

Ein Plädoyer gegen den Einsatz und die Entwicklung autonomer Waffensysteme

Studentische Ausarbeitung zum Seminar *Kriegsroboter, Drohnen und Co.* – zur
Ethik autonomer Waffensysteme bei Prof. Dr. Catrin Misselhorn

Michael Werner Czechowski (Matr.-Nr. 2832616)

30.11.18 · Institut für Philosophie · Universität Stuttgart

Inhaltsverzeichnis

Einleitung	1
Verantwortung	2
Bestehende Regularien	2
Kriegsvölkerrecht bzw. humanitäres Völkerrecht	2
Just War Theory (<i>bellum iustum</i>)	2
Einsatzregeln (<i>Rules of Engagement</i>)	4
Stand der Diskussion um autonome Waffensysteme	4
Convention on Conventional Weapons (CCW)	4
Länderpositionen	5
World Commission on the Ethics of Scientific Knowledge and Technology (COMEST)	6
Technik	8
Software	8
Fehlerquellen und Risiken beim Einsatz und Entwicklung von Software	9
Künstliche Intelligenz (KI)	10
Autonome und semi-autonome Waffensysteme	11
Modellbildung in KI-Systemen	11
Epistemische Opazität	12
Technikskepsis	12
Technikfolgenabschätzung	13
Ethische Einwände	13
Missbrauch und Korruption	14
Vermeintliche Vorteile	14
Philosophische Rechtfertigung des Einsatzes von LAWS	15
Philosophieskeptischer Einwand	16
Schlussplädoyer	17
Bibliographie	19

„WIR, DIE VÖLKER DER VEREINTEN NATIONEN - FEST ENTSCHLOSSEN, künftige Geschlechter vor der Geißel des Krieges zu bewahren, [...] den sozialen Fortschritt und einen besseren Lebensstandard in größerer Freiheit zu fördern, UND FÜR DIESE ZWECKE [...] unsere Kräfte zu vereinen, um den Weltfrieden und die internationale Sicherheit zu wahren, Grundsätze anzunehmen und Verfahren einzuführen, die gewährleisten, daß Waffengewalt nur noch im gemeinsamen Interesse angewendet wird [...] HABEN BESCHLOSSEN, IN UNSEREM BEMÜHEN UM DIE ERREICHUNG DIESER ZIELE ZUSAMMENZUWIRKEN.“

Präambel, UN-Charta

Einleitung

Weltweit befinden sich Entwicklungen zu Künstlicher Intelligenz (KI bzw. engl. AI für *artificial intelligence*) und maschinellem Lernen (engl. *machine learning*) auf dem Vormarsch. Immer mehr Bereiche des alltäglichen Lebens werden bereits zum jetzigen Zeitpunkt von solchen KI-gestützten Systemen ausgeführt oder automatisiert optimiert. Beispielsweise werden handschriftliche Texte auf Überweisungsbögen kaum oder nur in Ausnahmefällen von einem menschlichen Auge begutachtet und in maschinenlesbaren Text überführt. Hierbei sollten gleich zu Beginn KI-Systeme von rein algorithmischen, also auf eindeutigen Handlungsvorschriften basierenden, Systemen auseinander gehalten werden. KI-Systeme unterscheiden sich in der Weise, dass sie weit mehr sind als eine Aneinanderreihung von Wenn-Dann-Bedingungen und haben zudem als Ziel selber diese Unterscheidungsmerkmale im Arbeitsablauf herauszustellen, einzubinden und auf Dauer automatisiert zu optimieren.

Eine Sondergruppe der Vereinten Nationen (UN für engl. *United Nations*) befindet sich seit 2014 in einer andauernden Diskussionsrunde darüber was letale autonome Waffensysteme (LAWS engl. *Lethal Autonomous Weapons Systems*) sind, ob und wie die Entwicklung mit dem humanitären Völkerrecht (IHL für engl. *International Humanitarian Law*) vereinbar ist und was die potentiellen Risiken des Einsatzes dieser neuen Technologie auf die Friedenssicherung unseres Planeten zur Folge haben könnte. Zudem gilt es herauszuarbeiten, wie sich diese Risiken moralisch und ethisch evaluieren und gegebenenfalls eingrenzen lassen würden. 2013 wurde erstmals bei der Convention on Conventional Weapons (CCW) eine Einigung darüber erreicht, dass es für den Bereich von LAWS ein eigenes Mandat geben wird, das erstmalig vom 13. bis 16. Mai 2014 in Genf getagt hatte. (Vgl. „2014 Meeting of Experts on Lethal Autonomous Weapons Systems“, o. J.) Bei dem ersten Treffen äußerten sich Vertreter und Vertreterinnen aus über 30 UN-Mitgliedsstaaten zu diesem Thema. Zudem waren Gäste aus Bereichen des Militärs, Universitäten und Nichtregierungsorganisationen (NGO's für engl. *Non-governmental organization*) geladen, um ihre Einschätzungen zu diesem Thema zu teilen. Seitdem wurden sechs Verhandlungsrunden abgehalten. Die letzte fand vom 28. bis 31. August 2018 statt. Hierzu waren Unterhändler und Unterhändlerinnen von mehr als 75 UN-Mitgliedsstaaten in Genf geladen. (Vgl. Fras, 2018, S. 4.)

Die Debatte um die Entwicklung und den Einsatz von LAWS spaltet zur Zeit viele Bereiche der Wissenschaft und Gesellschaft, solange sie überhaupt darüber ausreichend informiert wird. Zahlreiche Petitionen und öffentliche Arbeitsniederlegungen aus den Forschungsgebieten der KI führen zu einem öffentlichen Druck, der diese Debatte immer kontroverser zu machen scheint. Auf der einen Seite entwickeln Militärs und privatwirtschaftliche Unternehmen weltweit genau diese neue Art des Waffentypus und auf der anderen Seite wird der Zweifel groß, dass die Vereinten Nationen nicht rechtzeitig, also bevor die ersten massentauglich produzierten LAWS auf den Markt kommen, zu einer Einigung kommen werden.

Ich sehe es als philosophische Herausforderung dieses Geflecht an ethischen, technischen, moralischen, zivilrechtlichen als auch völkerrechtlichen Spannungen aufzulösen, um zunächst auch Außenstehenden den Zugang zu diesem Diskurs zu ermöglichen. Meine Haltung ist klar, ich lehne den Einsatz von LAWS, als auch die Entwicklung aus Gründen, die ich später noch nennen werde, kategorisch ab und werde meine Einwände in meinem Schlussplädoyer wiedergeben. Hierzu werde ich mich schrittweise an das Thema annähern. Angefangen bei dem aktuellen Stand der Verhandlungen, worauf ich mich hauptsächlich auf die Materialien der United Nations Office of Geneva (UNOG) von 2017 stützen werde. Anschließend ist es von Bedeutung einen Überblick

über die technischen Begrifflichkeiten als auch die Zusammenhänge zwischen Technik, Software und KI zu geben. Bevor ich dann zu meinem Plädoyer komme, werde ich einen Auszug aus einer Liste von potenziellen Gefahrenquellen nennen und erklären.

Verantwortung

Bei der Entwicklung von letalen autonomen Waffensystemen (LAWS) gibt es eine unüberschaubare Reihe von Interessensgruppen, die gemeinsam für das diffuse Verantwortungsgeflecht sorgen. Aktuell ist es nicht ganz transparent, wer hinter den unterschiedlichen Befürwortern und Gegnern steckt, jedoch gilt es zumindest anhand der verfügbaren Informationen dieses nach Möglichkeit herauszustellen und zu beleuchten.

Vorab lässt sich jedoch sagen, dass die UN als eine der wichtigsten Interessensvertreter gilt neben den privatwirtschaftlichen Unternehmen, Wissenschaftlern und Wissenschaftlerinnen als auch Philosophinnen und Philosophen. Diese Positionen sollen hier zunächst einmal vorgestellt werden.

Bestehende Regularien

Spätestens nach der Einführung des Art. 2 Nr. 4 der UN-Charta von 1945 sind Kriege im herkömmlichen Sinne, sprich mit Waffengewalt ausgetragene Konflikte zwischen Staaten, kaum noch zu rechtfertigen. Der bis dahin geltende Briand-Kellogg-Pakt von 1928 wurde damit de facto außer Kraft gesetzt. Kriege gelten von daher als völkerrechtswidrig. Seit dem Genfer Abkommen zum humanitären Völkerrecht (früher Kriegsvölkerrecht) von 1949 haben sich die meisten Staaten weltweit auf Rechte geeinigt, die bewaffnete Konflikte und ihre Rechtfertigung unter sehr genaue Beobachtung stellt. In Bezug auf Waffengewalt wird hier zur weiteren Klärung ein Auszug aus aktuellen Regularien zu bewaffneten Konflikten aufgelistet.

