UNIVERSITÄT OSNABRÜCK

LV Geopolitik und Geostrategie
Fachbereich 1
49069 Osnabrück

Militärische und Sicherheits-Aspekte der Künstlichen Intelligenz

22.06.2020

Zusammenfassung
In diesem Arbeitspapier werden militärische und sicherheitstechnische Aspekte der künstlichen Intelligenz (KI)/Artificial Intelligence (AI) als neuer Bereich der Sicherheitspolitik vorgestellt. KI wird allgemein als die Fähigkeit von Maschinen verstanden, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern, und ist ein Schlüsselbereich fortgeschrittener Computertechnologie. Wichtige KI-bezogene Techniken umfassen neuronale Netze, Deep Learning, maschinelles Lernen, Edge Computing und Robotik. Die Konzepte und Definitionen der KI und ihre Auswirkungen auf Konstruktionsprozesse werden vorgestellt. Die USA und China konkurrieren um die Technologieführerschaft in der KI, gefolgt von Europa. Die KI-Strategien dieser Akteure werden vorgestellt, wobei in diesem Paper der Schwerpunkt auf militärischen Projekten liegt, zu denen eine Vielzahl unbemannter und autonomer Fahrzeuge, aber auch Programme für C2 (Command and Control) und Überwachung im Rahmen der Intelligence, Surveillance and Reconnaissance (ISR) gehören.
KI-Systeme haben ein spezifisches Cybersicherheitsprofil; sie können zur Erkennung von Cyberangriffen und zur automatisierten Cyberabwehr dienen, weisen jedoch auch komplexe Schwachstellen auf, die mit neuen Angriffstypen wie gezielten Manipulationen von Input-Daten und Bildern ausgenutzt werden können. KI-Systeme gewinnen auch für den Informationskrieg zunehmend an Bedeutung. Abschließend wird kurz das komplexe Feld der Maschinenlogik und -Ethik vorgestellt.

Contents

1. Grundlagen 3
1.1 Einführung 3
1.2 Was ist Künstliche Intelligenz? 3
1.2.1 Die Arbeitsdefinition des US-Verteidigungsministeriums DoD 3
1.2.2 ‘Starke’ und ‘Schwache’ KI 4
1.2.3 KI-bezogene Techniken 5
1.2.4 Der Einfluss auf Konstruktionsprozesse 6
1.2.4.1 Computer und Maschinen 6
1.2.4.2 Computer und Biologische Systeme 7
2. KI-Strategien 8
2.1 Einführung 8
2.2 Die KI-Strategie der Vereinigten Staaten 8
2.3 Die KI-Strategie Chinas 11
2.4 Die Verflechtung der USA und Chinas 12
2.5 Die Balance zwischen Cyber- und physischen Fähigkeiten 13
2.6 Die KI-Strategie der Europäischen Union 14
3. Militärische Aspekte 15
3.1 Eine einführende Fallstudie: Das Eurosur-Projekt 15
3.2 Praktische Anwendungen 16
3.2.1 Unmanned Aerial Vehicles (UAVs, Drohnen) 16
3.2.2 Autonome Fahrzeuge 19
3.2.3 Intelligence, Surveillance, and Reconnaissance (ISR) 19
3.2.4 Command and Control-Systeme 19
3.2.5 Logistik 19
4. Sicherheitsaspekte 21
4.1 Kurze Einführung 21
4.2 Cyber-Attacken 21
4.3 Wichtige Schwachstellen von KI-Systemen 22
4.3.1 Grundlegende Probleme der KI 22
4.3.2 Missionsstabilität 23
4.3.3 Daten-Manipulation 23
4.3.4 Hardware in der KI 24
4.4 Cyberverteidigung 25
4.4.1 Detektion von Cyberangriffen 25
4.4.2 Automatisierte Cyberabwehr 26
4.5 Informationskrieg 26
5. Ethik und Maschinen-Logik 28
6. Abschließende Bemerkungen 29
7. Literatur 30
7.1 Literaturquellen 30
7.2 Literaturempfehlungen 34

1. Grundlagen

1.1 Einführung

Künstliche Intelligenz (KI), englisch: **Artificial Intelligence (AI)** wird allgemein als die Fähigkeit von Maschinen verstanden, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern, und ist ein Schlüsselbereich fortgeschrittener Computertechnologie. Wichtige KI-bezogene Techniken umfassen neuronale Netze, Deep Learning, maschinelles Lernen, Edge Computing und Robotik. In diesem Papier werden militärische und sicherheitstechnische Aspekte der KI als neuer Bereich der Sicherheitspolitik vorgestellt.

