verantwortung und Künstliche Intelligenz, Konrad-Adenauer-Stiftung, Berlin 2023; im Übrigen auch Bundesverband der Deutschen Industrie e.V.: Künstliche Intelligenz in Sicherheit und Verteidigung – Handlungsempfehlungen der deutschen Industrie vom 15.01.2019, abrufbar unter: <https://bdi.eu/publikation/news/kuenstliche-intelligenz-in-sicherheit-und-verteidigung/>

⁵ Abrufbar unter: https://www.nato.int/cps/en/natohq/official_texts_17120.htm?selectedLocale=de

⁶ Abrufbar unter:

<https://www.bundeswehr.de/resource/blob/5102304/c78eb8cab206e9b3e5f024c6c9fc1fc7/auftrag-und-aufgaben-der-bundeswehr-data.pdf>

hat weiterhin den Auftrag, Deutschlands Verbündete zu schützen, sowie sicherheitspolitische Bedrohungen für deren offene Gesellschaften abzuwehren und in Kooperation zu freiem und sicherem Welthandel beizutragen. Sie stärkt die europäische Integration, die transatlantische Partnerschaft und die Zusammenarbeit zwischen Staaten weltweit.“ Dieser Auftrag unterstellt und setzt voraus, dass die Bundeswehr die nötigen Fähigkeiten und die erforderliche Legitimation besitzt, um den ihr vorgegebenen Auftrag erfolgreich ausführen zu können.

2. Eine allgemeingültige Begriffsbestimmung von KI existiert nicht. Die 2018 durch das damalige Bundeskabinett angenommene „Strategie Künstliche Intelligenz“ (KI-Strategie)⁷ führt dazu aus: „Sehr abstrakt ordnen sich KI-Forscher zwei Richtungen zu: der ‚schwachen‘ und der ‚starken‘ KI. Die ‚starke‘ KI formuliert, dass KI-Systeme die gleichen intellektuellen Fertigkeiten wie der Mensch haben oder ihn darin sogar übertreffen können. Ob es diese geben kann ist fraglich. Die ‚schwache‘ KI ist fokussiert auf die Lösung konkreter Anwendungsprobleme auf Basis der Methoden aus der Mathematik und Informatik, wobei die entwickelten Systeme zum Selbstlernen fähig sind. Dazu werden auch Aspekte menschlicher Intelligenz nachgebildet und formal beschrieben bzw. Systeme zur Simulation und Unterstützung menschlichen Denkens konstruiert. Die Bundesregierung orientiert sich bei ihrer Strategie an der Nutzung der KI für die Lösung von Anwendungsproblemen und damit an den Positionen der ‚schwachen‘ KI.“ Im militärischen Bereich dienen einige Anwendungen der Unterstützung des Menschen bei der

⁷ Abrufbar unter: <https://www.ki-strategie-deutschland.de/home.html>

Lagebildgewinnung oder der Analyse von Entscheidungsoptionen und Vorbereitung des Einsatzes von Wirkmitteln. Es kommt die KI-basierte Umsetzung menschlicher Entscheidungen in automatisierte ausgeführte Befehlsfolgen hinzu. ‚Starke‘ KI wird in diesem Dokument nicht betrachtet.

3. Die uneingeschränkte Handlungsverantwortung liegt beim Menschen,⁸ der sie im Verteidigungsfall entsprechend den dafür maßgeblichen völkerrechtlichen Rahmenbedingungen auszuüben hat. Diese völkerrechtlichen Rahmenbedingungen umfassen sowohl objektiv bestimmbare Restriktionen (wie insbesondere die Unterscheidung zwischen Kombattanten und Zivilisten), aber auch subjektiv wertende Kategorien (wie etwa das Übermaßverbot).⁹ Soldatische Ethik und Einsatzrecht im Kontext militärisch genutzter KI muss insbesondere den jeweiligen „Rules of Engagement“ (RoE) genügen, d.h. gewissen „Hard-Constraints“, die zu befolgen sind. Verantwortung zum Einsatz von KI in Waffensystemen fängt jedoch nicht beim ausführenden Menschen (in dem Fall Soldatin oder Soldaten) an,¹⁰ sondern hört dort auf. Bedarfsträger im Planungsprozess, Wissenschaftler und Ingenieure bei Forschung und Entwicklung, Bedarfsdecker und Politiker, die beschaffen und mandatieren, sowie die Planer und Ausführer militärischen Operationen haben jeweils ihren eigenen Anteil an der Verantwortung, die vielleicht wie im Versicherungswesen geteilt wird.

