



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2023

**Ethische und rechtliche Herausforderungen autonomer KI-Systeme im
Sicherheitsbereich.**

Burri, Thomas ; Christen, Markus

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-255936>

Journal Article

Published Version

Originally published at:

Burri, Thomas; Christen, Markus (2023). Ethische und rechtliche Herausforderungen autonomer KI-Systeme im Sicherheitsbereich. *Stratos*, 2023(1):94-100.

Expertise

Ethische und rechtliche Herausforderungen autonomer KI-Systeme im Sicherheitsbereich



THOMAS BURRI, Universität St. Gallen
MARKUS CHRISTEN, Universität Zürich

Abstract

Systems with autonomous capabilities are increasingly becoming an important element of modern armed forces. This applies to robotic systems – whether armed or not – as well as to systems for command support or cyber defense, each of which integrates artificial intelligence (AI) technologies. There are not only technical challenges to be solved, but also ethical

and legal issues. This paper provides a brief overview of the relevant developments regarding these aspects that may arise in the use of (semi-)autonomous AI systems in the security domain. In addition, recommendations are addressed to the legislature. The paper is based on a recently published expert report on these issues.

Schlüsselbegriffe Autonome Waffensysteme; Ethik; humanitäres Völkerrecht; LAWS; Robotik
Keywords autonomous systems; ethics; humanitarian law; LAWS; robotics



PROF. DR. THOMAS BURRI ist Professor für Völkerrecht und Europarecht an der Universität St. Gallen. Er erforscht das Recht und die Ethik der Robotik und Künstlichen Intelligenz seit mehr als einem Jahrzehnt.

E-Mail: thomas.burri@unisg.ch



PD DR. MARKUS CHRISTEN ist Geschäftsführer der «Digital Society Initiative» der Universität Zürich (UZH) und leitet das «Digital Ethics Lab» am Institut für Biomedizinische Ethik und Medizingeschichte an der UZH. Seine Forschungsgebiete sind Ethik von Informations- und Kommunikationssystemen, Neuroethik und Empirische Ethik.

E-Mail: christen@ifi.uzh.ch

Informationstechnologie und Robotik werden zunehmend zu einem entscheidenden Element der modernen Kriegsführung, insbesondere, wenn Streitkräfte der Industrieländer beteiligt sind. Die Sicherheitskräfte im weiteren Sinne – nebst den Streitkräften auch die Strafverfolgungsbehörden, die Behörden des Grenzschutzes oder Zivilschutzorganisationen – stützen sich dabei ebenfalls mehr und mehr auf Robotik-Systeme mit einer gewissen Form von Autonomie.

Diese Entwicklungen werfen ethische und rechtliche Fragen auf, die von internationalen Rechtsfachleuten, der Zivilgesellschaft und auch von den Streitkräften selbst zunehmend diskutiert werden. Dieser Beitrag gibt eine kurze Übersicht über diese Debatte und macht basierend darauf Empfehlungen, wie mit den ethischen und rechtlichen Unsicherheiten, welche die Dynamik der Technologieentwicklung begleiten, umgegangen werden kann.¹ Er basiert auf einem Bericht, der im vergangenen Jahr im Auftrag von armasuisse Wissenschaft+Technologie (W+T) erstellt worden ist; der im Bericht besprochene Zusammenhang zwischen Cyberspace und Robotik-Systemen ist in diesem Artikel aus Platzgründen ausgeklammert worden.²

Voraussichtlich kein Verbot autonomer Waffen

Eine erste Entwicklung betrifft ein allfälliges Verbot autonomer Waffensysteme. Im Rahmen der «Convention for Certain Conventional Weapons» (CCW) laufen die Diskussionen über autonome Waffensysteme (*lethal autonomous weapon systems*; LAWS) nunmehr seit mehr als acht Jahren. Trotz des engen Fokus der Diskussionen auf reine Waffensysteme reichte es bisher lediglich für die Verabschiedung einiger gemeinsamer Grundprinzipien, die letztlich nicht viel mehr beinhalten als ein Bekenntnis zur Anwendung des humanitären Völkerrechts auch auf autonome Waffensysteme. Ein solches Bekennt-

«Informationstechnologie und Robotik werden zunehmend zu einem entscheidenden Element der modernen Kriegsführung, insbesondere, wenn Streitkräfte der Industrieländer beteiligt sind.»

