Metis

Studie

Sicherheitspolitische Auswirkungen der Digitalisierung: Zukünftige Konfliktformen und Konfliktbearbeitung

Nr. 01 | Februar 2018

Institut für
Strategie & Vorausschau

Zusammenfassung

Maschinelle Autonomie ist der wichtigste Zukunftstrend in der fortschreitenden Digitalisierung. Mit Blick auf Waffensysteme verspricht Autonomie Kräftemultiplikation, höhere Operationsgeschwindigkeiten und präzisere Wirkung. Das Verdrängen menschlicher Verfügungsgewalt birgt jedoch operative Risiken kriegsvölkerrechtlicher und ethischer Natur in der Konfliktbearbeitung sowie strategische Risiken durch neue, eskalationsanfällige Konfliktformen. Die Studie empfiehlt daher für die Nutzung von Autonomie in Waffensystemen durch die Bundeswehr einen differenzierten Ansatz, um Vorteile für die Bundeswehr auszuschöpfen aber zugleich Risiken zu minimieren. Auf nationaler Ebene beinhaltet dieser Ansatz das Erstellen eines Richtliniendokuments, um, außer in Verteidigungssystemen, Auswahl und Bekämpfung von Zielen stets menschlicher Verfügungsgewalt zu unterstellen. Auf internationaler Ebene umfasst dieser das Bemühen um internationale, völkerrechtlich bindende, verifizierbare Regulierung.

Themenaufriss: Digitalisierung und Autonomie

Mit dem Fortschreiten von Informationszeitalter und Digitalisierung wachsen Umfang und Komplexität der Aufgaben, die vom Menschen an Computer und Maschinen delegiert werden. Robotik sowie insbesondere Künstliche Intelligenz (KI) sind gegenwärtig die Schlüsseltechnologien in diesem Prozess. Die zunehmende Bedeutung von Algorithmen und maschineller „Autonomie" sind in Form von Apples digitaler Assistentin Siri oder Teslas Fahrassistenzsystem Autopilot inzwischen ins Bewusstsein der breiten Öffentlichkeit gerückt.[1]

Diese Entwicklung wird gleichzeitig über- und unterschätzt. Überschätzt, weil der Begriff der künstlichen „Intelligenz" falsche Assoziationen weckt.[2] So ist Maschinelles Lernen (ML), das derzeit für Erfolge im Feld der KI hauptverantwortliche Verfahren, zwar ein mächtiges Werkzeug zur Mustererkennung (etwa in Bildern, Schrift oder Sprache).[3] Es ist aber auf eng definierte Problemstellungen begrenzt. Mit der flexiblen und generellen Kompetenz, die mit menschlicher Intelligenz einhergeht, ist es nicht vergleichbar. Unterschätzt wird währenddessen, welche Auswirkungen bereits solche „un-intelligenten" Systeme nichtsdestotrotz entfalten können. Man denke etwa an die algorithmengenerierten und leicht manipulierbaren „Filterblasen" in sozialen Netzen, die derzeit die Funktion der freien Medien als vierte Gewalt sowie etablierte Prozesse demokratischer Willensbildung beeinträchtigen.

Die kommende Phase der Digitalisierung wirft also schon jetzt drängende gesellschaftliche Zukunftsfragen auf. Dazu zählen im Zuge der „Industrie 4.0" auch

1 Die vorliegende Studie ist fokussiert auf die sicherheitspolitischen Implikationen des Phänomens maschineller Autonomie als zukünftig bedeutsamem Aspekt des digitalen Fortschritts. Der Cyber- und Informationsraum steht hier nicht im Zentrum.

2 Unter dem weiten und nicht einheitlich definierten Begriff der Künstlichen Intelligenz werden eine Vielzahl unterschiedlicher softwarebasierter Techniken und Verfahren zur Automatisierung von Aufgaben subsumiert, die bisher die Anwendung menschlicher Intelligenz erforderten. Im Folgenden wird daher auf den wenig hilfreichen KI-Begriff weitestgehend verzichtet. Stattdessen werden stets die konkreten, für die jeweilige Diskussion relevanten Techniken – z. B. maschinelle Bilderkennung – benannt.