⁸ Dieser Grundsatz ist in mehreren Bundeswehr-Grundsatz-Dokumenten verankert.

⁹ Singer, Tassilo: KI im operativen Kontext: Zur technischen Übertragbarkeit von Regeln des humanitären Völkerrechts auf Künstliche Intelligenz, in: NJW Sonderausgabe 2022: Künstlich-intelligente Maschinen, S. 7 ff

¹⁰ Vgl. Art. 86 (2), 87 (1) ZP I

4. Das vorliegende Dokument konzentriert sich daher auf die schwieriger zu beurteilenden Fälle des Einsatzes von KI im Rahmen teil- oder vollautomatisierter Waffensysteme, die dem auf dem Gefechtsfeld agierenden Menschen mehr als eine bloße Entscheidungshilfe zu verschaffen vermögen. Laut Sauer¹¹ geht es hierbei um die Beschreibung des Verhältnisses von Mensch und Maschine, denn Teil- oder Vollautomation bei Waffensystemen stellen nichts anderes als graduelle Aufgabenverschiebungen in der Mensch-Maschine-Relation aufgrund entsprechend programmierter Algorithmen dar. Es geht um die Frage, was auf einem zukünftigen Gefechtsfeld der Mensch allein entscheiden soll bzw. muss, und was zum Teil oder ganz die Maschine übernehmen kann und soll. Hiermit werden zwei Fragen aufgeworfen: 1.) in welchen Bereichen ist die Maschine leistungsfähiger als der Mensch, 2.) welcher menschlichen Kontrolle die Maschine in dem ihr übertragenen Rahmen bedarf („Verantwortbarkeit“ bzw. „Meaningful Human Control“) und 3) wie die Konstellationen der Zusammenarbeit zwischen Mensch und Maschine aussehen.¹²

5. Der Koalitionsvertrag „Mehr Fortschritt wagen“ vom 07.12.2021 geht zwar nicht gesondert auf die Frage der Bedeutung von KI für militärische Verteidigungstechnologien ein, setzt aber den Rahmen für das grundlegende Verständnis der aktuellen Bundesregierung. Hierzu heißt es: „Letale autonome Waffensysteme, die vollständig der

¹¹ Sauer, Frank: Drei Thesen zur nationalen Regulierung von Autonomie in Waffensystemen, in: Lammert, Norbert, Koch, Wolfgang (Hrsg.), Bundeswehr der Zukunft – Verantwortung und Künstliche Intelligenz, Konrad-Adenauer-Stiftung, Berlin 2023

¹² Quelle: H. W. Meerveld · R. H. A. Lindelauf · E. O. Postma · M. Postma, „The irresponsibility of not using AI in the military“, Ethics and Information Technology (2023) 25:14, abrufbar unter: <https://doi.org/10.1007/s10676-023-09683-0>

Verfügung des Menschen entzogen sind, lehnen wir ab. Deren internationale Ächtung treiben wir aktiv voran“.¹³

6. Europäische und NATO-Dimensionen: Auch, wenn in Deutschland Überlegungen zur verantwortlichen Nutzung digitaler Wehrtechnik stets von der Identität der Bundeswehr ausgehen müssen, bedarf es international geführter Diskurse, zum Beispiel im Kontext der EU und NATO. Aber wo stehen andere Länder bei der Nutzung von KI im Militär? Teilweise sind sie in den verschiedenen Erbringungsdimensionen weiter als Deutschland, was mit Blick auf die Interoperabilität digitaler Systeme im Bündnis und auf die Anschlussfähigkeit der Führungs-philosophien der Partnernationen zum Problem werden kann.

III. Rechtliche und ethische Determinanten militärischer KI-Nutzung

1. Völkerrecht: Ohne die Diskussion zu vertiefen, sei betont, dass das humanitäre Völkerrecht insbesondere in den Artikeln 35 und 36 des Zusatzprotokolls I zur Genfer Konvention Regeln zur Wahl von Mitteln und Methoden der Kriegführung aufstellt.¹⁴ Hierbei werden einerseits Regeln umfasst, die von automatisierten Systemen umgesetzt werden können, wie etwa die Unterscheidung zwischen zulässigen militärischen

¹³ Koalitionsvertrag zwischen SPD, Bündnis90/DIE GRÜNEN und FDP: Mehr Fortschritt wagen; abrufbar unter: <https://www.bundesregierung.de/breg-de/service/gesetzesvorhaben/koalitionsvertrag-2021-1990800>

