



Arbeitspapier Sicherheitspolitik, Nr. 26/2018

Künstliche Intelligenz in den Streitkräften: Zum Handlungsbedarf bei Autonomie in Waffensystemen

von Frank Sauer

Weltweit haben Streitkräfte damit begonnen, den Einsatz von Künstlicher Intelligenz (KI) für militärische Zwecke auszuloten. Die Bundeswehr bildet dabei keine Ausnahme. Bereiche, in denen KI ihre Stärken ausspielen und Prozesse in der Bundeswehr optimieren und beschleunigen kann, sind zum Beispiel Logistik, vorausschauende Wartung, Einsatzunterstützung und -führung, aber auch die Analyse großer Datenbestände zur Krisenfrüherkennung. Ein deutlich sensiblerer Bereich ist die Nutzung von KI zur Steigerung von „Autonomie“ in Waffensystemen – bis hin zu vollautonomen Waffensystemen, die Ziele ohne menschliche Kontrolle auswählen und bekämpfen. Dieses Papier eruiert den Neuigkeitswert und den funktionalen Charakter von „Waffensystemautonomie“. Daraufhin skizziert es die mit vollautonomen Waffensystemen verbundenen Risiken. Zuletzt unterbreitet das Papier drei Vorschläge, wie Deutschland sich – Risiken meidend, Chancen nutzend – international und national mit Blick auf die Bundeswehr positionieren sollte.

Wissenschaftler und Vertreter ziviler Technologieunternehmen, den Innovationsmotoren im Feld der KI, warnen schon seit Jahren unter großer medialer Aufmerksamkeit vor einem Paradigmenwechsel in der Kriegsführung. Dazu zählen der jüngst verstorbene Stephen Hawking ebenso wie etwa Elon Musk oder auch Demis Hassabis und Mustafa Suleyman, die Gründer von Googles KI-Schmiede DeepMind. Sie alle machen auf die völkerrechtlichen, sicherheitspolitischen und ethischen Risiken aufmerksam, sollten Waffensysteme zukünftig vermehrt außerhalb menschlicher Verfügungsgewalt operieren. Befördert durch diese Warnungen aus der Zivilgesellschaft finden auf Ebene der Vereinten Nationen (VN) im Rahmen der VN-Waffenkonvention in Genf bereits seit 2014 multilaterale Gespräche über eine mögliche Rüstungskontrolle für vollautonome Waffensysteme, im VN-Sprachgebrauch LAWS (*Lethal Autonomous Weapon Systems*) genannt, statt.

Waffenautonomie als Funktionalität

Bisher fehlt allgemein anerkannte (geschweige denn völkerrechtlich festgeschriebene) Definition des Autonomiebegriffs bei Waffensystemen. Es zeichnet sich aber im internationalen Fachdiskurs zunehmend die Durchsetzung einer funktionalen Autonomiedefinition ab – nicht zuletzt deshalb, weil schwergewichtige Spieler wie die USA oder das Internationale Komitee des Roten Kreuzes (IKRK) sich ein solches Verständnis von Waffenautonomie offiziell zu eigen gemacht haben. Vollautonome Waffensysteme würden dieser Definition zufolge nach ihrer Aktivierung mit Hilfe von Sensoren und Software selbständig, also im Unterschied zu ferngesteuerten Systemen ohne menschliche Kontrolle oder Aufsicht, den Entscheidungszyklus zur Zielbekämpfung durchlaufen. Dieser *targeting cycle* umfasst die Stationen *find, fix, track, target, engage* und *assess* – abgekürzt F2T2EA. LAWS wären also insbesondere im Rahmen der völkerrechtlich, ethisch und politisch besonders sensiblen Auswahl und Bekämpfung von Zielen – den, wie das IKRK sie nennt, „kritischen Funktionen“ des *targeting cycle* – der menschlichen Verfügungsgewalt entzogen.

Legt man diese Konzeption an, existieren bei genauem Hinsehen schon heute Waffensysteme, die als LAWS bezeichnet werden können. Das israelische System *Harpy* zum Beispiel richtet sich gegen Radarinstallationen und durchläuft, zumindest für diesen begrenzten Einsatzzweck (das Kreisen über einem Gebiet und Bekämpfen gegnerischer Luftabwehr), den *targeting cycle* bereits außerhalb der menschlichen Verfügungsgewalt – und gilt somit in Fachkreisen folgerichtig als Beispiel eines bereits existierenden LAWS. Es geht im Rahmen der LAWS-Diskussion also mitnichten um „Zukunftswaffen“.