3 ML benötigt aktuell noch viel Rechenleistung und gewaltige Datenmengen, um das sogenannte deep learning (mit neuronalen Netzen zur Klassifizierung von Daten) zu realisieren. Das Verfahren soll in naher Zukunft deutlich effizienter und schlanker werden.

ökonomische und soziale sowie (maschinen-)ethische, etwa bei der Nutzung von Robotern in der Pflege.

Maschinelle Autonomie in den Streitkräften

Auch für Streitkräfte birgt die Verschiebung in der Aufgabenteilung zwischen Mensch und Maschine neue Vorteile und Herausforderungen. Das Augenmerk dieser Studie liegt dabei nicht auf bereits existierenden Anwendungen, etwa beim maschinellen Zusammenführen und Analysieren von Daten oder in Führungsunterstützungssystemen. Es richtet sich stattdessen zukunftsgewandt auf das Phänomen der zunehmenden „Autonomie in Waffensystemen" (AWS).[4]

Vollautonomie in Waffensystemen, definiert in Anlehnung an die AWS-Direktive der USA und das Internationale Komitee des Roten Kreuzes (IKRK), bedeutet, dass ein Waffensystem nach Aktivierung mit Hilfe von Sensoren und Software selbständig, also im Unterschied zu ferngesteuerten Systemen ohne jedwede menschliche Kontrolle oder Aufsicht, einen kompletten Entscheidungszyklus (targeting cycle) durchlaufen kann: Das System sucht und findet Ziele aktiv, fixiert und verfolgt diese und führt auch die sogenannten kritischen Funktionen der Zielauswahl und -bekämpfung ohne menschliches Zutun aus (find, fix, track, select, engage, assess).

Waffensysteme, die in dieser Form „selbständig" Ziele bekämpfen, sind im Prinzip nicht neu. Verteidigungssysteme wie etwa PATRIOT sind seit Jahrzehnten im Einsatz.[5] Unter Zeitdruck können auch sie Ziele ohne menschliches Eingreifen bekämpfen (terminal defense). Sie sind allerdings – in der Regel – stationär, führen wiederholt die immer gleichen vorprogrammierten Aktionen aus und richten sich gegen Munition, also unbelebte Ziele.

Für diese Studie relevant ist Autonomie hingegen in mobilen Systemen, die über längere Zeit in dynamischen, unstrukturierten, offenen Umgebungen operieren und dabei den targeting cycle ohne menschliches Zutun absolvieren. Vereinzelt, etwa mit loitering munitions wie der israelischen Anti-Radar-Munition Harpy, ist auch diese Form der Waffenautonomie bereits im Einsatz (wenn auch im Falle Harpys mit begrenztem Einsatzzweck, nämlich dem Kreisen über einem Gebiet und der Bekämpfung gegnerischer Luftabwehrradars).

Autonomie:Vorteile und Herausforderungen

Welche Vorteile und welche Herausforderungen für Deutschland und die Bundeswehr birgt die AWS-Entwicklung auf operativer und strategischer Ebene?

Ganz allgemein bieten autonome Waffensysteme den Vorteil, eintönige, unangenehme sowie insbesondere gefährliche Aufgaben übernehmen zu können. Im Konkreten sind aktuell drei weitere Vorteile in der Diskussion.

Erstens dient AWS als Kräftemultiplikator: Ein/e Soldat/in soll demnach in Zukunft viele autonome Systeme oder Schwärme führen.