¹⁴ Zusatzprotokoll zu den Genfer Abkommen vom 12. August 1949 über den Schutz der Opfer internationaler bewaffneter Konflikte (Protokoll I) vom 8.6.1977, abrufbar unter: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/1982/1362_1362_1362/20180712/de/pdf-a/fedlex-data-admin-ch-eli-cc-1982-1362_1362_1362-20180712-de-pdf-a.pdf; im Übrigen Singer, ebda.

und nicht zulässigen zivilen Zielobjekten. Zum anderen verlangt das völkerrechtliche Regelwerk aber auch wertende bzw. abwägende Entscheidungen, wie etwa bei dem in Art. 57 Abs. 2 ZP I enthaltenen Exzessverbot, das die mil. Entscheidungsträger zum Abbruch eines Angriffs verpflichtet, wenn „damit zu rechnen ist, dass er auch Verluste unter der Zivilbevölkerung, die Verwundung von Zivilpersonen, die Beschädigung ziviler Objekte oder mehrere derartige Folgen zusammen verursacht, die in keinem Verhältnis zum erwarteten konkreten und unmittelbaren militärischen Vorteil stehen“. Angesichts der fortschreitenden KI-Technikentwicklung sind daher vor allem im Bereich der abwägenden Entscheidungsunterstützung immer wieder Äquivalenztests durchzuführen, die die entsprechenden KI-Fähigkeiten vor dem Hintergrund der völkerrechtlichen Anforderungen zu validieren vermögen.¹⁵ Darüber hinaus verpflichtet Artikel 36 die Staaten, alle neuen Waffen, Mittel und Methoden der Kriegsführung rechtlich zu prüfen, um festzustellen, ob ihr Einsatz durch das Völkerrecht verboten ist. Überlegungen zur Technikverantwortung sind demnach aus völkerrechtlicher Sicht gefordert.¹⁶

2. Überlebensfähigkeit des Nutzers: Ein Aspekt kommt in der Diskussion über militärische KI bisweilen immer noch zu kurz: Sauer¹⁷ wies bereits 2019 darauf hin, dass „im Sinne des Eigenschutzvorrangs natürlich autonome Verteidigung [...] legitim“ sein könne. Denn der

¹⁵ Singer, ebda., S. 9

¹⁶ V. Boulanin, „Implementing Article 36 weapon reviews in the light of increasing autonomy in weapon systems,“ in SIPRI Insights on Peace and Security. Solna, Sweden: SIPRI, No. 2015/1, Nov. 2015. [Online]. Available: <https://www.sipri.org/sites/default/files/files/insight/SIPRIInsight1501.pdf>

¹⁷ Sauer, Frank 2019: Mensch, Maschine, Krieg, in: zur sache bw 35, S. 24-29, abrufbar unter: <https://www.bundeswehr.de/resource/blob/179466/d8187a6e5d0b3150ada3aaae3e0530f9/zur-sache-bw-35-1-2019-data.pdf>.

sowohl naturrechtlich¹⁸ als auch in gängigen Strafrechtsordnungen stets anerkannte Grundsatz des Selbstverteidigungsrechts („Notwehr“) oder der erweiterten Selbstverteidigung („Nothilfe“) gilt natürlich auch für Kombattanten im Zustand einer kriegerischen Auseinandersetzung.¹⁹ Im Übrigen gilt der Grundsatz der Selbstverteidigung auch auf der kollektiven völkerrechtlichen Ebene: Wer als Staat angegriffen wird, darf sich verteidigen.²⁰ Im Fall der militärischen KI-Nutzung heißt dies (zitiert aus FAZ vom 31.01.2022): „Zu letzteren gehört, dass Deutschland und Europa technologisch in der Lage sein müssen, Angriffe mit KI-unterstützten Waffensystemen abzuwehren. Der generelle Ausschluss der KI hätte eine unverantwortbare strukturelle militärische Unterlegenheit zur Folge und würde operativ die Gefahr von Kollateralschäden etwa aufgrund mangelnder Präzision in der Zielerfassung erhöhen. Der militärische Einsatz von KI ist vor diesem Hintergrund nicht nur ethisch gerechtfertigt, sondern sogar geboten.“²¹ Wie noch auszuführen sein wird, betrifft die Determinante der

¹⁸ Erb, Volker: Notwehr, in: Staatslexikon, abrufbar unter: <https://www.staatslexikon-online.de/Lexikon/Notwehr>; Kopel, David B.: Das natürliche Recht auf Selbstverteidigung – Was die Welt aus dem Heller-Urteil lernen kann, abrufbar unter: <https://davekopel.org/Deutsch/Kopel%20Das%20nat%20C3%BCrlische%20Recht%20auf%20Selbstverteidigung.pdf>

¹⁹ Vgl. dazu Art. 31 (1) (d) Rom Statut, wonach die strafrechtliche Verantwortlichkeit zu verneinen ist.