«Auch der Einsatz autonomer Waffensysteme muss verhältnismässig sein und die Systeme müssen Kombattanten von Nichtkombattanten unterscheiden können. Bereits diese Grundprinzipien bewirken erhebliche Einschränkungen beim Einsatz.»

nis ist allerdings dem humanitären Völkerrecht ohnehin schon inhärent und ist damit in keiner Weise neu.

Dieser Mangel an aussagekräftigem Gehalt widerspiegelt die Uneinigkeit der Vertragsstaaten der CCW bezüglich der Regulierung der autonomen Waffensysteme. Länder, die führend in der Entwicklung solcher Systeme sind, stellen typischerweise die Vorteile von autonomen Waffensystemen in den Vordergrund. Länder, denen die entsprechende technologische Basis fehlt, betonen die Nachteile. Diese Positionen stehen einem konsensualen Verbot im Wege, insbesondere einem, das griffig ist und die meisten modernen bzw. absehbaren autonomen Systeme erfasst. Entsprechend ist es auch deswegen nicht überraschend, dass der Verordnungsvorschlag der Europäischen Union zur Regulierung Künstlicher Intelligenz, der aktuell in den Gesetzgebungsorganen der Union diskutiert wird, Waffensysteme vom Anwendungsbereich des Verordnungsvorschlages explizit ausklammert (abgesehen von sogenannten *dual-use*-Systemen).

Ein baldiges Verbot autonomer Waffensysteme ist also unrealistisch. Das Fehlen eines solchen Verbots hat allerdings weniger drastische Auswirkungen, als es bei einer oberflächlichen Betrachtung erscheint. Das Kriegsvölkerrecht erfasst solche Systeme grundsätzlich und beschränkt deren Einsatz mittels der allgemeinen humanitären Grundprinzipien. Auch der Einsatz autonomer Waffensysteme muss verhältnismässig sein und die Systeme müssen Kombattanten von Nichtkombattanten unterscheiden können. Bereits diese Grundprinzipien bewirken erhebliche Einschränkungen beim Einsatz. Allerdings wird dadurch eine Grundfrage überspielt, auf die auch die etablierten völkerrechtlichen Prinzipien keine unmittelbare Antwort geben: jene nach der Humanität bei der vollautonomen maschinellen Gewaltausübung, welche eng verknüpft ist mit der Wahrung der Menschenwürde.

«Allerdings wird dadurch eine Grundfrage überspielt, auf die auch die etablierten völkerrechtlichen Prinzipien keine unmittelbare Antwort geben: jene nach der Humanität bei der vollautonomen maschinellen Gewaltausübung, welche eng verknüpft ist mit der Wahrung der Menschenwürde.»

Kein Konsens in der ethischen Debatte

Auch die ethische Debatte zur Legitimität autonomer Waffensysteme – die zweite hier betrachtete Entwicklung – hat sich in den letzten Jahren deutlich intensiviert, ohne aber zu einer einheitlichen Beurteilung zu kommen. Es finden sich sowohl die Position, dass autonome Robotik-Systeme mit letalen Fähigkeiten grundsätzlich verboten sein sollte, als auch die Position, dass deren Einsatz geradezu geboten ist aufgrund des Ziels, Kriegsführung insgesamt ethischer zu gestalten. Der Beitrag von Daniel Trusilo gibt in dieser Nummer von *stratos* hierzu eine Übersicht.

Doch auch bezüglich des ethischen Diskurses lässt sich ein analoger Schluss ziehen wie bei der völkerrechtlichen Debatte: Eine einheitliche Linie für die ethische Bewertung von LAWS mag aktuell nicht bestehen; doch das hat ebenfalls keine zwingend negativen Folgen. Festzuhalten ist, dass ein ethischer Minimalkonsens dahingehend besteht, dass Entscheidungen, die durch autonome Robotik-Systeme getroffen werden, «nachvollziehbar» und in einem gewissen Sinn «steuerbar» sein sollen. Dies definiert sowohl technische Anforderungen an die Systeme, hat aber auch Implikationen bezüglich einer praktischen Ethik: Gefördert werden damit Instrumente zur ethischen Bewertung von neu entstehenden KI- und Robotik-Technologien, die relativ offen und flexibel sind und den Dialog zu diesen Themen auch ausserhalb der Ethik fördern.