Zweitens macht Autonomie Steuerungs- und Kommunikationsverbindungen optional. Diese sind sowohl störungs- als auch kaperungsanfällig und geben mitunter den Aufenthaltsort von Systemen preis. Außerdem liegt zwischen menschlichen Fernsteuerbefehlen und ihrer Ausführung stets eine Zeitverzögerung. Auf diese Verbindungen nicht angewiesene, autonome Systeme versprechen eine latenzfreie – also höhere – Operationsgeschwindigkeit sowie Operationsfähigkeit dort, wo Steuerung und Kommunikation schwierig, unmöglich oder ungewünscht sind.

Drittens soll die Fusion von Echtzeit-Aufklärung, Entscheidungsgeschwindigkeit und präzisem Waffeneinsatz ohne Zeitverzögerung es erlauben, Prinzipien des Kriegsvölkerrechts noch besser umzusetzen, also etwa zivile Opfer und Schäden an zivilen Objekten vermeiden.

Insbesondere dieser dritte Punkt ist jedoch unter Juristen und KI/Robotik-Experten höchst umstritten. Die Frage, ob Autonomie in Waffensystemen völkerrechtskonforme Streitkräfteoperationen erleichtert oder vielmehr doch erschwert, führt zu den Implikationen von AWS für die Konfliktbearbeitung.

Zukünftige Konfliktbearbeitung: Verdrängung der menschlichen Verfügungsgewalt?

Konfliktbearbeitung wird hier operativ, also im Sinne von Streitkräfteoperationen, verstanden. AWS führen hier tatsächlich einen neuen Faktor ins Geschehen ein. Denn Vorsicht: Vorschnelle Vergleiche mit bereits existierenden Zielsuchmunitionen oder Projektilen, die einer ballistischen Flugbahn folgen, führen in die Irre. AWS ist dem Entlassen von Munition auf eine nicht mehr beeinflussbare Flugbahn oder über einen Rückholpunkt hinweg nicht äquivalent. Denn Autonomie soll es – vereinfacht gesagt – ermöglichen, einem Waffensystem einen „Auftrag" zu erteilen. Das System operiert dann, womöglich über einen längeren Zeitraum, ohne menschliche Steuerung und Aufsicht, dabei eigene „Entscheidungen" über die Bekämpfung von Zielen treffend. Insofern liegt eine neue Situation vor (andernfalls würde die Entwicklung ja erst gar nicht angestrebt). Eine neuerliche und eigenständige juristische Bewertung ist daher erforderlich. Diese legt eine Reihe von Problemen nahe.

4 Häufig, insbesondere im Kontext der aktuellen Diskussion bei den Vereinten Nationen in Genf, findet auch das Akronym LAWS (für lethal autonomous weapon systems) Verwendung.

5 Selbst Minen lassen sich unter ein breites Verständnis von Waffenautonomie subsumieren – zumindest solche, die anhand bestimmter Signaturen eine „Ziel-Auswahl" treffen und nicht nur auf einem primitiven, opferaktivierten An-/Aus-Mechanismus ohne Selbstregulierungsschleife beruhen.

Erstens ist bisher noch ungeklärt, ob für ein legales autonomes Durchlaufen des targeting cycle überhaupt die technischen Voraussetzungen existieren, ob also AWS überhaupt völkerrechtskonform sein können. Zumindest nach dem aktuellen Stand der Technologie ist diese Frage zu verneinen. Kaum ein Vertreter aus den relevanten technischen Feldern hält es für möglich, die notwendigen Unterscheidungen (etwa zwischen Kombattant und Zivilist – was bisweilen schon für Menschen extrem schwierig, weil kontextabhängig, ist) und Beurteilungen über die Angemessenheit militärischer Mittel maschinell abzubilden. Zudem haben sich auf deep neural networks beruhende Mustererkennungssysteme, die derzeit den Stand der Technik im Bereich maschinelle Bilderkennung darstellen, als hochgradig manipulationsanfällig erwiesen.[6]