²⁰ Artikel 51 der UN-Charta: „Diese Charta beeinträchtigt im Falle eines bewaffneten Angriffs gegen ein Mitglied der Vereinten Nationen keineswegs das naturgegebene Recht zur individuellen oder kollektiven Selbstverteidigung, bis der Sicherheitsrat die zur Wahrung des Weltfriedens und der internationalen Sicherheit erforderlichen Maßnahmen getroffen hat. Maßnahmen, die ein Mitglied in Ausübung dieses Selbstverteidigungsrechts trifft, sind dem Sicherheitsrat sofort anzuzeigen; sie berühren in keiner Weise dessen auf dieser Charta beruhende Befugnis und Pflicht, jederzeit die Maßnahmen zu treffen, die er zur Wahrung oder Wiederherstellung des Weltfriedens und der internationalen Sicherheit für erforderlich hält“; s. auch Art. 42 (7) EUV

²¹ Bossong, Nora, Rieks, Ansgar, Koch, Wolfgang: Künstliche Intelligenz für die Landesverteidigung, in: FAZ vom 31.01.2022; ein weiteres Argument kann auch in Art. 57 (2) lit. a i) ZP I gefunden werden. Demnach müssen alle praktisch Möglichen Maßnahmen ergriffen werden, um zivile Kollateralschäden zu vermeiden bzw zivile Objekte und Zivilpersonen zu schützen (Feasibility Standard). Dies kann man auch so verstehen, dass zur Verfügung stehende bessere Sensorik/Sensor / Datenverarbeitung (durch KI) auch einzusetzen ist – wenn der entsprechende Effekt dann in einem besseren Schutzniveau resultiert. Andere sehen dies aber streng und lehnen dies ab unter Verweis, dass dieser Standard nur auf die im konkreten Fall zur Verfügung stehenden Mittel beschränkt ist (und damit der Gebotenheit hier nicht zwingend zuträglich ist).

Überlebensfähigkeit des Nutzers aber nicht nur die Frage des Einsatzes KI-unterstützter Waffensysteme im Allgemeinen, sondern auch die konkreten Anwendungsparameter im jeweiligen Einzelfall, festgemacht an den Kategorien der ethischen „Verantwortbarkeit“ bzw. der „Meaningful Human Control“.

3. „Verantwortbarkeit“ und „Meaningful Human Control“ des KI-Einsatzes: Während es sich bei dem Konzept der „Meaningful Human Control“ um ein vor allem im anglo-amerikanischen Raum entwickeltes Gedankenmodell handelt, ist das Konzept der „Verantwortbarkeit“ als Maßstab für das „Einhegen“ von handlungsbezogenen oder technischen Katastrophenrisiken im deutschen Ethik-Denken vielfach verankert, u.a. bei der „inneren Führung“ der Bundeswehr²² oder im Bereich der Umweltethik bei der Katastrophen- Vermeidung im Falle ziviler Kernkraft-Nutzung.²³ Eine andere, wiederum international gebräuchliche Ausprägung des Konzeptes der Technik- Verantwortbarkeit findet sich in Form des „Fail-safe“-Prinzips im Engineering: “A design feature or practice that, in the event of a specific type of failure, inherently responds in a way that will cause minimal or no harm to other equipment, to the environment or to people”.²⁴

4. Bezogen auf KI-Anwendungen bei Waffensystemen ist „Verantwortbarkeit“ als Konzept fundamentaler als etwa die Forderung

²² Koch, Wolfgang: Verantwortbarkeit als technisches Designprinzip künstlich intelligenter Maschinen, in: NJW Sonderausgabe 2022: Künstlich-intelligente Maschinen, S. 14 ff.