Kriegsvölkerrecht bzw. humanitäres Völkerrecht

Wie eingangs zitiert setzt sich die UN für den Weltfrieden und die Wahrung der internationalen Sicherheit entschlossen ein. Hierzu ist nicht nur die UN-Charta als wichtigstes Werk anzusehen, sondern gleichfalls das Genfer Abkommen von 1945, welches nach Beendigung des Zweiten Weltkrieges geschlossen wurde.

Just War Theory (*bellum iustum*)

Die Theorie des gerechten Krieges oder oft als Just War Theory betitelt widerspricht sich auf den ersten Blick mit dem Genfer Abkommen. Wie kann Krieg gerechtfertigt oder sogar gerecht sein? Nichtsdestotrotz gibt es sie und wird grundsätzlich in drei Phasen unterteilt. Zum einen dem Recht zum Krieg, dem Recht während eines Krieges und den Rechten und Pflichten nach Beendigung des bewaffneten Konflikts. Alex Leveringhaus fasste diese Theorie folgendermaßen zusammen:

1. Jus ad bellum (justice in the declaration of war). In its traditional formulation, jus ad bellum features six main criteria. These are (1) just cause, (2) proportionality,

- (3) necessity, (4) last resort, (5) right authority, and (6) reasonable likelihood of success.
2. Jus in bello (justice in the conduct of war). In its traditional formulation, jus in bello contains three main criteria. These are (1) distinction, (2) proportionality of means, and (3) necessity.
 3. Jus post bellum (justice in the aftermath of war). Unlike jus ad bellum and jus in bello, jus post bellum is less developed. There is considerable disagreement over which criteria should be included. Brian Orend suggests that jus post bellum's criteria should be conceived analogously to those found in jus ad bellum and jus in bello: (1) just cause for termination, (2) right intention (in terminating hostilities), (3) public declaration and authority, (4) discrimination, and (5) proportionality. (Leveringhaus, 2016, S. 12f)

Entscheidende Begrifflichkeiten sind hierbei die Definition der (1) Kriegsparteien, (2) Proportionalität, (3) Notwendigkeit, (4) Aussicht auf Beendigung des Konflikts und die (5) Unterscheidung von Kombattanten und der Zivilbevölkerung. Es gilt, dass lediglich ein Staat einem anderen Staat den Krieg erklären kann. Demnach sind bewaffnete Konflikte von Zivilpersonen gegen einen Staat nicht als Kriege zu werten und müssen auch nicht nach dieser Theorie bewertet werden. So wären solche Angriffe eher als terroristische Gewalttaten zu werten und müssten auch so auf nationaler oder internationaler Ebene geahndet werden. Proportional wäre ein Angriff genau dann, wenn die Art des Waffeneinsatzes nicht übermäßig größer ist als der der befeindeten Partei. Massenvernichtungswaffen sollten demnach nicht zum Gegenangriff bei einem Feuergefecht von Bodentruppen eingesetzt werden. Der Begriff der Notwendigkeit lässt sich in diesem Kontext sehr schwer definieren, da diese Form der Rechtfertigung auf jede Situation eigens angepasst werden muss und bei jedem Vorfall stets neu ausgehandelt werden muss. Klar ist jedoch, dass Angriffskriege und grundlose Waffenkonflikte von vorne herein ausgeklammert werden können. Wenn abzusehen ist, dass sich keine Beendigung des Konflikts abzeichnen sollte, muss darüber nachgedacht werden, ob die Lösung des Konflikts überhaupt mit einer Kriegserklärung oder Waffengewalt möglich ist. Diese Aussicht muss nämlich stets gegeben sein. Letztlich gilt es wortwörtlich die Zivilbevölkerung aus der Schusslinie zu nehmen. Krieg wird von Kombattanten geführt auch wenn Zivilisten immer wieder ins Ziel von Angriffen geraten. Die Minimierung von so genannten Kollateralschäden, nämlich dem Angriff von Krankenhäusern, Schulen oder Zivilisten im Allgemeinen, gilt als eines der höchsten Ziele bei der bewaffneten Auseinandersetzung zweier oder mehrerer Staaten.

Unter der Betrachtung dieser Regularien zu rechtem Krieg, reflektiert Leveringhaus den Einsatz von LAWS. Hierbei konzentriert er sich auf die soeben genannte Minimierung der Kollateralschäden.

As noted in the first chapter, just war theory, by offering a set of restrictions on the use of force, shares the humanitarian aim of reducing suffering, and there is a great deal of overlap between just war theory and humanitarianism here. From a just war perspective, humanitarian arguments in favour of certain weapons technologies are usually concerned with jus in bello rather than jus ad bellum and jus post bellum. (Leveringhaus, 2016, S. 62)

Aus dieser Sicht wäre der Einsatz von LAWS demnach gerechtfertigt, wenn hierdurch ein höheres humanitäres Ziel erreicht werden könnte. Also, wenn das humanitäre Völkerrecht vorsieht, dass humanitäres Leid einer jeden Bevölkerung zu reduzieren und der Einsatz von LAWS genau dieses

Ziel herbeiführen würde, so wäre dieser Einsatz geboten. Genau diese Argumentation ist im Wesentlichen die beliebteste Rechtfertigung für die Entwicklung und den potenziellen Einsatz von LAWS.

Einsatzregeln (*Rules of Engagement*)

Neben dem vorherrschenden humanitären Völkerrecht gibt es einsatzspezifische Regularien (ROE für engl. Rules of Engagement), die klar und deutlich für jeden Kombattanten und Streitkraft gelten und aufgrund ihrer Wichtigkeit besonders leicht verständlich formuliert sind. Sie sind als Imperative formuliert und sollen dem Kombattanten unmissverständliche Anhaltspunkte geben wann, wo, wie und wer als Ziel für eine militärische Auseinandersetzung mit Waffengewalt gewählt werden darf. (Vgl. Arkin, 2009, S. 81f.) Adressiert sind diese Regularien für bewaffnete Konflikte auf nationaler sowie multinationaler Ebene, als auch Friedensmissionen der UN.

Der Vollständigkeit halber gilt es noch die allgemeineren Regularien zu den ROE zu nennen. Wohingegen die ROE pro Einsatz neu formuliert werden, sind die SROE (Standing Rules of Engagement) vorausgehende allgemeine Regulierungen, die für jeden Einsatz von Waffengewalt gelten. (Vgl. Arkin, 2009, S. 82f.)

Stand der Diskussion um autonome Waffensysteme

Während privatwirtschaftliche Unternehmen wie Samsung (Südkorea) oder Boston Dynamics (Vereinigte Staaten von Amerika) direkt an der Entwicklung von LAWS beteiligt sind, spielen auch viele andere Forschungsbereiche und Unternehmen eine wichtige Rolle bei der Entwicklung.¹ Hierzu jedoch später mehr. Zunächst möchte ich auf die Rolle der Vereinigten Staaten aufmerksam machen, welche seit 2014 regelmäßig zu einem Sonderverhandlungstisch einlädt. Nämlich als Sondermandat des Convention on Conventional Weapons (CCW) und World Commission on the Ethics of Scientific Knowledge and Technology (COMEST). Im weiteren Verlauf wird der bisherige Arbeitsstand der beiden Arbeitsgruppen präsentiert.

Convention on Conventional Weapons (CCW)

In einer Zusammenfassung der venezolanischen Regierung und weiterer Vertragsstaaten wird der CCW zu Beginn des Treffens am 13. bis 17. November 2017 eindringlichst gewarnt vor weiterhin ungeklärter Risiken:

NAM² believes that the following elements shall be included in the substantive discussion of this matter: Ethical and moral concerns about lethal autonomous weapons systems. The discussion on the autonomy, lethality and critical functions of these weapons; [...] Common understandings, definitions and other concepts including of semiautonomous weapons; [...] Military technology and risk of an arms race of fully autonomous weapons, and the technology gap amongst States („Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be

¹Arkin (2008),

²Non-Aligned Movement

Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2017/WP.10“, 2017, S. 1)

Gefordert wird, dass eine Debatte um die moralischen und ethischen Konsequenzen eines Einsatzes und der Entwicklung von LAWS geführt wird. Außerdem seien wesentliche Begrifflichkeiten wie (1) Autonomie und (2) Letalität von LAWS nicht definiert. Zudem fehlt weiterhin eine umfängliche Definition darüber, was LAWS überhaupt sind und wie sie sich von semi-autonomen Waffensystemen wie beispielsweise Drohnen oder Raketenabwehrsysteme unterscheiden. Zusätzlich wird auf das Risiko eines Rüstungswettlaufs hingewiesen. Staaten die nämlich nicht beteiligt wären bei der Entwicklung von LAWS, könnten sich hierdurch bedroht fühlen und eigene Entwicklungen initiieren.