Auch die häufig anzutreffende Kontrastierung von „nur automatisch“ Ziele bekämpfenden Verteidigungssystemen (der Vergangenheit) und „autonomen“ Systemen mit offensiven Fähigkeiten (der Zukunft) hält der Überprüfung im Lichte einer funktionalen Autonomiedefinition nicht stand. Es ist zwar richtig: Seit Jahrzehnten werden Verteidigungssysteme gegen Raketen, Artilleriegeschosse oder Mörsergranaten eingesetzt, um die anfliegende Munition – unter extremem Zeitdruck und gegebenenfalls ohne menschliches Eingreifen – zu bekämpfen. Im Unterschied zu solchen beispielsweise auf Schiffen oder in Feldlagern aufgestellten Waffen werden LAWS häufig als mobile Waffensysteme – beziehungsweise Waffenplattformen – verstanden, die moderne Technologien – etwa *machine learning* – in den militärischen Einsatz bringen.¹

Aber Autonomie hängt gerade *nicht* von der Fortschrittlichkeit der eingesetzten Technologien ab. Zahlreiche schon lange existierende Verteidigungssysteme – etwa das auch von der Bundeswehr genutzte PATRIOT Flugabwehrsystem – waren und sind grundsätzlich zu Autonomie in den kritischen Funktionen befähigt und sind folglich nach der hier angenommenen Definition „vollautonom“. Das fiel bisher, salopp gesagt, bloß niemandem besonders auf. Erst jetzt, da Vollautonomie zunehmend auch in anderen, mobilen Systemen Einzug hält, werden ihre Implikationen deutlich. Genau so wenig wie an bestimmte Hochtechnologien ist Autonomie im Übrigen notwendigerweise an einzelne konkrete Waffensysteme gebunden. Die Autonomie ermöglichende „Intelligenz“ kann auch über einen Verbund aus Waffensystemen oder einen Schwarm verteilt sein.

Zwischenfazit: Waffenautonomie ist kein kategoriales Unterscheidungsmerkmal

Autonomie in Waffensystemen markiert somit weder eine spezifische noch eine revolutionär neue *Kategorie* von Waffen. Die Vorstellung, eine kategoriale Unterscheidung zwischen „autonomen Waffensystemen“ auf der einen und „nicht-autonomen Waffensystemen“ auf der anderen Seite vornehmen zu können, führt in die Irre. Nicht zuletzt, weil mit den neuen technischen Möglichkeiten, die mit KI aktuell Einzug halten, alte Systeme nachgerüstet (gleichsam „autonomisiert“) werden können und neue Systeme optional ferngesteuert oder eben autonom sein werden. Kurz: Man wird zukünftigen Waffensystemen nicht ansehen, „wie autonom“ sie sind. Das Beharren auf einer auf kategorialen Unterscheidungen basierenden Definition von LAWS, die zwischen autonom und nicht-autonom zu trennen versucht, wird somit nicht nur dem Gegenstand intellektuell nicht gerecht, es ist auch längst zum Hemmschuh in den rüstungspolitischen Gesprächen bei den VN in Genf geworden.

Demgegenüber ist es intellektuell schlüssiger und politisch sinnvoller, den USA und dem IKRK folgend eine funktionale Definition von Autonomie zu akzeptieren – auch, wenn manche es kontraintuitiv finden mögen, „alten“ Waffensystemen Vollautonomie zu attestieren. Die falsche Schlussfolgerung daraus wäre nun wiederum, den etablierten und bisher – völlig zu Recht – unumstrittenen Verteidigungssystemen urplötzlich ein Autonomie-Problem zu attestieren. Die Praxis der autonomen Abwehr von Granaten und Raketen warf schließlich bisher – und wirft auch weiterhin – wenig bis keine völkerrechtlichen, ethischen und sicherheitspolitischen Probleme auf. Im Gegenteil: Die Schutzfunktion dieser Systeme bei der Abwehr dieser unbelebten Ziele ist hochwillkommen und wird – im Lichte der Entwicklung von Flugkörpern mit Fluggeschwindigkeiten im vielfachen Überschallbereich – zukünftig mehr denn je militärisch benötigt werden.

¹ *Machine Learning* ist zurzeit für die meisten Fortschritte im Feld der KI verantwortlich.

Politische Entscheidungsträger sollten die LAWS-Debatte daher nicht als eine um Technologie, sondern als eine um bestimmte, akzeptable oder nicht-akzeptable militärische Praktiken bei der Nutzung von Waffensystemen mit autonomen Funktionen verstehen. Es gilt mit Blick auf Autonomie in Waffensystemen also vorausschauend zu fragen, wann und inwieweit menschliche Verfügungsgewalt bei der Nutzung eines Waffensystems reduziert oder gänzlich verdrängt wird – und was jeweils die Konsequenzen daraus sind. Neue Technologien aus dem Bereich der KI spielen also bei der LAWS-Entwicklung eine ermöglichende Rolle. Aber der echte Neuigkeitswert steckt in der über bisherige Anwendungsszenarien hinausgehenden militärisch-praktischen Nutzung autonomer Funktionen in Systemen – dort, wo Autonomie bisher nicht anzutreffen war. Das führt zum Szenario eines zukünftig womöglich vollautonom geführten Gefechts gegen Panzer, Schiffe, Flugzeuge oder Infanterie. Anders als bei der autonomen Abwehr von Munition, also unbelebten Objekten, wäre Waffenautonomie in einem solchen Kontext hochproblematisch.