²³ Vgl. u.a. Hülsmann, Heinz, Tschiedel, Robert: Kernenergie und wissenschaftliche Verantwortung, Athenaeum- Verlag, Bodenheim 1977; Altner, Günter: Ist Kernenergie ethisch verantwortbar?, abrufbar unter: https://link.springer.com/chapter/10.1007/978-3-642-93475-9_29; Zeit-Artikel „Al-Wazir: Weiterbetrieb von Atomkraftwerken ‚verantwortbar‘“ vom 28.09.2022, abrufbar unter: https://www.zeit.de/news/2022-09/28/al-wazir-weiterbetrieb-von-atomkraftwerken-verantwortbar?utm_referrer=https%3A%2F%2Fwww.google.com

²⁴ Vgl. Wikipedia zu „Fail-safe“, abrufbar unter: <https://en.wikipedia.org/wiki/Fail-safe>

nach einem „Human in the loop“ oder „Human on the loop“.²⁵ Denn auch Automation von Waffensystemen kann verantwortbar sein, wenn Reaktionszeiten für Menschen zu kurz oder die Datenfülle zu groß sind, um für sie einen Fail-safe-Zustand zu gewährleisten. Künstlich intelligente Maschinen müssen also gerade dann technische Beherrschbarkeit und verantwortungsvollen Einsatz gewährleisten, wenn es um die Einhaltung des „fail-safe“/„no harm“-Zieles für Mensch und Umwelt geht. Auch dort muss der Mensch eingebunden sein, nicht nur durch die Entscheidung, derartige Maschinen zu nutzen, sondern sie so zu entwerfen, dass ihr Einsatz effektiv und zugleich verantwortbar bleibt. Hierfür ist im Einsatz Wissen zu vermitteln, durch das Menschen einerseits automatisierter Entscheidungsunterstützung vertrauen können und andererseits die Grenzen kennen, die jede derartige Assistenz besitzt. Denn die auf künstlich intelligenten Maschinen basierenden Erkenntnisse müssen erklärbar sein.²⁶ Adäquates Training ist demnach unabdingbare Voraussetzung, KI verantwortungsvoll einzusetzen.

IV. Aspekte eines ethisch verantwortbaren Umgangs mit militärischer KI²⁷

1. Ethische Verantwortbarkeit: Diese umfasst nach Koch²⁸ vier Kategorien von relevanten Handlungs-Aspekten der Person, die sich für

²⁵ Koch, ebda.

²⁶ Koch, ebda.

²⁷ oberhalb der Betrachtung bestimmter Anwendungsfälle („use cases“)

²⁸ Koch, ebda.

den Einsatz eines mittels KI operierenden militärischen Wirkmittels bedient: einer kognitiven, einer volitiven, einer normativen und einer reflexiven Kategorie. Alle vier Kategorien sind der Unterstützung durch KI-Algorithmen zugänglich. Im Einzelnen meint dies:

- a. Kognitive Aspekte umfassen das zutreffende Erkennen eines militärischen Lagebildes vom Identifizieren/Auffinden und Verfolgen eines Objektes bis zum Identifizieren als zu bekämpfendes Ziel. Diese Aufgabe kann KI als wichtigen Unterstützungsbeitrag bei richtigem Design belastbar übernehmen.²⁹
- b. Volitive Aspekte meinen die Abläufe einer rationalen/informierten Entscheidung zum Bekämpfen eines Ziels anhand vorgegebener Entscheidungskriterien, also die Identifizierung von Handlungsalternativen und die rationale Auswahl daraus. Auch dies ist eine typische Aufgabenstellung, bei der die Kombination verschiedener Technologien einen entscheidenden Beitrag zu einer zeitnahen, situations- und auftragsgerechten Entscheidung leisten kann. Das Spektrum der Möglichkeiten reicht hier von geeigneter Visualisierung, über die Implementierung von regelbasierten Entscheidungselementen bis hin zur Berücksichtigung von Optionen, die ein

²⁹ Oft würde z.B. bereits eine direkte Beobachtung den Menschen einer Gefahr aussetzen. Der Einsatz von Drohnen und anderen Möglichkeiten der Fernaufklärung basiert auf der Auswertung von massenhaften Daten in Echtzeit, womit ein Mensch überfordert wäre. Diese Datenströme auf die relevanten Informationen zu reduzieren ist ein wesentliches Einsatzfeld für KI.

Optimierungsalgorithmus auf Basis von Lagebild und Auftrag dem menschlichen Entscheider vorlegt.