Länderpositionen

Neben der zuvor genannten Stellungnahme veröffentlichen die meisten Mitgliedsstaaten zu jedem Treffen ihre eigene Position und Einschätzung zur anhaltenden Diskussion.

First, on the question of identification of key characteristics and elaboration of a working definition of LAWS in order to facilitate the debate in the GGE and avoid confusion and misunderstandings on the scope of the fully autonomous systems. [...] We note that while fully autonomous lethal weapons systems do not exist yet, the GGE should continue to consider issues related to their potential development and regulation, including the question of human control over LAWS. In this context, it is important to ensure that we do not hamper innovation in high-technology industries, in particular research and development in robotics or other related areas in the civilian sector. („EU Statement on Lethal Autonomous Weapons Systems (LAWS)“, 2017, S. 2)

Die meisten Staaten sind der Überzeugung, dass die Treffen des Sondermandats des CCW wichtig und wegweisend seien und dass das Ausbleiben einer Arbeitsdefinition von LAWS dringend geändert werden muss. So auch die oben angeführte Meinung der EU-Abgesandten. Dennoch wird gleichzeitig eingeräumt, dass sich technologischer Wandel unvorhersehbar rapide beschleunigt, sodass eine reale Abschätzung der Folgen der Entwicklung von neuen Technologien nicht ablesbar ist. Zudem scheinen einige Mitgliedsstaaten davon überzeugt zu sein, dass ein verfrühtes Einstellen der Gespräche und ein kategorisches Verbot von LAWS zu voreilig sei. Schließlich sei noch nicht abzuschätzen, ob und wie LAWS zu Zwecken von humanitären Hilfen und der weltweiten Friedenssicherung genutzt werden könnten. Die Vereinigten Staaten von Amerika beispielsweise äußern, dass der Einsatz von LAWS bei der Unterscheidung von Kombattanten und Zivilisten helfen als auch die Proportionalität erhöhen könne. Zudem unterstreichen sie, dass der Nutzen und das Potenzial von LAWS noch nicht vollständig absehbar sei und ein verfrühter Abbruch dem Verstehen von solchen Technologien im Wege stehe.

„One thing is clear: any development or use of LAWS must be fully consistent with IHL³, including the principles of humanity, distinction, and proportionality. For this reason, the United States places great importance on the weapon review process in the development and acquisition of new weapon systems.“

The United States also continues to believe that advances in autonomy and machine

³International Humanitarian Law

learning can facilitate and enhance the implementation of IHL, including the principles of distinction and proportionality. One of our goals is thus to understand better how this technology can continue to be used to reduce the risk to civilians and friendly forces in armed conflict. („Group of Governmental Experts (GGE) on Lethal Autonomous Weapons Systems (LAWS) November 13-17, 2016 Opening Statement“, 2017, S. 1f)

In diesen Punkten deckt sich die Position der Russischen Föderation:

We believe that the key problem is the absence of an approved working definition and characteristics of such systems. Today our understanding of LAWS largely depends on the interpretation of each delegation. There are states that include semiautonomous and automated systems into this category and consider that such systems already exist and are widely used. Others argue that there are no such systems. It seems that in the context of such uncertainty it is very premature to speak about any specific and even intermediate results of our work. („Statement by the Russian Delegation on Agenda Item 8 of the Meeting of the States Parties to the Convention on Certain Conventional Weapons »Consideration of the Report of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems«, 2017, S. 1)

Die Meinungen scheinen weitestgehend darin übereinzustimmen, dass die Definition von LAWS ein akutes Problem darstellt. Vorher scheint es nach den Aussagen der Mitgliedsstaaten nicht möglich zu sein, die vollständigen Gefahren als auch Potenziale herauszustellen und zu untersuchen. Diese Meinung revidiert jedoch beispielsweise Dominique Lambert in seiner Rede als geladener Experte aus dem November-Treffen 2017. Auf seine Position wird jedoch erst später Bezug genommen, denn vorher gilt es den Blick auf eine eigene Kommission der UN zu richten, welche sich im Allgemeinen mit der Ethik von Technologien auseinandersetzt.

World Commission on the Ethics of Scientific Knowledge and Technology (COMEST)

2017 veröffentlichte die COMEST einen Bericht zum Thema „Ethik der Robotik“⁴. Dieser Bericht gibt einen umfangreichen Überblick zur Debatte um den Einsatz und die Entwicklung von LAWS.

Das 64 Seiten schwere Dokument gilt als aufschlussreiches Werk zur Eröffnung einer Diskussionsgrundlage auf dem Gebiet der Ethik der Robotik. Aus Gründen des Umfangs dieser Ausarbeitung wird jedoch nur auf die Empfehlungen der Kommission Bezug genommen.

Robots will have profound effects on society and on people’s everyday lives. In order to deal with these effects in a responsible way, citizens need to be equipped with adequate frameworks, concepts, and knowledge. For that purpose, public discussions need to be organized about the implications of new robotic technologies for the various dimensions of society and everyday life, including environmental impact of the entire robot production cycle, which help people to develop a critical attitude, and which sharpen the awareness of designers and policy-makers. (*REPORT OF COMEST ON ROBOTICS ETHICS*, 2017, S. 53)

Zunächst wird in den Vordergrund gestellt, dass die Diskussion um die LAWS nicht ohne Einbeziehen der Öffentlichkeit stattfinden kann. Gerade weil der Alltag eines jeden Menschen immer

⁴Inoffizielle Übersetzung des Autors

weiter von KI und ähnlichen Technologien durchgesetzt wird, gilt es eine öffentliche Teilnahme zu schaffen. Hierzu ist es entscheidend, dass keine Informationsvakuen entstehen und Entscheidungen ohne den kritischen Blick der Öffentlichkeit abgehandelt werden. Jedoch ist es genau das, was bereits der Fall ist. Die veröffentlichten Dokumente⁵ entsprechen keinem wissenschaftlichen Standard, sodass allein das Zitieren und die Orientierung in den unterschiedlichen Dokumenten sich eher umständlich gestaltet. Außerdem wird eine am Thema interessierte Person eher von der Flut an Berichten, Stellungnahmen und Mitschnitten überfordert. Es fehlen definitiv klare und verständliche Zusammenfassungen.

Two points stand out with respect to the possible use of autonomous weapons. Legally, their deployment would violate IHL. Ethically, they break the guiding principle that machines should not be making life or death decisions about humans. (*REPORT OF COMEST ON ROBOTICS ETHICS*, 2017, S. 54)

Nach aktuellem Stand würde der Einsatz von LAWS dem humanitären Völkerrecht widersprechen. Außerdem wäre es nicht ethisch verantwortbar die Entscheidung, über Leben oder Tod, vollständig einem nicht-menschlichen Akteur zu erteilen.

With respect to their technical capability, autonomous robotic weapons lack the main components required to ensure compliance with the principles of distinction and proportionality. Though it might be argued that compliance may be possible in the future, such speculations are dangerous in the face of killing machines whose behaviour in a particular circumstance is stochastic and hence inherently unpredictable. (*REPORT OF COMEST ON ROBOTICS ETHICS*, 2017, S. 54)

Das viel genannte Argument, LAWS könnten die Proportionalität von Angriffen und die Unterscheidung zwischen Kombattant und Zivilist besser treffen, wird hier entkräftet und skeptisch beäugt. Der aktuelle Stand der Technik kann dieses Kriterium nicht erfüllen und die reine Hoffnung darauf, der technologische Fortschritt könne diese spekulative Lücke schließen, scheint nicht angemessen zu sein.

Our strong, single recommendation is therefore that, for legal, ethical and military-operational reasons, human control over weapon systems and the use of force must be retained. Considering the potential speed of development of autonomous weapons, there is an urgent need (as the ICRC has urged) “to determine the kind and degree of human control over the operation of weapon systems that are deemed necessary to comply with legal obligations and to satisfy ethical and societal considerations” (ICRC, 2016). (*REPORT OF COMEST ON ROBOTICS ETHICS*, 2017, S. 54)

Letztlich wird ausdrücklich dafür plädiert sicherzustellen, dass die Ausführung militärischer Gewalt in der Hand von menschlichen Akteuren verbleibt. Die Kommission unterstützt hierbei die Empfehlung des ICRC⁶.

⁵Zahlreiche Dokumente waren zum Zeitpunkt der Ausarbeitung dieser Schriften auf den Webseiten der UNOG nicht verfügbar und erst durch den Hinweis auf das Fehlen dieser, wurden die Dokumente zugänglich gemacht.

⁶International Committee of the Red Cross

Technik

Nachdem ein ausführlicher Überblick über den Status quo der Debatte um den Einsatz und die Entwicklung von LAWS im Rahmen dieser kurzen Ausarbeitung gegeben wurde, gilt es sich dieser Technologie im Speziellen anzunähern. Hierzu wird ein technischer Blick auf Technologien wie Software und KI geworfen und gleichzeitig ein kritischer Einwand dagegen erhoben.