Autonome Systeme vermischen Kriegs- und Friedensrecht

Eine dritte relevante Entwicklung betrifft eine durch die Nutzung von KI- und Robotik-Technologie induzierte Verwischung der Grenze zwischen Kriegs- und Friedensrecht. Zwar gilt weiterhin der Grundsatz, dass das Kriegsrecht, einschliesslich des humanitären Völkerrechts, Anwendung im bewaffneten Konflikt findet, während das Friedensrecht ausserhalb von bewaffneten Konflikten, also etwa in den Fällen von Naturkatastrophen oder Demonstrationen, anwendbar ist. Doch bereits

das Recht selbst verwischt diese Trennlinie, indem es vorgibt, dass das Friedensrecht auch im bewaffneten Konflikt zu respektieren ist, wenn auch im Rahmen des Kriegsrechts und mit den Modifikationen, die der bewaffnete Konflikt bedingt.

Das Problem für die im Sicherheitsbereich eingesetzten Robotik- und KI-Systeme ist nun, dass die Abgrenzung der Kategorie «Waffensysteme für Kriegszeiten» gegenüber jener von «Systemen für Friedenszeiten» nur scheinbar eindeutig ist. Es besteht eine hohe faktische Durchlässigkeit zwischen diesen. Dies hat mit dem inhärenten *dual-use-Charakter* dieser Systeme zu tun. So können Robotik-Systeme, die über keinerlei Bewaffnung verfügen und nur im zivilen Bereich zum Einsatz kommen sollten (also etwa beim Katastrophenschutz), häufig mit geringem Aufwand so umgestaltet werden, dass sie als Waffensysteme zu gelten haben, etwa indem zusätzlich eine Waffe auf eine Plattform montiert wird. Die aktuelle Nutzung ziviler Drohnen in Kampfhandlungen in der Ukraine (z. B. als «Transporteure» von Handgranaten über feindliche Schützengräben) verdeutlicht das.

Aufgrund dieser Überschneidungen und der Durchlässigkeit sind daher Robotik-Systeme, die im Sicherheitsbereich zum Einsatz kommen sollen, im Hinblick auf die Einhaltung sowohl des Friedensrechts wie des Kriegsrechts zu betrachten, zu beurteilen und zu prüfen. In der Entwicklungsphase ist insbesondere auf mögliche Modifikationen von Systemen bzw. auf Vorkehrungen zur Verhinderung solcher Modifikationen zu achten. Ein entsprechendes Vorgehen empfiehlt sich auch für Robotik-Systeme, die über keinerlei Form von Autonomie verfügen.

Regulierungsversuche durch Ethik und Recht

Als vierter Trend ist die zunehmende Regulierung von Robotik-Systemen zu nennen, die sich allerdings auf unterschiedlichen Dimensionen ausprägt. Für Waffensysteme schreibt das humanitäre Völkerrecht in Artikel 36 des 1. Zusatzprotokolls zu den Genfer Konventionen (Zusatzprotokoll I 1977) eine Prüfung vor, ehe sie im bewaffneten Konflikt eingesetzt werden können. Dieser Waffenprüfung, die im Schweizer Recht von der Materialverordnung (s. Artikel 11 Materialverordnung VBS) implementiert wird, müssen auch bewaffnete Robotik-Systeme unterzogen werden.

Die *Weapons Review* nach Artikel 36 hat allerdings lediglich das Ziel sicherzustellen, dass das humanitäre Völkerrecht beim Einsatz des zu prüfenden Waffensystems eingehalten werden kann. Im Gegensatz dazu kann die Prüfung von Robotik-Systemen für Einsätze in Friedenszeiten im weitesten Sinne, etwa von Kraftfahrzeugen, umfassender ausfallen. Technische Standards sowie gesetzliche Vorgaben finden regelmässig Anwendung und enthalten detaillierte Regelungen (zu den technischen Standards siehe Gustafsson 2020; in Bezug auf autonome Systeme: Danks & London 2017).

«Allerdings steckt die Regulierung von fortgeschrittenen Robotik-Systemen – seien sie nun auf den Einsatz im bewaffneten Konflikt oder in Friedenszeiten ausgerichtet – noch in den Kinderschuhen, insbesondere wenn die Systeme über Autonomie verfügen oder Künstliche Intelligenz zur Anwendung kommt.»