Zweitens existiert für den Fall der Nutzung von AWS, anders als etwa beim Abfeuern einer Zielsuchrakete, aktuell keine juristische Mehrheitsmeinung mit Blick auf die Frage, wo bzw. bei wem die kriegsvölkerrechtliche Verantwortung für den Wirkmitteleinsatz zu verorten ist. Dies wäre bereits unter der Annahme fehlerlos operierender autonomer Systeme ein Problem. Da Fehler jedoch unvermeidlich sind, verursacht durch Soft- und Hardware oder den Nebel des Krieges sowie gegnerische Einflüsse, bergen AWS das Risiko, im Falle von auftretenden Rechtsverstößen eine inakzeptable rechtliche „Verantwortungslücke" zu erzeugen.

Drittens wird mit AWS auch die Schnittstelle zwischen Völkerrecht und Ethik berührt. Denn wenn Systeme bspw. nicht in der Lage sind, Personen, die aufgrund von Verwundung oder Kapitulation nicht mehr am Kampfgeschehen teilnehmen (völkerrechtlich „hors de combat"), verlässlich zu erkennen und angemessen zu behandeln, dann reicht dies über einen rein rechtlichen Sachverhalt sowie den Einzelfall hinaus. Die Autonomieentwicklung kommt hier in Konflikt mit dem generellen Gebot, Menschen – auch im Krieg – nicht ihrer Würde zu berauben und unnötigen Grausamkeiten auszusetzen.

Autonomie in Waffensystemen, die die kritischen Funktionen der Zielauswahl und -bekämpfung miteinschließt, ist also, von Verteidigung gegen Munition abgesehen, aufgrund der Verdrängung menschlicher Verfügungsgewalt (und somit rechtlicher und ethischer Verantwortung) operativ mit kriegsvölkerrechtlichen und ethischen Risiken behaftet. Diese gilt es aus Sicht Deutschlands und der Bundeswehr zu meiden. An Maschinen, deren Konformität mit dem Kriegsvölkerrecht fraglich ist und die weder Moral noch den Unterschied zwischen Leben und Tod begreifen, sollte die Bundeswehr Tötungsentscheidungen nicht delegieren.

Zukünftige Konfliktformen: Schlachtfeld-Singularität und Eskalationsdynamik?

Die strategischen Auswirkungen von Autonomie auf Konfliktformen in globaler Perspektive werden maßgeblich durch den oben bereits erwähnten Anstieg der Operationsgeschwindigkeit beeinflusst.

Autonomie zieht mehr Autonomie nach sich, weil Geschwindigkeit, definiert als die Fähigkeit zum Handeln innerhalb des gegnerischen Entscheidungszyklus, entscheidende Vorteile verspricht. Am Ende des Geschwindigkeitswettlaufs steht jedoch, wie es der für die AWS-Entwicklung in den USA federführende, stellvertretende US-Verteidigungsminister Robert Work einmal besorgt formulierte, das unausweichliche Delegieren von Tötungsentscheidungen an Maschinen, was auch aus US-Sicht ungewünscht bleibt. China nutzt dafür den Begriff der „Schlachtfeldsingularität" – den Zeitpunkt, an dem das zunehmende Gefechtstempo die menschliche Verfügungsgewalt unwiederbringlich hinter sich lässt.

Die zwei größten, damit verbundenen strategischen Risiken sind verschärfte Eskalationsdynamiken und Instabilitäten.

Die Eskalationsproblematik erwächst aus der unvorhersehbaren Interaktion zweier oder mehrerer autonomer Systeme. Im high frequency trading an den Finanzmärkten treten heute bereits regelmäßig unvorhergesehene Interaktionsprozesse zwischen zwei oder mehr autonom operierenden Handelsalgorithmen auf, was nicht selten sogar flash crashes und somit finanziellen Schaden verursacht. Dem kann an den Finanzmärkten regulativ begegnet werden; ohne eine kriegsvölkerrechtlich verbindliche und verifizierte Regulierung von AWS auf dem Schlachtfeld aber bedeutet dies, dass mit der Konfrontation generischer AWS zukünftig nicht-intendierte Wechselwirkungen, bis hin zu ungewolltem Waffengebrauch, drohen. Ein „flash war" würde so womöglich durch autonome Angriffe und Gegenangriffe mit hohem Tempo eskalieren, bevor der Mensch korrigierend eingreifen kann.