Die Entwicklung von Technologien ist nicht durch diese Technologien selbst zu erklären. Warum sich eine Technologie durchsetzt, eine andere dagegen nicht, hängt viel mehr von der Konstellation der Akteure, ihren Deutungen, Interessen und Machtressourcen ab als von einem immanenten Maßstab. (Kaminski, 2010, S. 31)

Software

Software wird in Abgrenzung zur Hardware als programmierbarer Bestandteil eines jeden Computersystems benötigt. Hardware stellt die physikalischen und chemischen maschinellen Bestandteile. Software ist der auf Algorithmen basierende Aufbau zur Steuerung der Hardware, wobei die Resultate der einzelnen Berechnungen nicht zwangsläufig in der physikalischen Welt einen Effekt haben müssen. So genannter Maschinencode besteht aus Nullen und Einsen, sprich Bits. Entweder ist eine Stromspannung vorhanden oder nicht. Programmcode jeder höheren Programmiersprache muss kompiliert werden in maschinenlesbare Aneinanderreihungen von Bits.

Aus der Informatik und den sich ergebenden Qualitätsstandards für performanten und fehlerresistenten Quelltext geht hervor, dass Software modulbasiert und aufeinander aufgebaut erstellt werden soll. Dahinter steht das Paradigma des „Don't repeat yourself“⁷. Das bedeutet, dass wenn beispielsweise jemals eine Funktion zur Addition von Dualzahlen geschrieben und veröffentlicht wurde, wird diese auf Funktion von vielen anderen Teilprogrammen verwiesen und nicht nochmals entwickelt. Soweit die Theorie. Man spricht hierbei von Abhängigkeiten. Hierzu gibt es in jeder Programmiersprache ein sogenanntes *Dependency Management System*. Funktionen, Module, Pakete, Teilprogramme werden in den Quellcode des Endprodukts injiziert⁸ und werden behandelt als wäre der Inhalt dieser injizierten Funktionalität Bestandteil desselben. Was jedoch oftmals geschieht, ist Folgendes:

1. **Prämisse 1** Komplexe Software hat oftmals Abhängigkeiten von mehreren zehntausend Paketen.
2. **Prämisse 2** Diese Pakete werden oftmals gar nicht mit der Software mitausgeliefert, um Speicherkapazitäten zu sparen und werden deshalb als Referenzlisten mit angeführt.
3. **Folgt aus Prämisse 1 & 2** Daher gibt es in so gut wie jeder Software eine Liste von teils unerfüllten Abhängigkeiten.
4. **Prämisse 3** Die entsprechenden Pakete sind lediglich als Referenzen hinterlegt und werden mittels eindeutiger Versionsnummer adressiert.
5. Wenn nun ein sogenannter Bug oder auch Fehler im Programmcode eines solchen Paketes gefunden wird, kommt es zur Behebung dieses Fehlers, nämlich einem sogenannten *Patch*.

⁷kurz: DRY

⁸sog. *Dependency Injection (DI)*

Und die vorherige Version wird als *deprecated* markiert, was so viel bedeutet, dass diese nicht mehr benutzt werden sollte ohne sich damit einem Risiko auszusetzen.

6. In diesem Fall muss jede Software, die in ihrer Abhängigkeitsliste genau auf diese ausgesonderte Versionsnummer referenzierte, ebenfalls aktualisiert und angepasst werden.
7. In diesem Zuge verändert sich das Endprodukt dieser Software.
8. **Konklusion** Letztlich führt es dazu, dass es zu keinem Zeitpunkt einen abgeschlossenen oder fertigen Quellcode gibt, der der gewünschten Funktionalität entspricht, da es sich um einen Prozess handelt.

Software hat somit niemals einen abgeschlossenen Zustand. Demnach kann niemals sichergestellt werden, ob Software vollständig das gewährleisten kann, was von ihr erwartet wird. Gesagt werden sollte, dass sich Experten und Expertinnen der Informatik genau darüber absolut im Klaren sind und gerade deshalb zahlreiche Qualitätssicherungsprozesse entwickelt haben. Es handelt sich hierbei um Methoden, die gerade diese Fehleranfälligkeit und Korrumpierbarkeit reduzieren soll. Jedoch hat es in der gesamten Zeit der Informatik noch keinen Code gegeben, der nicht auf die eine oder andere Weise doch geknackt werden konnte.⁹

Die Entwicklung von Software läuft somit ständig Gefahr korrumpiert, beschädigt, angegriffen oder fehleranfällig zu sein. Solange für diese Fälle keine umfänglichen Strategien zur Risikominimierung geäußert werden, bleibt Software, die kritische Aufgaben erfüllt, eine Gefahr.

Zu diesen Fällen äußerte sich bisher keiner der Interessensvertreter zur Entwicklung von LAWS. Die Diskussion beschränkt sich derweilen auf *Hypothesen*, die *nicht* mit der Realität der Entwicklung von Software zusammenhängen. Die Liste an potenziellen Gefahren ist noch länger. Durch das oben genannte Beispiel sollte verdeutlicht werden, dass es sich hierbei um berechtigte Einwände handelt, die im Kreise der Interessensvertreter wie der UN besprochen werden sollten.

Fehlerquellen und Risiken beim Einsatz und Entwicklung von Software

Zu den typischen Fehlerquellen und Risiken gehören (1) Anwender- und Anwenderinnenfehler, (2) Entwicklungsfehler, (3) technisches Versagen beispielsweise durch Überlastung und (4) zu geringe Sicherheitsmaßnahmen.

Aufgrund der vielen Fehlerherde ist es defacto nicht umsetzbar absolut fehlerresistente Software herzustellen oder auszuliefern. Gefahrenquellen existieren während der Programmierung als auch in der Anwendung, falls die Software nicht „ordnungsgemäß“ benutzt wird, wobei hier eher die Devise gilt, dass sämtliche Fehl- oder Schadeingaben abgefangen und in einem sogenannten *Error Handling* verarbeitet werden sollten. Potenziell sollten die Sicherheits- als auch Qualitätssicherungsmaßnahmen diesen bekannten Fehlerquellen entgegenwirken. Dennoch ist es jederzeit möglich, dass diese Schwachstellen von Dritten für eigene Interessen missbraucht werden können. Nicht nur, dass die entwickelnden Personen ungewollt Fehlfunktionen einbauen, sondern auch gewollt Fehlfunktionen, sprich sogenannte *back doors*, implementiert werden können. Diese können bis es zum Ernstfall kommt unentdeckt bleiben. Zudem kommt, dass proprietäre Software, sprich

⁹In der englischen Fachsprache ist vom *Cracker* häufig die Rede. Das sind Personen, die mutwillig Computersystem aufbrechen, um sich oder andere zu bevorteilen oder anderen zu schaden. Oftmals wird diese Bedeutung mit dem Begriff *Hacker* verwechselt, welcher lediglich aus Neugierde oder dem eigenen Spieltrieb zu augenscheinlichen selben Zwecken Computersysteme aufbricht, jedoch nach dem Aufbrechen, die Eigentümer der Software informiert und nicht wie der Cracker eine Hintertür offen lässt.

kompilierter rein-ausführbarer Softwarecode, sich der externen Qualitätssicherung entzieht. Dieser lässt sich nach Auslieferung nicht evaluieren und erfordert ein hohes Maß an Vertrauen gegenüber den ausliefernden Unternehmen.

Ein weiterer, jedoch bereits erwähnter, Fall ergibt sich aus dem Abhängigkeitsgeflecht moderner Software, welche in ihren Referenzlisten abhängig von anderen Softwarepaketen ist. Demnach muss vor jeder Aktualisierung der zu benutzenden Software eine Kontrolle durchgeführt werden. Hierzu werden absichtlich Fehlfunktionen provoziert und Fehleingaben getätigt, automatisiert versteht sich, um die Funktionalität zu prüfen und an die Benutzer und Benutzerinnen auszuhändigen. Doch auch in diesem Schritt können Fehler unterlaufen.

Es zeigt sich, dass obwohl Software nicht mehr aus dem modernen Alltag wegzudenken ist, keine vollständige Sicherheit gewährleistet werden kann. Daher ist es üblich geworden bei kritischen Anwendungen ein Supportkontingent mit anzubieten, um bei Störungen unmittelbar mit dem verantwortlichen Unternehmen in Kontakt treten zu können.

Künstliche Intelligenz (KI)

KI wird grundsätzlich unterschieden in schwache und starke KI. Eine starke KI wurde bisher erforscht, aber nicht umgesetzt, da davon auszugehen ist, dass eine starke KI der Funktionsweise des menschlichen Bewusstseins in nichts nachstehen sollte. (Vgl. Searle, 1984, S. 28ff.) Es ist zwar denkbar, eine solche Technologie zu entwickeln, jedoch konnte die Möglichkeit der Realisierung einer solchen KI noch nicht bewiesen werden. Schwache KI's sind im Gegensatz dazu bereits im Einsatz und unterstützen den Menschen bei alltäglichen Angelegenheiten. Im weiteren Verlauf soll der Begriff KI nur die schwache KI umfassen.