1.2 Was ist Künstliche Intelligenz?

1.2.1 Die Arbeitsdefinition des US-Verteidigungsministeriums DoD

Selbst für die menschliche Intelligenz gibt es keine Standarddefinition. Der Kern der Definitionen der menschlichen Intelligenz umfasst jedoch die mentale Fähigkeit, Probleme zu erkennen, zu analysieren und zu lösen. Ein Mensch ist dann intelligenter, wenn dies schneller und/oder bei komplexeren Problemen möglich ist.
Historisch gesehen war Konzept der künstlichen Intelligenz (KI) auf Maschinen ausgerichtet, die menschliche Intelligenz simulieren. Eine praktische Definition, die das allgemeine Verständnis von KI abdeckt, wurde vom US-Verteidigungsministerium (*Department of Defense DoD*) vorgenommen.
In der Zusammenfassung der DoD-KI-Strategie für 2018 heißt es: "*AI refers to the ability of machines to perform tasks that normally require human intelligence—for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action— whether digitally or as the smart software behind autonomous physical systems*."[1] Übersetzung: „*KI bezieht sich auf die Fähigkeit von Maschinen, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern - beispielsweise Muster erkennen, aus Erfahrungen lernen, Schlussfolgerungen ziehen, Vorhersagen treffen oder Maßnahmen ergreifen - ob digital oder als intelligente Software hinter autonomen physischen Systemen.* "
Viele Definitionen konzentrieren sich auf Aktivitäten, die menschliche Intelligenz erfordern, aber genau genommen haben bereits die einfachen Taschenrechner der 1970er Jahre etwas geleistet, das normalerweise menschliche Intelligenz erfordert. Aus der Literatur geht jedoch hervor, dass die KI-Forscher fortgeschrittenes und autonomes Rechnen meinen, wenn sie über KI sprechen.
Intelligente Agenten (**intelligent agents**) sind daher alle Geräte, die die Umgebung wahrnehmen und die Chance auf Zielerreichung maximieren können. Wenn eine Computeranwendung zur Normalität wird, wird sie typischerweise nicht mehr als KI betrachtet (**KI-Effekt**). Frühere Beispiele sind z.B. Taschenrechner, Übersetzungscomputer und Schachcomputer, aktuelle Beispiele sind Navigationssysteme und Heimassistenzsysteme wie *Alexa*, *Siri* usw.

[1] vgl. DOD 2018, S.5

Der *National Defense Authorization Act (NDAA)* für das Fiskaljahr 2019 enthält eine formale Definition der KI mit fünf Arten von KI-Systemen:[2]

1. Jedes künstliche System, das Aufgaben unter verschiedenen und unvorhersehbaren Umständen ohne nennenswerte menschliche Aufsicht ausführt oder aus Erfahrungen lernen und die Leistung verbessern kann, wenn es Datensätzen ausgesetzt ist.
2. Ein künstliches System, das in Computersoftware, physischer Hardware oder einem anderen Kontext entwickelt wurde und Aufgaben löst, die eine menschenähnliche Wahrnehmung, Erkenntnis, Planung, Lernen, Kommunikation oder physisches Handeln erfordern
3. Ein künstliches System, das so konzipiert ist, dass es wie ein Mensch denkt oder handelt, einschließlich kognitiver Architekturen und neuronaler Netze.
4. Eine Reihe von Techniken, einschließlich maschinellem Lernen, mit denen eine kognitive Aufgabe angenähert werden soll.
5. Ein künstliches System, das für rationales Handeln ausgelegt ist, einschließlich eines intelligenten Software-Agenten oder eines physischen Roboters, der Ziele durch Wahrnehmung, Planung, Argumentation, Lernen, Kommunikation, Entscheidungsfindung und Handeln erreicht.