- c. Normative Aspekte stellen die Messlatte für die Entscheidung dar, beinhalten also vor allem die zu beachtenden völkerrechtlichen Auswahlkriterien (wie das Verbot der Bekämpfung ziviler Objekte oder die Einhaltung des Übermaßverbotes). Auch hier dienen KI-Algorithmen als Werkzeug zur richtigen, dabei jederzeit nachvollziehbaren Identifizierung von Entscheidungsalternativen, also insbesondere deren Aufbereitung.
- d. Reflektive Aspekte schließlich beinhalten die letzte Plausibilitätsüberprüfung vor der Umsetzung der Entscheidung, zur Wahrung oberster völkerrechtlicher Grundsätze (wie Menschenwürde), d.h. zur Einhaltung des Übermaßverbotes, zum Ausschluss einer Verwechslung des Zielobjekts, aber auch zur Herstellung des Gesamtkontextes etwa zur Geschwindigkeit des Angriffs und damit zur Eigen-Gefährdung des Entscheiders (in Relation zu seinen Entscheidungsspielräumen).

2. Verantwortbare Erweiterung bzw. Reduzierung des reflektiven Spielraums: Im Kalten Krieg gab es den berühmten, im Nachhinein bekannt gewordenen Fall, in dem der russische Oberstleutnant Stanislaw J. Petrow am 26.09.1983 als zuständiger Offizier in der Kommandozentrale der sowjetischen Satellitenüberwachung einen vom System gemeldeten Angriff der USA mit nuklearen

Interkontinentalraketen auf die UdSSR korrekt als Fehlalarm einstuft und durch sein Stoppen eines Gegenschlages einen Atomkrieg verhindert.³⁰ Dies ist ein Beispiel für die Erweiterung des reflexiven Spielraums gegenüber den Indikatoren und Entscheidungsempfehlungen, die das System selbst dem zuständigen „Man in/on the loop“ anzeigt, durchaus auch in dem Bewusstsein, dass es sich bei einem solchen Eingreifen im Nachhinein gesehen um eine Fehlentscheidung handeln kann. Verantwortbar bleibt es dennoch, insbesondere unter dem Blickwinkel des „fail-safe“/„no harm“-Zieles für Mensch und Umwelt.

3. Genauso kann es aber auch den umgekehrten Fall geben, in dem das System dem „Man in/on the loop“ anzeigt (oder angesichts der Angriffs-Geschwindigkeit nicht einmal mehr anzeigen kann), dass das übergeordnete Ziel des Selbstschutzes – „fail-safe“/„no harm“ – für ihn nur erreichbar ist, wenn er das System automatisiert die Zielbekämpfung ausführen lässt. Auch diese Variante muss nach der unter III. 2 genannten Ziel-Kategorisierung als ethisch verantwortbar gelten, denn niemand muss sich angesichts entsprechend vorhandener technischer Schutzmöglichkeit nur zur Wahrung des menschlichen Eingriffsprivilegs einer umso akuterer Gefahr für das eigene Leben aussetzen. Nicht-Nutzung erschiene hier gerade nicht verantwortbar.

³⁰ Hier wiedergegeben nach Wikipedia, abrufbar unter:
https://de.wikipedia.org/wiki/Stanislaw_Jewgrafowitsch_Petrow

V. „Verantwortbarkeit“ und „Meaningful Human Control“ im Konzept-Vergleich

1. „Meaningful Human Control“: Eine in der ethischen Diskussion um sog. letale autonome Waffen (international kurz „LAWS“ – oder im deutschen Umgangssprachgebrauch „Killerroboter“) gebräuchliche Kategorie zur Beschreibung voraussetzender menschlicher Kontrolle ist in der internationalen Literatur der Begriff der „Meaningful Human Control“³¹ oder in der Sprache der US-Regierung des „Appropriate Level of Human Judgement“.³² Im Anschluss an die Vorgaben des humanitären Völkerrechts gelten als Schlüsselemente einer „Meaningful Human Control“.³³
 - a. Kontext-Kontrolle über das Wirken der KI-unterstützten Waffe im Hinblick auf völkerrechtlich maßgebliche Rahmenbedingungen ihres Einsatzes;
 - b. Verständnis der Waffe im Hinblick auf ihre Technologie und des Prozesses ihres Einsatzes;
 - c. Verständnis des „operational Environment“, in dem die Waffe ihre Wirkung entfaltet;
 - d. Kontrolle im Sinne von „Predictability and Reliability“ des Waffensystems in dem Sinne, dass es den ihm aufgegebenen

³¹ Eklund, Amanda M.: Meaningful Human Control of Autonomous Weapon Systems, Swedish Defence Research Agency 2020, abrufbar unter: <https://www.fcas-forum.eu/publications/Meaningful-Human-Control-of-Autonomous-Weapon-Systems-Eklund.pdf>

³² United States (2018), Human-Machine Interaction in the Development, Deployment and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, working paper for the Group of Governmental Experts (2018-08-27–31), CCW/GGE.2/2018/WP.4, 2018-08-28.