Waffenautonomie und die Risiken

Hält die bisher nur vereinzelt in Systemen wie Harpy oder beschränkt auf den Flugabwehrbereich wie bei PATRIOT existierende Waffenautonomie zukünftig auch in mobilen Systemen Einzug, die über längere Zeit operieren und die kritischen Funktionen beim Bekämpfen belebter Ziele oder direkt gegen Menschen ohne menschliche Verfügungsgewalt komplettieren, dann wirft das zahlreiche gravierende Probleme auf. Aus kriegsvölkerrechtlicher Sicht erzeugen solche Waffensysteme eine Verantwortungslücke. Denn es ist unklar, wer im Rahmen der Anwendung militärischer Gewalt die Verantwortung trägt, sollten sie Zivilisten ein dem militärischen Ziel unangemessenes, nicht rechtfertigbares und somit illegales Leid zufügen. Auch bedroht das globale militärische Wettrennen um Vollautonomie in Waffensystemen die globale militärische Stabilität und erhöht die Gefahr nichtintendierter Eskalationen. Von den Finanzmärkten kennen wir die unvorhersehbaren Interaktionen zwischen Algorithmen schon, die bisweilen in sogenannten *flash crashes* münden – ein maschinell losgetretener *flash war* ist ohne die menschliche Verfügungsgewalt als Sicherheits- und Entschleunigungsinstanz ein reales Risiko. Schließlich bleibt noch der wichtigste, moralische Einwand – nämlich der Verweis darauf, dass es die Würde des Menschen verletzt, Entscheidungen über Leben und Tod auf dem Schlachtfeld an Algorithmen zu delegieren. Denn das Töten im Krieg auf Maschinen auszulagern und automatisch „abarbeiten“ zu lassen, macht Menschen zu Objekten. Für die Getöteten mag es keinen Unterschied machen, ob ein Mensch oder ein Algorithmus ihren Tod bewirkt hat. Aber die Gesellschaft, die eine solche Praxis erlaubt und mit dem Töten im Krieg ihr kollektives menschliches Gewissen nicht mehr belastet, riskiert nicht weniger als die Aufgabe der grundlegendsten zivilisatorischen Werte und humanitären Prinzipien.

Fazit: Zum Handlungsbedarf bei der differenzierten Nutzung von Waffenautonomie

Die Weichen für die zukünftige Nutzung von Autonomie in Waffensystemen werden heute gestellt. Im Falle des Abtretens von Verfügungsgewalt an Algorithmen, insbesondere mit Blick auf die kritischen Funktionen von Waffensystemen, ist im Lichte der oben skizzierten Risiken die richtige Entscheidung, an der menschlichen Kontrolle grundsätzlich festzuhalten – die prominente Ausnahme bilden die oben erwähnten Verteidigungssysteme gegen Granaten und Raketen. Es bedarf also eines differenzierenden und besonnenen Vorgehens bei der Vermählung von KI mit Waffen. Deutschland ist im Lichte dieser Erkenntnis bei den VN in Genf im Rahmen der *Convention on Certain Conventional Weapons* (CCW) rüstungskontrollpolitisch aktiv. Dort eruiert die Staatengemeinschaft Bedarf und Möglichkeiten multilateraler Rüstungskontrolle. Aktuell fordern im CCW-Rahmen 26 Staaten ein völkerrechtlich bindendes Verbot von ohne menschliche Verfügungsgewalt operierenden Waffen. Österreich hat sich zu dieser Forderung bekannt und sich bei den letzten Gesprächen im August 2018 dafür sehr aktiv eingesetzt. China sendet seit April 2018 ähnliche, wenngleich unklare Signale. Besonders die USA und Russland lehnen jedoch jedwede Regulierung strikt ab. Deutschland hat sich der Gruppe der Verbotsbefürworter bei den VN bisher nicht angeschlossen, erteilt aber laut der Koalitionsverträge von 2013 und 2018² vollautonomen Waffensystemen eine klare Absage und fordert ihre internationale Ächtung.