Allerdings steckt die Regulierung von fortgeschrittenen Robotik-Systemen – seien sie nun auf den Einsatz im bewaffneten Konflikt oder in Friedenszeiten ausgerichtet – noch in den Kinderschuhen, insbesondere wenn die Systeme über Autonomie verfügen oder Künstliche Intelligenz zur Anwendung kommt. Im Hinblick auf solche fortgeschrittenen Systeme ist das Recht erst im Begriff, Vorschriften zu entwickeln. Dem zwischenzeitlich bestehenden Bedürfnis nach Leitlinien kommt aktuell primär die sich relativ rasch und flexibel entwickelnde Ethik nach, die mitunter sehr konkrete Gestalt annehmen kann, wie das von den Autoren dieses Artikels zusammen mit dem Schweizer Drohnen- und Robotik-Zentrum (SDRZ VBS) entwickelte Beurteilungsinstrument zeigt.³ Allerdings bleibt die Ethik häufig auf eher abstrakter Ebene – siehe nur die Myriade von mittlerweile bestehenden ethischen Grundprinzipien zu Künstlicher Intelligenz.⁴ Trotz ihrer Abstraktheit liefern die ethischen Grundprinzipien den Nutzerinnen und Nutzern in Abwesenheit rechtlicher Vorgaben immerhin einen moralischen Kompass. Die diesbezügliche Ethik ist allerdings zuweilen schwer von eher prospektiven rechtlichen Abhandlungen auf der Basis von allgemeinen Rechtsprinzipien und Menschenrechten zu unterscheiden (s. z. B.

Toronto Declaration 2018 und Montreal Declaration 2018).

Doch das Recht bleibt in der Regel nicht in den Kinderschuhen stecken. Es holt nach einer gewissen Anlaufzeit auf (siehe CCW GGE 2020, Vorschlag KI Verordnung). Dies führt zu einer Verrechtlichung der bis anhin primär von der Ethik gelenkten Materie (z. B. Vöneky 2020). Durch das sich entwickelnde, aufholende Recht werden allerdings ethische Instrumente nicht überflüssig. Einerseits belässt das Recht häufig Spielraum, bei dessen Ausnutzung ethisch-moralische Richtlinien weiterhin hilfreich sein können – etwa, wenn das Recht lediglich Grundprinzipien vorgibt oder eine Sorgfaltpflicht konkret ausdifferenziert werden muss, ohne dass das Recht spezifische Vorgaben macht. Andererseits können solche Richtlinien auch nützlich sein, wenn das Recht gewisse Aspekte bewusst von Regulierung freihält, was beispielsweise im Kriegsrecht durchaus vorkommt (etwa bezüglich des Einsatzes von Nuklearwaffen).

Komplexe Mensch-Maschine-Interaktionen

Eine fünfte Entwicklung hat einen direkten Bezug zum komplexen Regulierungsproblem autonomer KI- und Robotik-Systeme im Sicherheitsbereich. Entgegen dem medial oft verbreiteten Narrativ einer Ersetzung menschlicher Tätigkeiten durch Maschinen ist von einer zunehmend komplexer werdenden Interaktion zwischen Mensch und Maschine auszugehen. Im Sicherheitsbereich ist zu erwarten, dass eine zunehmende Vielfalt von Robotik-Systemen mit breitem Auftrags- und Funktionsspektrum und in engem Zusammenspiel mit dem Menschen zum Einsatz kommen wird. Die Einbettung von Robotik-Systemen in bestehende Kommandostrukturen, Einheiten und Verbände mit dem Ziel, neue Fähigkeiten in diese Organisationen zu integrieren, wird ebensolche komplexe Mensch-Maschine-Interaktionen zur Folge haben. Die Vorstellung einer eindimensional gerichteten Interak-

«Entgegen dem medial oft verbreiteten Narrativ einer Ersetzung menschlicher Tätigkeiten durch Maschinen ist von einer zunehmend komplexer werdenden Interaktion zwischen Mensch und Maschine auszugehen.»

tion zwischen Roboter und Operateur bzw. Operateurin wird dieser Entwicklung immer weniger gerecht; insbesondere dann, wenn das Robotik-System Lernfähigkeit aufweist. Zu erwarten ist eine graduelle Rollenverschiebung der menschlichen Entscheidung von der Teleoperation zur beaufsichtigenden Kontrolle im Rahmen vorbestehender (militärischer bzw. polizeilicher) Verantwortlichkeiten und Kommandobeziehungen.