³³ Eklund, ebda., S.29 ff.

menschlichen Willen erfüllt und aller Voraussicht nach nicht unkontrolliert von ihm abweicht;

- e. Menschliche Aufsicht und die Möglichkeit des beaufsichtigenden Menschen, in den Prozess von der Analyse bis zur Wirkung einzugreifen;
- f. Zurechenbarkeit der ausgelösten Wirkung zu einer menschlichen Verantwortung in der vom humanitären Völkerrecht gebotenen Form;
- g. Beachtung der menschlichen Würde als dem obersten Gebot alles ethischen Handelns.

2. Konzept der „Meaningful Human Control“ vs. Konzept der „Verantwortbarkeit“: Das Konzept der „Meaningful Human Control“ umfasst – wie das Konzept der „Verantwortbarkeit“ – die kognitiven, volitiven und normativen Aspekte der im Wesentlichen durch das humanitäre Völkerrecht bestimmten Vorgaben für den Einsatz von Waffen gegen Menschen.³⁴ Die Frage, ob zwischen beiden Konzepten ein wesentlicher Unterschied besteht, bezieht sich vor allem auf den reflektiven Bereich: „Meaningful Human Control“ verweist hier auf das Gebot der menschlichen Würde als obersten Grundsatz für jegliches – auch kriegerisches – Handeln. Das Konzept der „Verantwortbarkeit“ schließt dies im reflektiven Bereich ebenso ein, allerdings erweitert um den Aspekt des gebotenen Selbstschutzes des für die Wirkung der

³⁴ Eklund, ebda., S. 32 f.

Waffe verantwortlichen Akteurs. Dieser letztgenannte Aspekt jedoch erweist sich in seiner unveräußerlichen, schon aus dem Naturrecht abzuleitenden Gültigkeit als ein unverzichtbarer Dreh- und Angelpunkt der gesamten Diskussion über Kontrolle und Verantwortbarkeit von KI-unterstützten Waffensystemen.

VI. Folgerungen

Die wichtigste Erkenntnis aus den vorherigen Ausführungen ist, dass der „Human in/on the Loop“ für den Fall, dass er angegriffen wird, sein Leben (und ggfs. das anderer) nicht der Aufrechterhaltung seines Kontrollprinzips opfern muss. Wird sein Leben gerade dadurch bedroht, dass er sein menschliches Kontrollrecht durchhält, so kann er automatisierte Entscheidungsvorgänge der Maschine abrufen, um ihm nach Möglichkeit sein Leben zu erhalten. Eine entsprechend autorisierte Maschine sollte also unter Zeitdruck und unter den entsprechenden operativen Umständen Bekämpfungsmaßnahmen zur Verteidigung einleiten und erst dann den Menschen hinzuziehen können – sich allerdings natürlich nicht jederzeit und beliebig selbst dazu ermächtigen können. Dies ist eigentlich der entscheidende ethische Grundsatz im Mensch-Maschine-Verhältnis, der aus der überragenden Bedeutung des Selbsterhaltungs- und Selbstverteidigungsgrundsatzes abzuleiten ist. In einigen eng begrenzten Anwendungskontexten, etwa mit Blick auf die automatische Verteidigung gegen anfliegende Raketen oder Mörsergranaten, ist genau diese Form der Mensch-Maschine-Interaktion im Übrigen bereits seit Jahrzehnten in vielen Streitkräften

üblich. Man denke als nur ein Beispiel an das israelische System Iron Dome.³⁵

Auch hierbei mag es diffizile Konfliktfälle geben. So mag etwa der Fall vorstellbar sein, dass die Abgabe der Kontrolle an die Maschine zwar dem individuell kämpfenden Akteur das Leben rettet, während die Aufrechterhaltung der menschlichen Kontrolle zwar den individuellen Akteur das Leben kostet, dafür aber einer anderen Gruppe das Leben rettet. Solche Fälle erscheinen jedoch konstruiert und mögen der Ex-Post-Analytik einer KI-gestützten Maschine entspringen. Sie ändern jedoch nichts an der einfachen Tatsache, dass es von Rechts wegen keinem individuellen Menschen zugemutet wird, eine Entscheidung gegen sein eigenes Überleben zu treffen. Dieses muss daher auch bei der Programmierung der Maschine höchste Priorität haben.

