## Künstliche Intelligenz in Grenzen halten

In den Unterhaltungsmedien wird des Öfteren eine Zukunft aufgezeigt, in der die Menschen die Kontrolle über einen Algorithmus bzw. eine Technologie verloren haben und dieser Zustand negative Folgen für die Menschheit hat. Allerdings setzen sich längst nicht nur Unterhaltungsmedien mit der unkontrollierten sowie unregulierten Entwicklung neuer Technologien sowie deren Gefahren auseinander. Zunehmend melden sich hochrangige IT-Experten und Forscher zu Wort und warnen vor der rücksichtslosen Entwicklung neuer Technologien. Es wird im besonderen Maße auf die Gefahren eingegangen, die mit der Entwicklung von künstlicher Intelligenz einhergehen (vgl. Brundage et al. 2018, S. 3ff). Aus diesem Grund befasst sich der folgende Artikel zum einen mit der Frage, welche spezifischen Probleme in näherer Zukunft durch den unkontrollierten Einsatz von künstlicher Intelligenz entstehen können, und zeigt zum anderen auf, welche Lösungsvorschläge zum Eindämmen bzw. Eliminieren dieser Gefährdung vorhanden sind.

Als Schwellenpunkt, an dem die KI als größte Bedrohung gesehen werden kann, ist der Eintritt der sogenannten Singularität. Unter einer technischen Singularität versteht man den Punkt an dem eine Superintelligenz entsteht, die den Menschen in allen Kategorien übertrifft (vgl. Mainzer 2016: S. 208ff). Zwar ist noch nicht vollkommen eindeutig ob oder viel mehr wann dieser Punkt erreicht werden wird und was nach diesem Punkt für eine Gefahr für die Menschheit entsteht, allerdings wird dieses mögliche Ereignis sehr kritisch von führenden Experten betrachtet, da ein intelligenteres Wesen die Menschen unterdrücken könnte (vgl. Dörner 2017). Es ist nicht von der Hand zu weißen, dass die aktuelle Technologie höchstwahrscheinlich noch Jahrzehnte von einer Superintelligenz, falls diese eintreten sollte, entfernt ist (vgl. Grace et al. 2018, S. 1). Allerdings übertreffen spezielle Algorithmen bereits heutzutage Menschen in vielen Tätigkeiten, wie beim Diagnostizieren von bestimmten Krebsarten oder beim Spielen des hochkomplexen Spiels GO (vgl. Sauer und Braun 13.09.18; vgl. DeepMind). Besonders der Fall AlphaGo, der Name des Algorithmus, welches das Spiel Go beherrscht und von der Firma DeepMind entwickelt wurde, zeigt die rasante und exponentielle Entwicklung von künstlichen Intelligenzen auf (vgl. DeepMind). Die erste Version des Algorithmus AlphaGo hat mehrere Monate gebraucht um mithilfe von menschlicher Dateneingabe das Spiel Go zu lernen, um den menschlichen Weltmeister dieser Disziplin zu schlagen (vgl. DeepMind). Einige Monate nach dem Ereignis, wurde der neue verbesserte Algorithmus AlphaGo Zero vorgestellt. Hat die Vorgängerversion noch Monate gebraucht um ein Weltmeisterlevel zu erreichen, schaffte AlphaGo Zero das in nur 72 Stunden. Zwei Monate danach wurde wiederum ein neuer Algorithmus namens AlphaZero vorgestellt, der nicht nur innerhalb von 8 Stunden auf dem Weltmeisterlevel war, sondern sich zudem noch zwei weitere Spiele parallel beigebracht hat (vgl. DeepMind). Mithilfe dieses Beispiels sowie dem Fakt, dass sich die Rechenleistung von KI-Systemen alle dreieinhalb Monate