KI's sind letztlich automatisiert entwickelte Modelle, die zur Diskriminierung, also Unterscheidung, von zwei oder mehr Tatbeständen benutzt werden. Mit Hilfe des maschinellen Lernens, also der effizienten Kopplung von Parametern aus potenziell unendlichen Informationsmatrizen, werden Modelle gebildet, die beispielsweise in der Bilderkennung dabei behilflich sein können handschriftlich geschriebene Zahlen und Buchstaben zu unterscheiden. Hierzu wird ein elektronisches Bild auf seine einzelnen Bildpunkte und seine Positionen im Bild heruntergebrochen. Daraus entsteht eine Informationsmatrix mit Parametern oder vielmehr Bildpunkten zur Helligkeit beziehungsweise Dunkelheit. Einfachheitshalber wird sich auf dieses binäre System zwischen hell und dunkel konzentriert. Entweder ist ein Pixel ausgefüllt oder nicht. Allein auf einem 100 mal 100 Pixel großen Bild aus binären Werten ergibt sich eine Menge von 2^{10000} möglichen Zuständen in der ersten Schicht. Bisherige Computersysteme versuchten letztlich alle Kombinationsmöglichkeiten zu berechnen und mit stochastischen und probabilistischen Verfahren Aufschluss darüber zu geben, um welche Zahlen oder Buchstaben es sich im Einzelfall handelt. Hierzu wurden einfache For-Schleifen und Wenn-Dann-Beziehungen benutzt um die gesuchten Zahlen zu klassifizieren.

Diese potenziell unendlich lange Verknüpfung von Wenn-Dann-Beziehungen ist ineffizient. Deshalb, so die Erwartungen an die KI, wird nach effizienteren Methoden geforscht, die ein früheres Antizipieren dieser Kombinationsmuster herauszustellen versucht. Das ist das Kerngebiet von maschinellem Lernen und dem Nutzen von neuronalen Netzen.

¹⁰Das Ergebnis dieser Rechnung hat 3011 Dezimalstellen.

Autonome und semi-autonome Waffensysteme

Wie bei der Unterscheidung von starker und schwacher KI, gibt es auch eine Unterscheidung zwischen autonomen und semi-autonomen Waffensystemen. Fälschlicherweise werden oftmals Drohnen und automatisierte Raketenabwehrsysteme mit LAWS gleichgesetzt. Jedoch haben semi-autonome Waffensysteme einen entscheidenden Unterschied zu LAWS, sie feuern nämlich zu keinem Zeitpunkt selbstständig oder wählen niemals ein Ziel ohne den Eingriff eines menschlichen Akteurs aus. Auch wenn Drohnen beispielsweise unbemannt sind, so werden diese zu einem großen Teil von Menschen gesteuert. Das Auslösen von Waffengewalt obliegt dem oder der Pilotin. Die Verantwortung liegt unmittelbar in Zusammenhang mit einer Mensch-Maschinen-Interaktion. Wenn es zu Fehlern kommt, kann die Verantwortungskette nachvollzogen werden.

Gerade hierin eröffnet sich ein Hauptproblem von LAWS: Wer gilt als verantwortlich für den Einsatz von Waffengewalt? Undurchsichtiger wird es jedoch, sobald es zu Fehlfunktionen käme. Wer hierbei die Verantwortung tragen würde, ist zu diesem Zeitpunkt ungeklärt.

Modellbildung in KI-Systemen

Die zwei Kernaufgaben eines KI-Systems bestehen darin (1) Vorhersagen zu treffen und (2) sich selbst zu optimieren. Anders als in stochastischen und probabilistischen Verfahren wird auf eine höchst effiziente Weise nach einem Unterscheidungsmodell gesucht, dass in der Lage ist Fragen eines menschlichen Akteurs zu beantworten. Mit Trainingsdaten, die bereits das Unterscheidungsmerkmal dem lernenden System offenlegt und anhand dieser das Modell gebildet wird, wird im zweiten Schritt dieses Modell geprüft und auf seine Fehlerquote getestet. Um bei dem Beispiel der Interpretation von handschriftlichen Texten zu bleiben, wird ein bereits vom Menschen klassifiziertes Set an Bildern vom Entscheidungsmodell der trainierten KI ausgewertet. Da anfänglich die hinreichenden Merkmale zur Unterscheidung von Buchstaben und Zahlen noch nicht vollständig erfasst wurden, muss das System in mehreren Zyklen trainiert werden.

Da die Kombinationsmöglichkeiten und Gewichtung der einzelnen Bildpunkte in Relation zu allen anderen Punkten eine unüberschaubare Zahl an Interpretationen zulässt, wird oftmals erwartet eine versteckte vereinfachte Erklärung der wesentlichen Unterscheidungsmerkmale herauszustellen. Es ist grundsätzlich unerwünscht, dass letzten Endes nur das KI-System in der Lage ist, die Unterscheidungen zu treffen. Vielmehr wird eine Erwartungshaltung von Nachvollziehbarkeit geweckt. Wenn beispielsweise ein System die Liquidität eines Klienten einstuft, möchte die Kreditbank ebenfalls wissen, was den liquiden vom illiquiden Klienten unterscheidet, um gleichzeitig einen Weg zur Änderung der Situation vorzuschlagen. Ansonsten bliebe das Diskriminierungsmodell jeder KI eine Blackbox, die eigenwillig ihre Unterscheidungen trifft. Glücklicherweise gibt es in der Informatik den Bereich des *Reverse Engineering*, um trainierte Systeme nachvollziehbarer zu machen.¹¹

We assume that there is a hidden model with a number of hidden, or latent, causes that interact to generate the data we observe. Though the data we observe may seem big and complicated, it is produced through a process that is controlled by a few variables, which are the hidden factors, and if we can somehow infer these, the data

¹¹Die Forschungsgemeinschaft, bestehend aus dem Fraunhofer HHI, der TU Berlin und SUTD Singapur, untersucht beispielsweise, ob es Methoden gibt nicht-lineare Vorhersagemodelle besser verstehen zu lernen. Nachzulesen auf <http://heatmapping.org/>

can be represented and understood in a much simpler way. Such a simple model can also make accurate predictions. (Alpaydin, 2016, S. 60f)

Neben der Schwierigkeit dieses versteckte Modell zu finden, entscheidet zusätzlich die epistemische Opazität über die Transparenz und Nachvollziehbarkeit der Modellbildung, was im nächsten Abschnitt erläutert wird.

Epistemische Opazität

In der Wissenschaft hat man sich daran gewöhnt, dass die Natur nicht vollständig nachvollziehbar ist und mit naturwissenschaftlichen Methoden versucht wird Phänomene zu erklären. Seit dem Auftreten der KI und dem maschinellen Lernen ist jedoch eine Methodenundurchsichtigkeit hinzugekommen, die das wissenschaftliche Arbeiten in diesen Bereichen maßgeblich verändert hat.

In many computer simulations, the dynamic relationship between the initial and final states of the core simulation is epistemically opaque because most steps in the process are not open to direct inspection and verification. This opacity can result in a loss of understanding because in most traditional static models our understanding is based upon the ability to decompose the process between model inputs and outputs into modular steps, each of which is methodologically acceptable both individually and in combination with the others. [...] Is the presence of this kind of opacity necessarily a defect? To address this, consider that within models some loss of detail can enhance understanding. (Humphreys, 2004, S. 147f)

Während die relative epistemische Opazität Hinweise darauf gibt nach einer bestimmten Zeit die Methoden nachvollziehbar zu machen, verschränkt sich die essenzielle epistemische Opazität vollständig der Nachvollziehbarkeit.

There are at least two sources of epistemic opacity. The first occurs when a computational process is too fast for humans to follow in detail. [...] Epistemic opacity of this second kind plays a role in arguments that connectionist models of cognitive processes cannot provide an explanation of why or how those processes achieve their goals. The problem arises from the lack of an explicit algorithm linking the initial inputs with the final outputs, together with the inscrutability of the hidden units that are initially trained. (Humphreys, 2004, S. 148f)

Zu nennen wäre noch, dass die epistemische Opazität nicht als Eigenschaft eines von KI-Systemen erzeugten Modells darstellt. Erwartungsgemäß wurde in diesem Kapitel gezeigt, mit welchen Problemen sich Software und KI-Systeme auseinandersetzen und im wissenschaftlichen Bereich neue Erkenntnisse ermöglicht, welche ohne diese Disziplinen der Wissenschaft entzogen geblieben wäre. Im kommenden Kapitel soll eine kritische Sichtweise zu dem vorherrschenden Technikoptimismus entgegengesetzt werden.