Diese Entwicklungen gehen einher mit der Befürchtung eines Kontrollverlustes, was den Nährboden für die Forderung nach «sinnvoller menschlicher Kontrolle» (*meaningful human control*) bereitet hat, insbesondere für Fälle, in denen tödliche Gewalt ausgeübt wird (also im Fall von *Lethal AWS*). Der Kontrollbegriff wird dabei in der Regel abgestuft verwendet (Altman 2019): Ein Mensch ist *in the loop*, wenn er oder sie den Angriff befiehlt; wenn das Waffensystem weitgehend selbstständig agiert, während der Mensch die Aktion nur beaufsichtigt und nötigenfalls eingreifen kann; gibt es keine Aufsicht oder Einflussnahme, ist der Mensch *out of the loop*.

Inwiefern eine «Aufsicht» autonomer Systeme aber praktisch umsetzbar ist, darf heute bezweifelt werden. Eine aktuelle Untersuchung von 40 Regelwerken zur Erreichung einer solchen «menschlichen Aufsicht» fand hierzu zwei grundlegende Schwächen (Green 2022): Erstens gibt es empirische Anhaltspunkte dafür, dass Menschen gar nicht in der Lage sind, die gewünschten Überwachungsaufgaben zu erfüllen. Zweitens legitimieren Richtlinien zur menschlichen Aufsicht den Einsatz fehlerhafter und umstrittener Algorithmen durch Regierungen, ohne die grundlegenden Probleme solcher Instrumente anzugehen. Anstatt Menschen vor den potenziellen Schäden algorithmischer Entscheidungsfindung zu schützen, vermitteln solche Richtlinien ein falsches Gefühl der Sicherheit bei der Einführung von Algorithmen und ermöglichen es Anbietern und Behörden, sich der Verantwortung für algorithmische Schäden zu entziehen.

Diese Beobachtungen zeigen, dass die Begriffe «Kontrolle» und «Aufsicht» autonomer Systeme mehrdimensional gedacht werden müssen, unter Einbezug der verschiedenen Autonomiedimensionen technischer Systeme. Dabei ist zu bedenken, dass es bereits im Fall von herkömmlichen Systemen, die vordefinierten Regeln folgen und nicht auf Lernfähigkeit basieren,

fraglich ist, ob der Output durchwegs vorhersehbar ist (Altman 2019). Das Ergebnis der Interaktion zweier getrennt programmierter Systeme lässt sich oft nicht vorhersagen. Im Falle von Systemen, die mit maschinellem Lernen und systembedingter Intransparenz arbeiten, kann sich diese Problematik potenzieren.

Ausweitung des Fokus weg von reiner Gewaltanwendung

Hier ist schliesslich auf eine sechste Entwicklung hinzuweisen: Auch wenn die Befürchtungen von Kontrollverlust vorab durch bewaffnete autonome Systeme genährt wurden, wird der Fokus langsam ausgeweitet, zumal das beschriebene «Aufsichtsproblem» sich nicht nur bei der Anwendung tödlicher Waffengewalt im bewaffneten Konflikt stellt. Die Streitkräfte und Sicherheitsbehörden haben begonnen, anstelle der Autonomie von Robotern, die gedanklich eng mit Gewaltanwendung verknüpft ist, das Thema der Künstlichen Intelligenz in den Blick zu nehmen. Damit einher geht eine gewisse Abkehr von der Gewaltanwendung, denn die Künstliche Intelligenz hat – zumindest gegenwärtig, wohl aber auch in näherer Zukunft – eher vorbereitenden, empfehlenden Charakter. Sie zielt nicht auf möglichst hohe Eigenständigkeit, sondern wirkt unterstützend, im Verbund mit menschlichen Operateuren bzw. Operateurinnen. Nur im Rahmen der unmittelbaren Gefahrenabwehr zu Verteidigungszwecken mittels autonomer Systeme (siehe z. B. das Iron Dome System) mag die Situation anders gelagert sein.

Dies verlangt, dass die Rolle der Robotik bzw. der Künstlichen Intelligenz im Sicherheitsbereich in differenzierterem Lichte als jenem der reinen Gewaltanwendung durchleuchtet und geklärt werden muss. Dem für Künstliche Intelligenz typischen Problem der Diskriminierung mittels algorithmischer Systeme (*bias*), der Anfälligkeit dieser Systeme sowie ihrer Verständlichkeit sollte grössere Aufmerksamkeit zuteilwerden.