verdoppelt, lässt sich die rasante Entwicklung der Algorithmen aufzeigen (vgl. Stöcker 20.05.18). So mag es zwar nach Expertenschätzungen noch Jahrzehnte dauern bis die Singularität eintritt, allerdings werden die Algorithmen stetig und mit zunehmender Geschwindigkeit verbessert, wodurch der Menschheit neue mächtige Werkzeuge an die Hand gegeben werden. Diese Werkzeuge können neben förderlichen Aktivitäten allerdings auch für ethisch verwerfliche sowie allgemein schädigende Taten eingesetzt werden (vgl. Brundage et al. 2018, S. 3ff). Der Bereich, bei dem der Einsatz von KI wohl am kritischsten diskutiert wird, ist der des Militärs. Zwar ist die Menschheit schon im Besitz von halbautonomen Waffen, die nach Zieleingabe und Abschuss komplett autonom ihre Entscheidungen und Folgeschritte durchführen und der Mensch somit "out of the Loop" ist, allerdings entscheidet immer noch der Mensch zuvor, welches Ziel wann abgeschossen werden soll (vgl. Schaub und Wenzel Krisoffersen 2017, S. 14f). Bei vollautonomen Waffen ist der Mensch von der Zielerfassung bis zum Einschlag bzw. Angriff der Waffe "out of the Loop", d.h. die Waffe entscheidet selbst, welches Ziel sie wann und wie angreift, ohne weitere Befehle von einem Menschen zu benötigen (vgl. Schaub und Wenzel Krisoffersen 2017, S. 5). Die Entwicklung hin zu autonomen Waffen wird sehr kontrovers betrachtet, da etwa nicht geklärt ist, wer die Verantwortung über Entscheidungen der autonomen Waffen tragen muss, oder ob ein Algorithmus über Leben und Tod entscheiden darf (vgl. Prössl 27.08.17). Neben vollautonomen Waffen verändert der Einsatz von KI die Bereiche des Datenschutzes sowie die Überwachung der Bevölkerung. Das Internet und dessen Regeln werden maßgeblich von Tech-Giganten geformt und beeinflusst. Die Staaten und deren Regierungen tun sich schwer mit der Geschwindigkeit neuer Innovationen mitzuhalten und diese durch Gesetzesentwürfe zu regulieren (vgl. Meyer 28.11.16). So konnten die Unternehmen in den unregulierten Märkten bis vor kurzem jahrelang Daten der Benutzer ohne deren Einwilligung speichern und verwenden (vgl. Martin-Jung 2010). Durch den Einsatz von besseren und schnelleren Algorithmen können die Daten der Bevölkerung abermals schneller und effektiver ausgelesen und verwertet werden (vgl. Conrad 2017, S. 742f). Durch den zusätzlichen Einsatz von KI-gestützten Systemen, wie etwa die aktuelle erprobte Gesichtserkennung auf den Straßen Chinas, werden die Menschen zum gläsernen Bürger dessen Privatsphäre zunehmend eingeschränkt wird (vgl. Ankenbrand 09.04.18). Eine weitere Gefahr, die mit der Sammlung von Daten und verbesserten Algorithmen einhergeht, ist die Beeinflussung der politischen Meinung der Bürger. (vgl. Meyer 28.11.16; vgl. Breuer 01.03.18). Durch die Manipulation, etwa über Social Media, könnten die Grundprinzipien einer Demokratie ausgenutzt werden, um Wahlen zugunsten der Personen ausgehen zu lassen, die den effektivsten Algorithmus besitzen. Zudem birgt KI eine Gefahr für die digitale Sicherheit von Unternehmen und einzelnen Individuen durch verbesserte Hacks oder Kopierung menschlicher Identitäten (vgl. Brundage et al. 2018, S. 6). Einen weiteren Bereich, auf den die Entwicklung besserer Algorithmen bzw. einer besseren künstlichen Intelligenz starken Einfluss haben wird, ist der des Arbeitsmarktes. Zwar werden bereits heutzutage Arbeitsstellen durch Algorithmen ersetzt,