Technikskepsis

Die Debatte um den Einsatz und die Entwicklung von LAWS ist von zahlreichen Kontroversen durchzogen. Grundsätzlich gilt es die Folgen und Risiken aus technischer, ethischer und

gesellschaftlicher Perspektive her zu analysieren und abzuschätzen.

Technikfolgenabschätzung

Gerade im Zuge, wenn über den Einsatz und die Entwicklung neuer Waffensysteme diskutiert wird, ist es nicht zu vermeiden, sich auch in das Gebiet der Technikfolgenabschätzung (TA) zu begeben. TA wird als eigenes Forschungsgebiet innerhalb der Technikphilosophie und Techniksoziologie verortet.

In dieser Spannungssituation [zwischen naivem Fortschrittsoptimismus und Technikskepsis seit Beginn der Industriellen Revolution] existieren beide Seiten der Medaille gleichzeitig und nebeneinander. Zum einen setzt die moderne Gesellschaft weiter auf technischen Fortschritt und unterliegt dabei Systemzwängen vor allem ökonomischer Art aufgrund des extremen Wettbewerbsdrucks in der globalisierten Wirtschaft. Trotz der Ambivalenzen und dadurch erforderlichen Regulierungen wird der technische Fortschritt weiter beschleunigt, indem mehr Geld in Forschung und Entwicklung investiert wird. Die Gesellschaft ist auf technischen Fortschritt auch daher angewiesen, um die negativen Folgen der älteren Technik kompensieren zu können [...] Zum anderen ist das Bewusstsein durchaus vorhanden, dass auch Technik, die zu besten Absichten entwickelt und eingesetzt wird, in ihrem Einsatz unerwartet und nicht beabsichtigte Folgen mit sich bringen kann. [sic] (Grunwald, 2010, S. 20)

Deutlich wird, dass technologischer Fortschritt kein immanentes Moment besitzt, aus dem es seinen Antrieb gewinnt. Dieser Fortschritt kann zum einen aus einem Systemzwang resultiert, welcher dafür sorgt, dass Schäden aus älteren Errungenschaften der Technik wieder wettgemacht werden sollen, zum anderen auch partikulären Interessen vorausgehen können. Besonders in einer globalisierten Welt, in der die Verflechtung von Forschung, Wirtschaft und Gemeinwohl in einer fragilen Co-Abhängigkeit nicht mehr rückgängig zu machen scheint, aber gleichzeitig der Wille darin besteht gemeinsam zu kooperieren, um das Wohl der gesamten Menschheitsbevölkerung und Umwelt zu fördern, wird die Frage immer lauter, ob dieser Fortschritt noch Fortschritt und nicht ein unausweichliches Fortschreiten bedeutet. Jede doch so gut intendierte Errungenschaft auf den Gebieten der Wissenschaft kann gewollt oder ungewollt ein konträres Resultat zur Folge haben.

Ethische Einwände

Dominique Lambert, ebenfalls Mitglied des Expertengremiums des Sondermandats des CCW, weist auf einen bisher ungehörten Punkt hin, nämlich die Umkehrung des Definitionsversuchs.

Die ethische Bewertung von LAWS erfordert beispielsweise die Festlegung und Klärung eines allgemeinen Wertesystems, das als Grundlage für die Beurteilung dient, ob ein bestimmtes Waffensystem akzeptabel ist oder nicht.

Dies ist nicht einfach, denn in einem multikulturellen und multinationalen Kontext kann es eine Tendenz geben, die Existenz eines solchen Referenzwertesystems zu leugnen. Diese Schwierigkeit ist real, aber man könnte denken, dass es eine Lösung gibt, die eine Wertegemeinschaft um eine ziemlich allgemeine Vorstellung vom Menschen herum erzeugen könnte, von dem, was menschlich ist und was nicht. [...]

Ich denke, wir sind vielleicht nicht in der Lage, genau zu definieren, was LAWS sind. Gesetze, die aus Gründen der menschlichen Kohärenz nicht erwünscht sind, können jedoch zumindest präventiv dargestellt und definiert werden. Hier sind es nicht die Definitionen, die den ethischen Fragen und Rahmenbedingungen vorausgehen, sondern die Ethik wird die Grundlage für die negativen Definitionen von Maschinen bilden, die wir sicherlich nicht im Namen einer minimalen und gemeinsamen Vorstellung davon, was Menschlichkeit ausmacht, wollen würden. Das ist vielleicht das, was dem, was wir als signifikante menschliche Kontrolle bezeichnen, einen Sinn geben könnte, wenn es um die kritischen Funktionen von LAWS geht. [...]

Eine Ethik der gemeinsamen Menschheit kann es uns daher ermöglichen, in Abwesenheit einer Definition von bewaffneten Robotern, die es nicht gibt, eine Mindestdefinition von LAWS festzulegen, die wir nicht wollen würden! Der Schritt wäre minimal, aber gerechtfertigt, was nicht so schlimm wäre. (Aus dem Französischen übersetzt Lambert, o. J., S. 1ff)

Anstelle weiter daran zu scheitern eine akzeptable Arbeitsdefinition von LAWS zu entwickeln, schlägt er vor, zunächst einmal zu definieren, welche Auswirkungen bei dem Einsatz und der Entwicklung von LAWS grundsätzlich zu vermeiden wäre. Somit gibt er Anlass zur kritischen Selbstreflexion des seit vier Jahren tagenden Sondergremiums des CCW.

Missbrauch und Korruption

Stuart Russel, Professor für Informatik an der University of California in Berkeley und im Exptegremium des November-Treffens des Sondermandats des CCW, macht in seiner öffentlichen Stellungnahme deutlich, dass die Gefahren und Risiken bei der Entwicklung von LAWS den Vorteilen überwiegen. Als Unterzeichner und Mitinitiator des offenen Briefs „Forschungsprioritäten für stabile und wohltätige künstliche Intelligenz“ und Co-Autor von „Research Priorities for Robust and Beneficial Artificial Intelligence“ positioniert er sich klar für einen menscheitszugewandten Einsatz von KI. Er sieht in der Erforschung dieser Technologie ein enormes Potenzial zu Verbesserung der Lebensbedingungen für die gesamte (nachkommenden) Menschheit.

Finally, lethal autonomous weapons allow the extension of cyberwarfare into the physical domain, because software control systems can be infiltrated and modified. For example, a nation's autonomous weapons might be turned against its own civilian population. With no possibility of attribution to an external adversary or individual, one can imagine that the nation's government would be less popular after such an event. (Russell, 2017, S. 2)

Er unterstreicht nochmals die Gefahr der möglichen Korrumpierung und des Diebstahls einer solchen massentauglichen Waffe. Der bereits ausgefochtene Cyberkrieg, welcher zwischen den Staaten ausgeführt wird, könnte sich somit auch auf die physikalische Welt ausbreiten.

Vermeintliche Vorteile

Um sich einer vermeintlichen Polarisierung zu entziehen, gilt es ebenfalls die Gegenseite oder vielmehr die Fürsprecher und Fürsprecherinnen anzuhören. Hierzu wurde im November-Treffen des Sondermandats des CCW LTC Christopher Korpela als Vertreter des US-amerikanischen

Militärs geladen. Zwar ist nur seine dreiseitige Bildschirmpräsentation zugänglich auf der Webseite des UNOG, das sollte jedoch ausreichend dafür sein, seine Position wiederzugeben:

- All weapons systems development is driven by requirements
- There are many operational advantages of unmanned and remotely operated systems
- Autonomous systems leverage the same operational advantages of unmanned systems
- AWS offer potential humanitarian benefits, such as less collateral damage
- Weapons employment is the commander's role and responsibility
- Operational environment and context are key (Korpela, 2017, S. 2)

Seine Argumentation besticht durch seine Präzision den Kernpunkt seiner Position in einer sehr komprimierten, jedoch aussagekräftigen Weise verständlich zu machen. Demnach wird die Entwicklung von Waffensystemen erst durch ihre Anforderung, oder besser gesagt durch ihre Nachfrage bestimmt. So Korpela, haben LAWS äquivalent zu unbemannten Waffensystemen Vorteile, welche jedoch in seiner Argumentation ausbleiben. Es ist anzunehmen, dass während seines Vortrags diese genannt wurden, diese Informationen jedoch nicht der Öffentlichkeit zur Verfügung stehen. Sein Hauptargument, nämlich die Reduzierung von Kollateralschäden und die Potenzialerwartung zur Förderung von humanitären Vorteilen durch den Einsatz von tötenden unbemannten und autonom-agierenden Computersystemen, scheint keiner weiteren Prämissen zu benötigen, denn der Waffeneinsatz werde lediglich durch den oder die Befehlshaberin erteilt und von seiner oder ihrer Rolle verantwortet. Letztlich entscheiden auch die Umgebung und der Kontext über die Nutzung von LAWS.