Was für den Verteidigungsbereich im eigentlichen Sinne gilt, erscheint auch für den Einsatz von Robotik zu Friedenszeiten angebracht, denn bei dieser Verwendung macht die Fokussierung auf Gewaltanwendung noch weniger Sinn. Sie ist aber auch hier weit verbreitet. Im Rahmen des autonomen Fahrens zum Beispiel absorbierte das Trolley-Problem⁵, das letztlich eine besondere Spielart der Gewaltausübung durchexerziert,

viel Aufmerksamkeit (Bonneton et al. 2016, Awad et al. 2018, Kochupillai et al. 2020) – obwohl das autonome Fahren seiner Anlage nach gerade darauf ausgerichtet ist, *keine* Gewalt auszuüben bzw. weniger Gewalt als bis anhin beim Kraftfahrzeugbetrieb. Es dürfte im Friedensbereich *a fortiori* Sinn machen, die Diskussion im Lichte der unmittelbar relevanten, wenn auch vielleicht weniger spektakulären Problemstellung zu führen. Man denke zum Beispiel an etwaige Diskriminierungen beim Einsatz der Robotik im Rahmen der Katastrophenhilfe oder an einen möglichen Mangel an Erklärbarkeit gewisser Handlungsmuster. ◆

Endnoten

- 1 Danksagung: Diese Arbeit wurde von armasuisse W+T via Schweizer Drohnen- und Robotik-Zentrum des Departements für Verteidigung, Bevölkerungsschutz und Sport (VBS) und dem Nationalen Forschungsprogramm 77 «Digitale Transformation» des Schweizerischen Nationalfonds unter Projektnummer 407740_187494 unterstützt.
- 2 Siehe dazu auch: Sicherheitsrobotik für die Schweiz: <https://www.ar.admin.ch/de/ueber-armsuisse/medien.detail.nsb.html/92203.html>. Der Bericht ist verfügbar auf SSRN: <https://ssrn.com/abstract=4164188>
- 3 Hintergründe zu diesem Beurteilungsschema liefert: Christen et al. (2017); angewendet wurde es in Trusilo und Burri (2021).
- 4 AlgorithmWatch unterhält eine Datenbank der ethischen Prinzipien bezüglich künstlicher Intelligenz: <https://inventory.algorithmwatch.org/> (zuletzt besucht am 8. März 2023); vergleiche auch mit Jobin et al. 2019, Fjeld et al. 2020.
- 5 Das «Trolley-Problem» ist ein bekanntes Gedankenexperiment der Ethik. Skizziert wird die Situation, dass eine ausser Kontrolle geratene Strassenbahn auf fünf Gleisarbeiter zurast, die dies nicht bemerken und durch die Bahn getötet würden – ausser ein Beobachter würde die Bahn auf ein Nebengeleise umleiten, auf der allerdings dann ein Arbeiter getötet würde. Die sich stellende Frage ist: Ist es moralisch gerechtfertigt, dass der Beobachter die Strassenbahn umleitet, dadurch fünf Menschen rettet, aber direkt den Tod eines anderen Menschen verursacht? Das Gedankenexperiment existiert in zahlreichen Varianten.