allerdings ist noch nicht klar, wie sich die allgemeine Marktsituation verändern wird, da auch neue Arbeitsplätze geschaffen werden (vgl. Mewes 10.03.18; vgl. Buxmann und Schmidt 2018, S. 30f). Jedoch könnte dieser Veränderungsprozess bei unregulierter Durchführung negative Folgen, wie Arbeitslosigkeit, für die Bevölkerung nach sich ziehen. Um diese Problemfelder anzugehen, gibt es allgemeine Lösungsansätze, die nicht nur versuchen die Gefahren einer eintretenden Singularität zu minimieren, sondern zusätzlich versuchen, die Probleme, die sich durch die stetige Verbesserung der KI ergeben, frühzeitig einzudämmen. Betrachtet man die aufgezeigten Bereiche lässt sich erkennen, dass die Probleme meist ethische, gesellschaftliche und rechtliche Fragen aufwerfen, für die es Antworten benötigt. Aus diesem Grund bedarf es strikte Richtlinien und Regelungen, die den Unternehmen und Staaten einen Kompass geben, an dem sie sich bei ihren Entscheidungen und dem Umgang mit KI orientieren können (vgl. Erdélyi und Goldsmith 2018). Dies fordern Politikexperten sowie IT-Unternehmen, wie Microsoft, die besonders auf den Schutz der Mitarbeiter und deren Arbeitsplätze eingehen (vgl. Boyle 10.04.18; vgl. Bass 18.01.18). Auch die European Commission's Initiatives in Artificial Intelligence hat bereits grobe Richtlinien für die Entwicklung und Anwendung von künstlicher Intelligenz verschriftlicht (Huet). Dabei wird z.B. auf die Notwendigkeit von einer geprüften Sicherheit der Algorithmen sowie auf den Datenschutz eingegangen (Huet). Zwar wird über die Pflicht von Richtlinien immer häufiger diskutiert, allerdings gibt es zum heutigen Zeitpunkt noch wenige konkrete Gesetze, die formuliert worden bzw. in Kraft getreten sind. Eine Organisation, die bereits konkretere Gesetze zur Digitalisierung sowie den Einsatz intelligenter Algorithmen veröffentlicht hat, ist die Digital Charta (vgl. Digital Charta). In Artikel 5 wird etwa festgelegt, dass nur Menschen und keine Algorithmen über ethische Thematiken oder sogar Leben entscheiden dürfen (vgl. Digital Charta). Hierbei handelt sich aber wiederum vielmehr um Initiatoren aus Politik und Wirtschaft, die eine Beispielform der digitalen Grundrechte formuliert haben, als um Regierungsvertreter, die aktiv Gesetze verabschieden. Ein Bereich, in dem bereits konkretere Abkommen durch Verhandlungen mit Regierungsvertretern festgelegt worden sind, ist der des Militärs. In dem bereits 5. Treffen der Gruppe der Regierungsexperten der Genfer Waffenkonvention, haben sich 26 Staaten, darunter China und Österreich, für einen Sperrvertrag von vollautonomen KI-Waffen ausgesprochen (vgl. Ermert 23.04.18). Neben Regierungsexperten sprechen sich zusätzlich vermehrt Bürger und Wissenschaftler gegen vollautonome Waffen aus. Beispiele hierfür sind die Gründung der Vereinigung StopKillerRobots sowie eine Unterschriftensammlung seitens Google-Mitarbeitern, die sich gegen die Zusammenarbeit mit dem Pentagon wehren (Ermert 18.02.18; Ex-Google-Chef lobt Pläne des Militärs 18.04.18). Wie bereits erwähnt setzt sich auch Microsoft mit Richtlinien für den Einsatz und die Entwicklung von KI auseinander. Doch auch andere große Tech-Unternehmen wie Google, Facebook, Apple und Amazon erlegen sich selbst Richtlinien auf oder bilden Forschungsorganisationen wie OpenAI und AI-Partnerschaften, die zur sicheren Entwicklung von KI dienen sollen (vgl. OpenAI; vgl. Tilley 27.01.17). OpenAI will z.B. sicherstellen,