Autonomie in Waffensystemen befördert darüber hinaus Instabilität, mit potenziell besonders gravierenden Konsequenzen für die nukleare Ebene. Unter dem Stichwort entanglement werden hier Effekte diskutiert, die aus den zunehmenden Kapazitäten konventioneller Waffensysteme – darunter Autonomie – für die strategische Ebene erwachsen, etwa nicht-nukleare Gefahren für Nuklearwaffen sowie C3I Systeme. Konkret bringt Autonomie so etwa im Bereich der Seekriegsführung neue Möglichkeiten zur U-Boot-Bekämpfung mit sich. Mit der Sea Hunter wird etwa im Rahmen des DARPA Anti-Submarine Warfare Continuous Trail Unmanned Vessel (ACTUV)

6 Die derzeit an der Schnittstelle zwischen maschinellem Lernen und Computersicherheit unter dem Schlagwort adversarial examples betriebene Forschung legt zudem nahe, dass maschinelle Bilderkennung Gegnern somit auch neue Angriffsflächen bietet, etwa indem dieser autonomen Systemen verlässlich falsche Tatsachen vorgaukeln oder sie sogar, im Falle von im Feld lernenden Systemen, durch wiederholtes Täuschen gezielt „umtrainieren" kann.

Programms derzeit ein autonomer Trimaran getestet. Seine Fähigkeit, getauchte und mit ballistischen Nuklearraketen bestückte U-Boote zu detektieren und zu verfolgen, schränkt die gesicherte nukleare Zweitschlagsfähigkeit anderer Nuklearwaffenstaaten ein. Verschärft wird das entanglement-Problem durch die zunehmende Bereitschaft, nicht-nukleare Angriffe – wozu neben den Waffen selbst, wie oben gesehen, auch Frühwarn- und Kontrollsysteme gezählt werden – nuklear zu vergelten. Signifikante, nicht-nukleare strategische Angriffe, wie es in der Nuclear Posture Review der Trump Administration heißt, sollen durch die USA neuerdings auch nuklear beantwortet werden können. Diese Haltung fand sich auf russischer Seite aufgrund des konventionellen rüstungstechnologischen Vorsprungs der USA schon länger; nun spiegeln die USA sie ihrerseits, was der Stabilität zwischen den beiden größten Nuklearmächten weiter abträglich sein dürfte.

Deutschland muss also ein über die unmittelbar eigenen Streitkräfte hinausreichendes, übergeordnetes Interesse an der Eindämmung der durch Autonomie in Waffensystemen entstehenden Risiken haben. Nicht zuletzt, weil die AWS zugrundeliegende dual-use-Technologie proliferationsanfällig, bis hin zu nicht-staatlichen Akteuren, ist.

Empfehlungen: Vorteile nutzen, Risiken meiden

Die Bundesregierung lehnt gemäß aktuellem Koalitionsvertrag „[a]utonome Waffensysteme, die der Verfügung des Menschen entzogen sind, [ab]" und zielt (weiterhin) darauf ab, solche „weltweit [zu] ächten"; ein Standpunkt, den auch Generalleutnant Ludwig Leinhos auf der diesjährigen Münchner Sicherheitskonferenz bekräftigt hat. Denn Autonomie in Waffensystemen bietet einerseits zahlreiche militärische Vorteile. Andererseits gehen mit ihr operative und strategische Risiken einher, wenn der targeting cycle gänzlich – also bis hin zum Auswählen und Bekämpfen von Zielen – der menschlichen Verfügungsgewalt entzogen wird. Ratsam ist daher eine differenziert-abgestufte Nutzung von Autonomie in Waffensystemen, mit besonderer Vorsicht hinsichtlich der letzten beiden Phasen des targeting cycle.[7]