Philosophische Rechtfertigung des Einsatzes von LAWS

Ronald C. Arkin, Professor Direktor des Mobile Robot Laboratory an der School of Interactive Computing, College of Computing am Georgia Institute of Technology veröffentlichte 2009 ein bedeutendes Werk zur Beantwortung von ethischen Fragen bezüglich der Entwicklung von LAWS.

Aus einem von ihm veröffentlichten Artikel, welches als gekürzte Version seines oben genannten Werkes zu bezeichnen ist, wird deutlich, welche Aufgabe er sich zum Ziel gemacht hat:

This paper presents the background, motivation and philosophy for the design of an ethical autonomous robotic system capable of using lethal force. The system is governed by the Laws of War and Rules of Engagement using them as constraints. It is a goal of this research, which is funded by the Army Research Office, to yield ethical performance by autonomous systems that eventually exceed that of human soldiers.

Specific design models for the implementation of this approach appear in [1] which includes an ethical governor, ethical behavioral control, an ethical adaptor and responsibility advisor. (Arkin, 2008, S. 7)

Unter Einbezug des geltenden humanitären Völkerrechts beleuchtet Arkin die unterschiedlichen Dimensionen der Entwicklung von LAWS und versucht ein ethisches Grundgerüst zu konstruieren, auf dem es gerechtfertigt wäre, diese neue Art von Waffensystemen zu verwenden und zu entwickeln. Hierzu entwirft er mittels Pseudocode [Programmier Quelltext, welcher keiner

bekannten Programmiersprache zugehörig ist, jedoch ohne große Fachkenntnis auch von Laien verstanden werden kann.] Programme, die die ethischen und regulativen Richtlinien der gerechten Kriegstheorie, Rules of Engagement und dem humanitären Völkerrecht beachten und befolgen. (Arkin, 2009, S. 132) Da grundsätzlich bei LAWS von KI ausgegangen wird und innerhalb von KI-gestützten Modellen klassische Wenn-Dann-Beziehungen weniger notwendig sind, scheint sein Pseudocode wenig aussagekräftig zu sein und unterschätzt die Komplexität des maschinellen Lernens.

This research is funded under Contract #W911NF-06-1-0252 from the U.S. Army Research Office. (Arkin, 2008, S. 7)

Zudem ist die Einordnung seiner Position aufgrund der Finanzierung durch Drittmittel durch das U.S. Army Research Office unter Umständen von den Interessen des Geldgebers gefärbt. Schließlich ist es nach heutigem Stand nicht unüblich, dass universitäre Forschungsvorhaben von externen Geldern zumindest subventioniert werden. Dennoch gilt es an solchen Stellen mit einem kritischen aber nicht feindseeligen Blick genauer hinzuschauen, wessen Interessen aus welchen Gründen bekräftigt werden. Aufgrund dieser Intransparenz und der zwar gut gemeinten technischen Herangehensweise von Arkin, bleibt offen, ob seine Argumentation zum Wohle der Menschheit oder gewissen Partikularinteressen ausgeht.

Seine Argumentation verliert hierdurch nicht an Gültigkeit oder Schlüssigkeit. Allerdings büßt sie an Glaubwürdigkeit dadurch ein, dass nicht auf eine Argumentationsgrundlage Rücksicht genommen wird, die den Aspekt der Menschlichkeit, sowie einen dementsprechenden gesellschaftlichen Fortschritt, mit einschließt. Da es sich bei der Entwicklung von LAWS um einen vorwiegend weltpolitischen Diskurs handelt, ist es geboten auch im Sinne der Menschheit zu argumentieren und sich von diskursgefährdenden Einzelinteressen freizumachen. Daher möchte ich im Folgenden einen weitaus skeptischeren Blick auf den institutionalisierten Bereich von Forschung, Universität und Wissenschaft werfen.

Philosophieskeptischer Einwand

Vittorio Hösle, Professor für Kunst und Schrift am Institut für deutsche und russische Sprachen und Literaturen als auch für Philosophie und Politikwissenschaft der University of Notre Dame, besticht durch seinen unermüdlichen Versuch, Wissenschaft und Philosophie in einem Atemzug zu vereinen, und die Institutionalisierung dieser kritisch zu hinterfragen.

So ist es beklemmend, wie wenig die Gegenwartsphilosophie zur Lösung aktueller Fragen leistet. Die Schere zwischen beschleunigter wissenschaftlich-technischer Entwicklung und stagnierendem, wenn nicht regredierendem ethischem Bewußtsein öffnet sich immer mehr, und so kann man sich nicht wundern, daß es bisher immer noch zu wenig gut begründete und zugleich detaillierte und konkrete Abhandlungen zu ethischen Fragen der technischen Zivilisation gibt. Eine ethisch-naturwissenschaftliche Doppelkompetenz ist selten zu finden, obgleich zur Lösung jener Fragen eine Überbrückung der Kluft zwischen den beiden Kulturen gefordert ist; (Hösle, 1994, S. 23)

In unserem Fall der ethischen Diskussion um den Einsatz und Entwicklung von LAWS wäre es von entscheidender Wichtigkeit Personen und Forschungsgemeinschaften aufzusuchen, die sich genau dieser Herausforderung stellen, nämlich dem Brückenschlag von KI und Ethik.

Wer eine privilegierte gesellschaftliche Position einnimmt, wer mehr Gelegenheit hatte, sich zu bilden, der hat auch mehr Pflichten: In Anbetracht dieses nahezu tautologischen Satzes bedrückt die kollektive Verdrängung der Existenzfragen der Menschheit zu dem Behufe, sich um so ungenierter mit den eigenen Partikularinteressen auseinanderzusetzen, zwar bei jedem Menschen als Ausdruck einer tiefen Verachtung des Allgemeinen, wird aber wörtlich „Bekenner der Liebe zur Weisheit“ bedeutet. Nicht die Beschäftigung mit dem Partikularen ist das eigentlich Empörende [...], sondern der Widerspruch zwischen dem Anspruch eines Berufs, der den höchsten Ideen der Menschheit gewidmet sein sollte, und der existenziellen Gleichgültigkeit gegenüber diesen Ideen, die bei einigen offenbar nur noch extrinsischen und egoistischen Motivationen bemüht werden. Die Würde des Politischen im erhabenen Sinne des Wortes, d.h. des Öffentlichen, verfallen zu sehen, wird stets schmerzen, aber besonders traurig ist es, wenn Philosophen [und Philosophinnen], die in der Theorie jenes Allgemeine vertreten sollten, dessen praktisches Pendant das Politische ist, zu Funktionären kleinlicher Partikularinteressen verkommen. (Hösle, 1994, S. 25f)

Gerade der letzte Satz verdeutlicht die zuvor geäußerte Skepsis gegenüber Arkin. Auch wenn er sich als Philosoph etablieren konnte, ist nicht ausgeschlossen, dass seine mangelnde Überzeugungskraft gerade in dem Sich-Verschließen vor dem Politischen Ausdruck findet. Die Frage bleibt offen, ob er im vollen Bewusstsein um die Auswirkungen seiner Schrift argumentiert hat oder nicht doch sich für die Einzelinteressen eines von Hierarchie gesteuerten Machtapparats wie dem Militär eingesetzt hat.

Schlussplädoyer

Aus gegebenem Anlass, dass zu diesem Zeitpunkt bereits an der Entwicklung von autonomen Waffensystemen geforscht wird, die zur Zielsetzung haben, den menschlichen Akteur aus der unmittelbaren Waffengewalt zu entfernen, ist es geboten diesen Diskurs in die breite Öffentlichkeit zu tragen. Ein Sondermandat der Vereinten Nationen tagt bereits seit 2014 regelmäßig zu dieser Thematik und stagniert schon bei der Aufgabe eine geeignete Definition dieses neuen Typus von Waffentechnologien zu finden. Nach aktuellem humanitären Völkerrecht ist es geboten, dass die Entscheidungsgewalt für den Befehl zur Tötung bei einem Menschen zu liegen hat. Demnach müsste für den potenziellen Einsatz dieser Waffen das humanitäre Völkerrecht angepasst werden. Das Stimmungsbild der Mitgliedsstaaten verrät in welcher Verunsicherung sich die Teilnehmer und Teilnehmerinnen befinden, denn grundsätzlich verweigern sich die wenigsten einem kategorischen Verbot der Entwicklung. Im Grunde will niemand das Potenzial zur Verbesserung der humanitären Lage in Kriegssituation unerforscht lassen. Anhänger und Anhängerinnen der optimistischen Seite der Verhandlungen erhoffen sich eine Reduzierung von Kollateralschäden, sprich eine Reduzierung der Gefahren für Zivilbevölkerung. Zudem wird erhofft, dass Gegenschläge oder humanitäre Interventionen mit einer höheren Proportionalität durchgeführt werden können. Das heißt so viel wie die Eliminierung oder Neutralisierung der Keimzellen der Bedrohung oder ein höheres Maß an Zielgerichtetheit der Angriffe auf das Gefahrenzentrum.