dass jeder die Vorteile einer ausgereiften KI erhält, die allgemeine Sicherheit gewährleisten und Kooperationen mit anderen Forschungsinstituten bilden (vgl. OpenAI). Beachtenswert ist hierbei, dass sich die Tech-Unternehmen zum ersten Mal selber umgreifende einschränkende Richtlinien vorgeben, was nur die Bedeutsamkeit der Technologie sowie die Bedrohlichkeit, die von ihr ausgeht, unterstreicht. Neben Gesetzen und Leitlinien die einen ethischen Rahmen für die Unternehmen bilden sollen, d.h. was für ein Algorithmus darf man entwickeln und welchen nicht, müssen auch seitens der Technologie Fortschritte gemacht werden, um Probleme zu vermeiden. Eine Kombination aus Richtlinien und Technologie ist das sogenannte "Values by Design" (vgl. Meyer 28.11.16). Dadurch könnte man direkt die Grundrechte in den Code einpflegen, wodurch im idealen Fall der Algorithmus moralisch richtig entscheidet. Eine der größeren Herausforderungen stellt die Blackbox-Problematik der KI dar, denn bislang kann man aufgrund der Architektur der KI nicht sehen, wie der Algorithmus auf ein Endergebnis gekommen ist, sondern nur das Endergebnis selbst sowie welche Daten es dafür benutzt hat (vgl. Beuth 25.10.17). Allerdings gibt es auch hier wiederum Verfahren, die z.B. durch leichte Input-Veränderungen die Funktionsweise dieser Blackbox deutlich machen sollen (vgl. Beuth 25.10.17). Eine KI, die sich stetig selbst verbessert und im schlimmsten Falle außer Kontrolle gerät, wird wohl als größte endgültige Gefahr angesehen. Doch auch in diesem Feld gibt es bereits Forschungen, die sich als wirksam herausgestellt haben (vgl. Wiltz 23.01.18). Die Algorithmen einer KI lernen von Fehlern, d.h. wenn man einen Algorithmus ausschaltet, wird er davon lernen und versuchen das Ausschalten zu umgehen, da es ihn davon abhält seine Aufgabe zu erfüllen. Durch das System der "safe interruptibility" wird dem Algorithmus beigebracht, nicht von dem Ausschalten zu lernen (vgl. Wiltz 23.01.18). Allerdings gibt es hier weitere Hürden an denen geforscht wird um ein komplettes Abschalten möglich zu machen.

Letztendlich lässt sich sagen, dass sich viele Bereiche des Lebens durch die rasante Entwicklung von KI verändern werden. Um die Probleme und Gefahren dieses Prozesses zu minimieren, bedarf es rechtzeitig an strikten Richtlinien, an die sich Staaten und Unternehmen halten, sowie technischer Forschung. Allgemein benötigt es ein Verständnis innerhalb der Regierungen, warum Regulierungen und Richtlinien sinnvoll sind und welche Folgen ein unbegleiteter Prozess haben kann. Bislang gibt es noch zu wenige Vorschläge für umgreifende Richtlinien seitens der Staaten. Die beste Lösung wären wohl globale Richtlinien, die mithilfe aller Staaten sowie den größten Tech-Firmen, erstellt werden. Es muss etwa Pläne für Mitarbeiter geben, die durch die KI ersetzt wurden und eine Kontrolle der Algorithmen geben, bevor diese Aufgaben in heiklen Bereichen wie in der Medizin oder im Recht eingesetzt werden. Bei Einhaltung aller Vorsichtsmaßnahmen kann die sichere Entwicklung und Einführung der KI in näherer Zukunft gewährleistet werden. Allerdings ist es nur schwer zu sagen, was bei einem Eintritt einer Singularität geschehen wird.

# Literaturverzeichnis


Ankenbrand, Hendrik (09.04.18): China hat das wertvollste KI-Startup der Welt. Online verfügbar unter http://www.faz.net/aktuell/wirtschaft/unternehmen/china-ueberwachung-durch-gesichtserkennung-15533068.html, zuletzt aktualisiert am 09.04.18, zuletzt geprüft am 22.09.18.