Die Empfehlung dieser Studie lautet folglich, die Vorteile von Autonomie in den unkritischen Phasen des targeting cycle (etwa Navigation, Erkennen von Zielen u.ä.) zu suchen. Den Risiken einer autonomen Zielauswahl und -bekämpfung sollte zugleich doktrinär und rüstungskontrollpolitisch wie folgt vorgebeugt werden:

Maßnahmen mit Blick auf operative Risiken

- Entwickeln und Veröffentlichen eines zunächst fünf Jahre gültigen und dann zu überprüfenden Richtliniendokuments, in dem die Bundeswehr einen Umgang mit Autonomie in Waffensystemen festschreibt (wie es Verbündete wie USA, GB und NLD tun). Demgemäß sollten die beiden letzten Phasen des targeting cycle stets menschlicher Verfügungsgewalt unterliegen; mit anderen Worten, die Bundeswehr sollte sich keine nach diesem Verständnis „vollautonomen" Waffensysteme erlauben.
- Fortgesetzte Nutzung von präzise definierten Verteidigungssystemen[8] als Ausnahme[9] festschreiben und verregeln.
- Initiieren eines Forschungsprogramms, um vorausschauend die für die Wahrung menschlicher Verfügungsgewalt über Waffensysteme notwendigen Parameter in zukünftigen Mensch-Maschine-Interaktionen in Waffensystemen (bspw. das Future Combat Air System) auszuloten. Forschungsfrage: Wie weit kann menschliche Entscheidungs- und Verfügungsgewalt im targeting cycle maschinell unterstützt werden, bevor sie ihren Wesenskern einbüßt?[10]

Maßnahmen mit Blick auf strategische Risiken

- Identifizieren vertrauensbildender Maßnahmen und best practices zur Wahrung der menschlichen Verfügungsgewalt über Waffensysteme in Kooperation mit alliierten Militärs.
- Fortsetzen und Intensivieren deutscher Bemühungen um eine internationale, völkerrechtlich bindende Regulierung von Autonomie in Waffensystemen, die das Auswählen und Bekämpfen von Zielen ohne menschliche Verfügungsgewalt außerhalb von Verteidigungssystemen verbietet.
- Ausloten der Möglichkeiten zur effektiven Rüstungskontrolle durch geeignete Methoden zur Verifikation eines solchen Verbots.

7 Die Implikationen von Autonomie in früheren Phasen des targeting cycle könnten Gegenstand einer späteren Metis Studie sein. Es wäre noch genauer zu analysieren, wie eine rechtlich, ethisch und militärisch ausgewogene Aufgabenteilung zwischen Mensch und Maschine in diesen Phasen gestaltet werden muss.

8 Etwa: „Fest installierte Systeme, die auf die automatische Ausführung einer kleinen und genau vorbestimmten Anzahl vorprogrammierter Aktionen beschränkt sind und in einfach strukturierten Umgebungen ausschließlich in Reaktion auf Beschuss und gegen unbelebte militärische Objekte unter höchstem Zeitdruck eingesetzt werden dürfen".

9 Für weitere, eventuell unter eine funktionale AWS-Definition fallende Altsysteme der Bundeswehr, beispielsweise bestimmte Seeminen, wären mögliche Ausnahmen im Einzelfall gesondert zu prüfen.

10 Vgl. FN 7.

Metis

IMPRESSUM

Herausgeber

Metis Institut
für Strategie und Vorausschau
Universität der Bundeswehr
München
metis.unibw.de

Autor

Dr. Frank Sauer
metis@unibw.de

Creative Director

Christoph Ph. Nick, M.A.
c-studios.net

Titelbild

Arif Wahid auf Unsplash

ISSN-2627-0587