Auch wenn die potenziellen Vorteile im Einklang mit dem humanitären Völkerrecht stehen könnten, bleibt zweifelhaft, ob das gewünschte Ziel auch tatsächlich realisierbar ist und nicht ungewollterweise neue Gefahrenherde mit der Entwicklung von autonomen Waffensysteme eröffnet

werden könnten. In der internationalen Forschungsgemeinschaft lässt sich ein folgenschwerer Trend beobachten, nämlich die steigende Vernachlässigung der Begründbarkeit von einzelnen Forschungsvorhaben. Es sei nicht zu vergessen, dass in einer stark sektionalisierten und hochgradig spezialisierten Forschungslandschaft kaum noch zu überblicken ist, welche Interessen letztlich verfolgt werden und was sich im Zusammenschluss der einzelnen Forschungsergebnisse ergeben könnte. Das Ausmaß, welche Kombinationsmöglichkeiten für Einzelinteressen genutzt werden könnten, ist (1) nicht vorhersehbar und (2) nicht öffentlich offengelegt der einzelnen Interessengruppen.

Auch wenn das Sondermandat der Vereinten Nationen Expertengremien einlädt, um den Stand der Technik wiederzugeben und zu erläutern, ist es kaum vorstellbar, dass die gesamte Interessensgemeinschaft aus Befürwortern und Befürworterinnen sich einem öffentlichen Audit unterzieht und offenlegt, welche Vorhaben geplant sind und mit welchen Risiken für die Menschheit bereits zu einem frühen Stadium wie jetzt gerechnet wird. Außerdem gilt es sämtliche Gefahren in einem anhaltenden Prozess zu sammeln und zu gewichten, um den Diskurs endlich auf einer vollständigen Informationsgrundlage führen zu können.

Die Verschleierung von Partikularinteressen gilt es zu verhindern und hiergegen auch geeignete Strategien zu entwickeln, um einen ernst gemeinten Diskurs auf einer ehrlichen und offenen Kommunikationsbereitschaft aufzubauen. Gemeint ist nicht einen demokratischen Diskussionsprozess anzuleiten, welcher letztlich doch durch Vorenthaltung von wesentlichen Informationen verzerrt werden könnte, sondern vielmehr eine andere Form der ethischen Bewertung zu entwickeln, welche sich tatsächlich auf die reflektierte Meinung der Menschheit beruft. Dieses Unterfangen, so scheint es, ist hoffnungslos. So flüchtig die Meinungen einer jeden Person sind, so intransparent auch die Einzelinteressen der forschenden Wissenschaften auch sein mögen, um so ernüchternder ist die Erkenntnis, dass derartige Entscheidungen wie die Entwicklung von menscheitsvernichtenden Technologien nicht aktiv verhindert werden, sondern durch unterlassene Intervention einfach zugelassen wird.

Technologien lösen technische Probleme, Technologien kompensieren vergangene Schäden älterer Technologien, doch Technologien lösen keine sozialen oder sozioökonomischen Probleme wie Armut oder Hunger, noch nicht.

Anstelle die Energien der weltweiten Forschungsgemeinschaften für die Entwicklung autonomer Waffensysteme zu verheizen, schlage ich vor, die frei werdenden Ressourcen zusammenzuschließen, um dem Hauptziel der Vereinten Nationen gerecht zu werden und „künftige Geschlechter vor der Geißel des Krieges zu bewahren“.

Bibliographie

2014 Meeting of Experts on Lethal Autonomous Weapons Systems. (o. J.). Zugriff am 29.11.2018. Verfügbar unter: [https://www.unog.ch/___80256ee600585943.nsf/\(httpPages\)/a038dea1da906f9dc1257dd90042e261?OpenDocument](https://www.unog.ch/___80256ee600585943.nsf/(httpPages)/a038dea1da906f9dc1257dd90042e261?OpenDocument)

Alpaydin, E. (2016). *Machine learning: the new AI* (MIT Press essential knowledge series). Cambridge, MA: MIT Press.

Arkin, R. C. (2008). Governing lethal behavior: Embedding ethics in a hybrid deliberative/reactive robot architecture. In *Proceedings of the 3rd international conference on human robot interaction - HRI '08* (S. 121). Gehalten auf der The 3rd international conference, Amsterdam, The Netherlands: ACM Press. doi:10.1145/1349822.1349839

Arkin, R. C. (2009). *Governing lethal behavior in autonomous robots*. Boca Raton: CRC Press.

EU Statement on Lethal Autonomous Weapons Systems (LAWS). (2017, November 17).. Zugriff am 28.11.2018. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/8DDFEE147FB37799C125823B003FFAB0/\\$file/2017_GGE+LAWS_Statement_EU.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/8DDFEE147FB37799C125823B003FFAB0/$file/2017_GGE+LAWS_Statement_EU.pdf)

Fras, D. (2018). Angst vor dem Killer-Roboter. Frankfurt am Main.

Group of Governmental Experts (GGE) on Lethal Autonomous Weapons Systems (LAWS) November 13-17, 2016 Opening Statement. (2017, November 17).. Zugriff am 28.11.2017. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/6E9C8002759032A8C12582490031466C/\\$file/2017_GGE+LAWS_Statement_USA.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/6E9C8002759032A8C12582490031466C/$file/2017_GGE+LAWS_Statement_USA.pdf)

Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2017/WP.10. (2017, November 13).. Zugriff am 28.11.2017. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/A980151CB5E662D4C12581D80025D4F3/\\$file/2017_GGEonLAWS_WP9_NAM.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/A980151CB5E662D4C12581D80025D4F3/$file/2017_GGEonLAWS_WP9_NAM.pdf)

Grunwald, A. (2010). *Technikfolgenabschätzung: eine Einführung* (Gesellschaft, Technik, Umwelt) (Zweite, grundlegend überarbeitete und wesentlich erweiterte Auflage.). Berlin: edition sigma.

Hösle, V. (1994). *Die Krise der Gegenwart und die Verantwortung der Philosophie: Transzendentalpragmatik, Letztbegründung, Ethik* (Ethik im technischen Zeitalter) (2., um ein Nachw. erw. Aufl.). München: Beck.

Humphreys, P. (2004). *Extending ourselves: computational science, empiricism, and scientific method*. New York: Oxford University Press.

Kaminski, A. (2010). *Technik als Erwartung: Grundzüge einer allgemeinen Technikphilosophie* (Edition panta rhei). Bielefeld: Transcript.

Korpela, C. (2017, November 14). LAWS - Military Effects Panel. Zugriff am 29.11.2018. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/B0CE5EDF3E2983C1C1258249003146D8/\\$file/2017_GGE+LAWS_Statement_ChrisKorpela.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/B0CE5EDF3E2983C1C1258249003146D8/$file/2017_GGE+LAWS_Statement_ChrisKorpela.pdf)

Lambert, D. (o. J.). Intervention de Dominique Lambert. Genf. Zugriff am 29.11.2018. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/2A56AEC91DF9D07FC12582490035877C/](https://www.unog.ch/80256EDD006B8954/(httpAssets)/2A56AEC91DF9D07FC12582490035877C/)

\$file/2017_GGE+LAWS_Statement_DLambert.pdf

Leveringhaus, A. (2016). *Ethics and autonomous weapons* (Palgrave pivot). London New York, NY: Palgrave Macmillan.

REPORT OF COMEST ON ROBOTICS ETHICS. (2017).. Zugriff am 28.11.2017. Verfügbar unter: <http://unesdoc.unesco.org/images/0025/002539/253952E.pdf>

Russell, S. (2017, November 13). AI and Lethal Autonomous Weapons Systems - Remarks delivered at the meeting of the Group of Governmental Experts on Lethal Autonomous Weapons Systems, Geneva, November 13, 2017. Zugriff am 28.11.2018. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/3292E33C04873441C12582490031469D/\\$file/2017_GGE+LAWS_Statement_StuartRussel.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/3292E33C04873441C12582490031469D/$file/2017_GGE+LAWS_Statement_StuartRussel.pdf)

Searle, J. R. (1984). *Minds, brains, and science* (The 1984 Reith lectures). Cambridge, Mass: Harvard University Press.

Statement by the Russian Delegation on Agenda Item 8 of the Meeting of the States Parties to the Convention on Certain Conventional Weapons »Consideration of the Report of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems«. (2017, November 22).. Zugriff am 28.11.2017. Verfügbar unter: [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/37365361B9432DC2C125823B00418F0C/\\$file/2017_GGE+LAWS_Statement_Russia.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/37365361B9432DC2C125823B00418F0C/$file/2017_GGE+LAWS_Statement_Russia.pdf)