² „Autonome Waffensysteme, die der Verfügung des Menschen entzogen sind, lehnen wir ab. Wir wollen sie weltweit ächten.“ (7027 bis 7028)

In Genf versucht Deutschland nun gemeinsam mit Frankreich über eine politische Deklaration und weichere Regulierung einen Mittelweg zwischen Verbotsbefürwortern und -gegnern zu bahnen und so im Prozess weiter schrittweise auf ein Verbot hinzuarbeiten. Auf nationaler Ebene haben hochrangige deutsche Militärs, wie zuletzt der Inspekteur des Organisationsbereichs Cyber- und Informationsraum Generalleutnant Ludwig Leinhos auf der Münchner Sicherheitskonferenz 2018, die Position der Bundesregierung bekräftigt. Was den nationalen Regelungsbedarf angeht, so reicht für die Bundeswehr ein Verweisen darauf, auch im Lichte zunehmender Waffensystemautonomie stets die Vorgaben des Internationalen Humanitären Völkerrechts einzuhalten, nicht aus. Man wird damit der Tragweite der Entwicklung – dem Einzug einer neuen Funktionalität und der Neubestimmung des Mensch-Maschine-Verhältnisses im Krieg – nicht gerecht. Es bedarf vielmehr neuer Regeln. Die Vorgaben des Koalitionsvertrags könnten in drei Schritten national und international umgesetzt beziehungsweise weiterverfolgt werden:

Schritt 1) Erarbeiten eines Bundeswehr-Leitliniendokuments zu Autonomie in Waffensystemen

Dieses Dokument sollte ein funktionales und auf den kritischen Funktionen von Zielauswahl und Zielbekämpfung beruhendes Autonomieverständnis anführen. Mit ihm würde eine zukunftsweisende Vorschrift zur Wahrung menschlicher Verfügungsgewalt über alle Waffensysteme der Bundeswehr verankert. Konkret würde damit eine Mensch-Maschine-Interaktion festgeschrieben, die maximal das softwareseitige Selektieren von Zielen, die vom Menschen dann zur Bekämpfung freigegeben werden müssen, erlaubt. Nur bei aktuellen und zukünftigen Verteidigungssystemen, siehe Schritt 2, wäre ein höherer Grad an maschineller Autonomie akzeptabel.³ Dieses Dokument könnte alle fünf Jahre im Lichte technologischer Entwicklungen geprüft und bei Bedarf aktualisiert werden.

Schritt 2) Fortgesetzte, regulierte Nutzung von vollautonomen Verteidigungssystemen

Das Leitliniendokument sollte die fortgesetzte Nutzung von vollautonomen Verteidigungssystemen als Ausnahme deklarieren, aber diese vorausschauend mittels folgender Rahmenbedingungen und Beschränkungen für Konstruktion und Betrieb regulieren. Demnach dürfen vollautonome Verteidigungssysteme der Bundeswehr:

- nur unter Zeitdruck defensiv gegen direkte Angriffe durch unbelebte militärische Objekte eingesetzt werden;
- nur wenige, genau definierte und in ihrem Effekt vorhersehbare vorprogrammierte Aktionen wiederholt unabhängig von externen Einflüssen oder Kontrollen durchführen;
- nur verwendet werden in Umgebungen, die relativ übersichtlich sind und ein sehr geringes Risiko ziviler Schäden aufweisen;
- nur fixiert installiert wie zum Beispiel auf einem Schiff oder in einem Feldlager, aber nicht mobil als eigenständiges Waffensystem, operieren.

Schritt 3) Intensivieren der deutschen Arbeit hin zu einem völkerrechtlich bindenden und überprüfbaren CCW-Protokoll für Autonomie in Waffensystemen

Der von Deutschland und Frankreich in Genf eingeschlagene Kurs markiert angesichts der sich zusehends verhärtenden Fronten zwischen Verbotsbefürwortern und -gegnern in der *CCW-Group of Governmental Experts* (GGE) derzeit die Alternative zu entweder einem (teils als verfrüht empfundenen) Sofort-Verbot oder totalem Stillstand. Doch auch der deutsch-französische Mittelweg birgt ein Risiko. Schon sein erster Schritt, eine unverbindliche politische Deklaration, könnte, wenn als *letzter* Schritt missbraucht, die Gespräche ebenfalls weit unter dem eigentlich durch Deutschland laut Koalitionsvertrag angestrebten Ziel des Verbots abwürgen. Deutschland sollte daher 2019 seine Ambition klarer signalisieren und zusätzlich auf ein Verhandlungsmandat für die GGE drängen – damit aus den Genfer Gesprächen bei der VN-Waffenkonvention zumindest echte Verhandlungen werden.

Dr. Frank Sauer forscht und lehrt an der Universität der Bundeswehr München und twittert @drfranksauer. Der Autor gibt seine persönliche Meinung wieder.

³ Für andere, eventuell unter eine funktionale Definition fallende Altsysteme der Bundeswehr, beispielsweise bestimmte Seeminen, wären Ausnahmen im Einzelfall zu prüfen.