Bass, Dina (18.01.18): Microsoft Says AI Advances Will Require New Laws, Regulations. Online verfügbar unter https://www.bloomberg.com/news/articles/2018-01-18/microsoft-says-ai-advances-will-require-new-laws-regulations, zuletzt aktualisiert am 18.01.18, zuletzt geprüft am 23.09.18.

Beuth, Patrick (25.10.17): Die Automaten brauchen Aufsicht. Online verfügbar unter https://www.zeit.de/digital/internet/2017-10/kuenstliche-intelligenz-deepmind-back-box-regulierung, zuletzt aktualisiert am 25.10.17, zuletzt geprüft am 28.09.18.

Boyle, Alan (10.04.18): Policy experts trade ideas for intelligent ways to regulate artificial intelligence. Online verfügbar unter https://www.geekwire.com/2018/policy-experts-debate-regulate-artificial-intelligence-intelligently/, zuletzt aktualisiert am 10.04.18, zuletzt geprüft am 23.09.18.

Breuer, Ingeborg (01.03.18): Wie Computer die Meinungsbildung beeinflussen. Online verfügbar unter https://www.deutschlandfunk.de/demokratie-und-kuenstliche-intelligenz-wie-computer-die.1148.de.html?dram:article_id=411990, zuletzt aktualisiert am 01.03.18, zuletzt geprüft am 22.09.18.

Brundage, Miles; Avin, Shahar; Clark, Jack; Toner, Helen; Eckersley, Peter (2018): The Malicious Use of Artificial Intelligence: Forecasting, Prevention,and Mitigation. Online verfügbar unter https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v_50335.pdf, zuletzt geprüft am 21.09.18.

Buxmann, Peter; Schmidt, Holger (2018): Künstliche Intelligenz. Mit Algorithmen Zum Wirtschaftlichen Erfolg. Berlin, Heidelberg: Gabler.

Conrad, Conrad Sebastian (2017): Künstliche Intelligenz – Die Risiken für den Datenschutz. Online verfügbar unter https://www.datenschutz-notizen.de/wp-content/uploads/2017/12/Kuenstliche-Intelligenz-Die-Risiken-fuer-den-Datenschutz_DuD.pdf, zuletzt geprüft am 22.09.18.

DeepMind: AlphaGo. Online verfügbar unter https://deepmind.com/research/alphago/, zuletzt geprüft am 21.09.18.

Digital Charta: Digital Charta. Online verfügbar unter https://digitalcharta.eu/, zuletzt geprüft am 23.09.18.

Dörner, Stephan (2017): Superintelligenz: Diese kommende Erfindung könnte das Ende der Menschheit bedeuten. Online verfügbar unter https://t3n.de/news/superintelligenz-ki-ai-787316/, zuletzt aktualisiert am 20.01.2017, zuletzt geprüft am 21.09.18.

Erdélyi, Olivia J.; Goldsmith, Judy (2018): Regulating Artificial Intelligence: Proposal for a Global Solution. Association for the Advancement of Artificial. Online verfügbar unter http://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_13.pdf, zuletzt geprüft am 23.09.18.

Ermert, Monika (23.04.18): Autonome Waffen: Bundesregierung spricht sich bedingt gegen Wettrüsten mit smarten Waffen aus. Online verfügbar unter https://www.heise.de/newsticker/meldung/Autonome-Waffen-Bundesregierung-spricht-sich-bedingt-gegen-Wettruesten-mit-smarten-Waffen-aus-4029629.html, zuletzt aktualisiert am 23.04.18, zuletzt geprüft am 23.09.18.

Ermert, Monika (18.02.18): Missing Link: Eindringliche Warnung vor autonomen Waffen. Online verfügbar unter https://www.heise.de/newsticker/meldung/Missing-Link-Eindringliche-Warnung-vor-autonomen-Waffen-3971911.html?seite=all, zuletzt aktualisiert am 18.02.18, zuletzt geprüft am 23.09.18.

Ex-Google-Chef lobt Pläne des Militärs (18.04.18). Online verfügbar unter http://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/ex-google-chef-schmidt-lobt-das-militaer-15548612.html, zuletzt aktualisiert am 18.04.18, zuletzt geprüft am 24.09.18.