Empfehlungen

1. Da sich bisher nur ein minimaler rechtlicher und ethischer Konsens über autonome Robotik-Systeme im Sicherheitsbereich herauskristallisiert hat, ist es notwendig (und möglich), eine ethische Orientierung zu geben. Fundierte ethische Rahmenwerke können Orientierung bieten, zur Schadensminimierung beitragen, die menschliche Autonomie bekräftigen und grundlegende Überlegungen zur Gerechtigkeit in Robotik-Systemen berücksichtigen. Wir empfehlen, solche ethischen Rahmenwerke weiterzuentwickeln, zu stärken und universell zu verbreiten.
2. Wir empfehlen eine praxis- und anwendungsorientierte Form der Ethik. Das Potenzial ethischer Bewertungen, die sich an konkreten Anwendungsfällen orientieren, muss voll ausgeschöpft werden. Die Praxisethik muss die Robotikentwicklung normativ begleiten und wo nötig begrenzen.
3. In dem Masse, wie sich die Gesetzgebung weiterentwickelt, müssen die bestehenden Instrumente für die ethische Bewertung von Robotik-Systemen angepasst werden, um neu geschaffene oder geänderte Gesetze zu berücksichtigen und zu integrieren.
4. Bei der Entwicklung von Robotik-Systemen für den Einsatz im Sicherheitssektor sollte sichergestellt werden, dass die Gesetze für bewaffnete Konflikte und das in Friedenszeiten geltende Recht eingehalten werden.
5. Ethische und rechtliche Bewertungen sollten regelmässig Teil des Lebenszyklusmanagements von Robotik-Systemen sein. Als Teil eines permanenten Überwachungsprozesses müssen ethische und rechtliche Bewertungen regelmässig und gelegentlich wiederholt werden, wenn die Umstände dies erfordern.
6. Neue Robotik-Systeme, die für den Sicherheitssektor entwickelt werden, sollten so früh wie möglich im Lichte der rechtlichen und ethischen Aspekte bewertet werden. Agenturen, die neue Robotik-Systeme beschaffen wollen, sollten ihren Einfluss in diesem Sinne geltend machen.
7. Die rechtliche und ethische Diskussion über Robotik-Systeme sollte ausgeweitet werden. Sie sollte widerspiegeln, dass Robotik-Systeme über die Anwendung von Gewalt hinaus weiteren Zwecken dienen und als zusammengesetzte Systeme vorgelagerte, nachgelagerte und ergänzende Funktionen im Zusammenhang mit der Anwendung von Gewalt erfüllen. Auch das geltende Recht und die Ethik stossen an ihre Grenzen. Daher ist ein pragmatischer Ansatz erforderlich, der unrealistische Erwartungen an ein baldiges Verbot der autonomen Gewaltanwendung vermeidet.

Literaturverzeichnis

- Altmann, Jürgen (2019). «Autonomous Weapons Systems – Dangers and Need for an International Prohibition», in Christoph Benzmüller und Heiner Stuckenschmidt (eds), *KI 2019, 11793 Lecture Notes in Artificial Intelligence*, 1–17
- Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon und Iyad Rahwan (2018). «The Moral Machine Experiment», *Nature* 563, 59–64
- Bonnefon, Jean-François, Azim Shariff und Iyad Rahwan (2016). «The Social Dilemma of Autonomous Vehicles», *Science* 352(6293), 1573–1576
- Burri, Thomas, und Markus Christen (2022). «Rechtliche und ethische Dimensionen der Robotik im Sicherheitsbereich / Legal and ethical dimensions of robotics in the security sector» (Juli, 2022). Universität Zürich, Digital Society Initiative. Verfügbar auf SSRN: <https://ssrn.com/abstract=4164188>
- Christen, Markus, Thomas Burri, Joseph Chapa, Raphael Salvi, Filippo Santoni de Sio, und John Sullins (2017). «An Evaluation Schema for the Ethical Use of Autonomous Robotic Systems in Security Applications» (November 1, 2017). Available at SSRN: <https://ssrn.com/abstract=3063617>
- Danks David und Alex John London (2017). «Regulating Autonomous Systems: Beyond Standards», *IEEE Intelligent Systems* 32(1), 88–91
- Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy und Madhulika Srikumar (2020). «Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI», available online: <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420>
- Green, Ben (2022). «The Flaws of Policies Requiring Human Oversight of Government Algorithms», *Computer Law & Security Review*, 45, available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3921216
- Gustafsson, Ingrid (2020). «How Standards Rule the World: The Construction of a Global Control Regime», Elgar
- Jobin, Anna, Marcello Ienca und Effy Vayena (2019). «The global landscape of AI ethics guide lines», *Nature Machine Intelligence* 1, 389–399
- Kochupillai, Mrinalini, Christoph Luetge und Franziska Pozzler (2020). «Programming Away Human Rights and Responsibilities? «The Moral Machine Experiment» and the Need for a More «Humane» AV Future», *Nanoethics* 14, 285–299
- Montreal Declaration for a Responsible Development of Artificial Intelligence (2018). Available online: https://docs.wixstatic.com/ugd/ebc3a3_c5c1c196fc164756afb92466c081d7ae.pdf
- The Toronto Declaration (2018). «Protecting the right to equality and non-discrimination in machine learning systems». Available online: https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf
- Trusilo, Daniel und Thomas Burri (2012). «The Ethical Assessment of Autonomous Systems in Practice», 4 (MDPI) J 4, 749–763
- Vöneky, Silja (2020). «Key Elements of Responsible Artificial Intelligence – Disruptive Technologies, Dynamic Law», *Ordnung der Wissenschaft* 1, 9–22