VII. Fazit

Die öffentliche Diskussion über das Mensch-Maschine-Verhältnis bei KI-gestützten letalen Waffen muss präzise, differenzierend und auf sachlicher Ebene geführt werden. Sie muss dem Umstand Rechnung tragen, dass im Falle eines Falles im Rahmen der Verhältnismäßigkeit volle Unterstützung durch künstlich intelligente Automation ethisch erlaubt sein muss, wenn diese für das Überleben der betroffenen Individuen oder Kollektive in einer Verteidigungssituation erforderlich ist. Diese Überlegung folgt gleichsam der Logik einer

³⁵ Vorausgesetzt ist, dass menschliche Überwachung der Systeme stattfindet und die Bediener umfassend im Umgang mit den Systemen geschult werden; vgl. dazu <https://www.jstor.org/stable/resrep06103>

„Schiebelehre“, die das Mensch-Maschine-Verhältnis kontextspezifisch in einen verantwortungsbewussten Ausgleich bringt. Ein wichtiges Element zur Bestimmung der jeweils angemessenen „Position“ der „Schiebelehre“ ist die individuelle, aber auch kollektive Bedrohung und die Kontrollfrage, wie eine Nicht-Anwendung vorhandener Technologie zu bewerten wäre. Hinzu kommen Faktoren wie die Beschaffenheit des operativen Umfelds, etwa mit Blick auf die Wahrscheinlichkeit ziviler Opfer, oder die Zielprofile, auf die ein Waffensystem ausgelegt ist. Als konkretes Beispiel stelle man sich eine Fregatte auf hoher See vor, wo keine Kollateralschäden drohen und die Handhabung der Automation im Wirken gegen feindliche Zeile in Abwesenheit akuter Bedrohungen weniger restriktiv sein kann als in komplexeren und unübersichtlicheren Gefechtssituationen wie etwa im urbanen Gelände. Auch hier legen die unterschiedlichen Tactics, Techniques, and Procedures³⁶ im Umgang mit bereits seit Jahren existierenden Systemen – etwa Phalanx auf hoher See und Centurion an Land – den Schluss nahe, dass das Problem weder so neu noch militärisch so schwer lösbar ist, wie bisweilen suggeriert wird. Das Gut der Verteidigungs- und Überlebensfähigkeit bildet bei der Erarbeitung von Protokollen für den auch weiterhin verantwortungsbewussten Umgang mit Automation ein unveräußerliches Maß. Industrie und Forschung spielen eine entscheidende Rolle dabei, den aktuellen Stand der Technik in der KI (hinsichtlich Modellen und Architekturen) voranzutreiben, die richtigen Designentscheidungen in Abhängigkeit von Anforderungen und

³⁶ Sauer, Frank 2022: Autonomie in Waffensystemen: Wie gelingt Regulierung?, in: Tagesspiegel Background, <https://background.tagesspiegel.de/cybersecurity/autonomie-in-waffensystemen-wie-gelingt-regulierung>

neuesten Fähigkeiten zu treffen und in definierter Qualität vortrainierte Systeme zur weiteren Verbesserung an das Militär zu übergeben. Da die Algorithmik von KI-Systemen maßgeblich von Daten bestimmt wird und der Industrie nicht alle Möglichkeiten offenstehen, einsatznahe Daten zu sammeln, kommt dem Militär eine ebenso entscheidende Rolle dabei zu, die Qualität der Systeme zu prüfen und durch die Sammlung weiterer Daten und ein Nachtrainieren zu verbessern. Die Verwaltung und Politik schließlich müssen Rahmenbedingungen schaffen, dass eine kontinuierliche Überprüfung und Weiterentwicklung der Modelle möglich wird. Dem stehen aktuell noch die althergebrachten Beauftragungsmodelle entgegen.³⁷ Dies kann durch einen offenen Dialog zwischen Industrie, Forschung und Bundeswehr, die Einbeziehung von und den konstanten Austausch mit Operateuren in der Entwicklung und der Einhaltung höchster Standards an Sicherheit, Ethik und Recht gelingen.

³⁷ Die Herausforderung ist erkannt und soll im Kontext einer Software Defined Defense Strategie adressiert werden.



BDI

BDLI 

BDSV 

bitkom