Grace, Katja; Salvatier, John; Dafoe, Allan; Zhang, Baobao; Evans, Owain (2018): When Will AI Exceed Human Performance? Evidence from AI Experts. Oxford University, AI Impacts, Yale University. Online verfügbar unter https://arxiv.org/pdf/1705.08807.pdf, zuletzt geprüft am 21.09.18.

Huet, Cécile: Smart Policies. Online verfügbar unter https://www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/conference-agenda/ai-intelligent-machines-smart-policies-huet.pdf.

Mainzer, Klaus (2016): Künstliche Intelligenz - Wann übernehmen die Maschinen? Berlin, Heidelberg: Springer (Technik im Fokus).

Martin-Jung, H. (2010): Googles Zugeständnisse. Online verfügbar unter https://www.sueddeutsche.de/digital/datenspeicherung-googles-zugestaendnisse-1.689683, zuletzt aktualisiert am 17.05.2010, zuletzt geprüft am 23.09.18.

Mewes, Bernd (10.03.18): KI-gesteuertes Marketing: Zalando streicht 250 Arbeitsplätze. Online verfügbar unter https://www.heise.de/newsticker/meldung/KI-gesteuertes-Marketing-Zalando-streicht-250-Arbeitsplaetze-3990425.html, zuletzt aktualisiert am 10.03.18, zuletzt geprüft am 23.09.18.

Meyer, Jan-Bernd (28.11.16): Gefährdet KI die Demokratie? - Interview mit Yvonne Hofstetter. Online verfügbar unter https://www.computerwoche.de/a/wollt-ihr-die-totale-ueberwachung,3327610, zuletzt aktualisiert am 28.11.16, zuletzt geprüft am 22.09.18.

OpenAI: OpenAI. Online verfügbar unter https://openai.com/, zuletzt geprüft am 24.09.18.

Prössl, Christoph (27.08.17): Streit um autonome Kriegsmaschinen. Online verfügbar unter https://www.tagesschau.de/inland/autonome-waffen-101.html, zuletzt aktualisiert am 27.08.17, zuletzt geprüft am 22.09.18.

Sauer, Candy; Braun, Anja (13.09.18): Computer erkennt Hautkrebs besser als Ärzte. Online verfügbar unter https://www.swr.de/wissen/kuenstliche-intelligenz-erkennt-hautkrebs-zuverlaessiger-als-fachaerzte-neue-studie/-/id=253126/did=21779082/nid=253126/ngqxjm/index.html, zuletzt aktualisiert am 13.09.18, zuletzt geprüft am 21.09.18.

Schaub, Gary; Wenzel Krisoffersen, Jens (2017): In, On, or Out of the Loop? Denmark and Autonomous Weapon Systems. University of Copenhagen. Online verfügbar unter https://cms.polsci.ku.dk/publikationer/in-on-or-out-of-the-loop/In_On_or_Out_of_the_Loop.pdf, zuletzt geprüft am 22.09.18.

Stöcker, Christian (20.05.18): Moore's Law ist eine lahme Ente. Online verfügbar unter http://www.spiegel.de/wissenschaft/technik/kuenstliche-intelligenz-moore-s-law-ist-eine-lahme-ente-kolumne-a-1208312.html, zuletzt aktualisiert am 20.05.18, zuletzt geprüft am 21.09.18.

Tilley, Aaron (27.01.17): Why Apple Joined Rivals Amazon, Google, Microsoft In AI Partnership. Online verfügbar unter https://www.forbes.com/sites/aarontilley/2017/01/27/why-apple-joined-

rivals-amazon-google-microsoft-in-ai-partnership/#494deaaf5832, zuletzt aktualisiert am 27.01.17, zuletzt geprüft am 24.09.18.

Wiltz, Chris (23.01.18): How Can Engineers Stop AI from Going Rogue? UBM America. Online verfügbar unter https://www.designnews.com/electronics-test/how-can-engineers-stop-ai-going-rogue/54158381258137, zuletzt aktualisiert am 23.01.18, zuletzt geprüft am 26.09.18.