1.2.2 'Starke' und 'Schwache' KI

Die sogenannte "schwache" KI kann ein beobachtetes Verhalten reproduzieren und Aufgaben nach einem Training ausführen[3], d.h. Systeme, die maschinelles Lernen, Mustererkennung, Data Mining oder die Verarbeitung natürlicher Sprache anwenden. Intelligente Systeme, die auf "schwacher" KI basieren, umfassen z.B. Spamfilter, selbstfahrende Autos und Industrieroboter. Im Gegensatz dazu wäre „starke" KI ein intelligentes System mit echtem Bewusstsein und Denkfähigkeit.
Die aktuelle KI von 2020 ist immer noch eine „schwache" KI mit programmierten Maschinen, die schnelle Berechnungen durchführen, die es ihnen ermöglichen, Aktionen mithilfe von Datenbanken und statistischen Modellen zu interpretieren, nachzuahmen oder vorherzusagen, aber immer noch keine Vorstellung von sich selbst haben und nicht reflektieren können, d.h. sie kann nicht wirklich "Ich" und "Warum" denken oder meinen.

Auf der anderen Seite umfassen menschliche Handlungen viele sich wiederholende und routinemäßige Aktivitäten, die standardisiert werden können und daher bereits jetzt für die KI zugänglich sind. Darüber hinaus ist die Entscheidungsfindung oft nur die Wahl zwischen Standardoptionen. Sogar Dinge, die Menschen als komplexe Aktivität wahrnehmen, z.B. Autofahren von Stadt A nach Stadt B, besteht meist aus langen Abfolgen von Routinetätigkeiten und Standardentscheidungen, zum Beispiel: Das Auto kommt an eine Ampel: anhalten oder fahren?,… .dann fahren…. Eine Kreuzung kommt: links oder rechts abbiegen? ... dann wieder fahren ... und so weiter ...

---

[2] vgl. NDAA 2019, Section 238. Originaltext: 1. Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
2. An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action
3. An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
4. A set of techniques, including machine learning that is designed to approximate a cognitive task.
5. An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.

[3] vgl. Perez at al 2019, S.6

Dies gilt in ähnlicher Weise auch für die Industrieproduktion und Maschinenaktivitäten.
Zusammenfassend lässt sich sagen, dass bereits aktuelle KI-Systeme in der Lage sind, menschliche Aktivitäten in wesentlichen Bereichen des täglichen Lebens, der Kommunikation, des Handels, der Industrie usw. zu unterstützen oder zu ersetzen und alle Arten der Maschinennutzung zu unterstützen oder zu steuern, was das massive Wachstum der KI und ihr enormes Potenzial erklärt.

1.2.3 KI-bezogene Techniken

Wichtige KI-bezogene Techniken sind **neuronale Netze, Deep Learning, maschinelles Lernen, Edge Computing** und **Robotik**.

Neuronale Netze: Das menschliche Gehirn verarbeitet Eingaben mit miteinander verbundenen Knoten von Nervenzellen, den Neuronen. Die Verarbeitung umfasst die Signalübertragung, aber auch das Filtern durch inhibitorische Neuronen. Schließlich können eingehende Eingabemuster mit bekannten Mustern verglichen werden, um eine Reaktion zu erzeugen. Als vereinfachtes Beispiel: Wenn die Augen auf der Straße ein Objekt mit vier Rädern sehen, werden Signale von der Netzhaut der Augen zum optischen Kortex im hinteren Gehirn und von dort zum benachbarten interpretativen Kortex und zu den Gedächtnisbereichen im Hippocampus übertragen, was schließlich ermöglicht, das Objekt als "Auto" zu klassifizieren, auch wenn das spezifische Automodell noch nie zuvor gesehen wurde.
Das gleiche Prinzip wird in KI-Anwendungen verwendet: Die Eingabe (Input) wird durch mehrere verborgene Schichten (hidden layers) von Computerbereichen (Knoten) übertragen und gefiltert, bevor das Ergebnis (z.B. Objektklassifizierung, Entscheidung) als Output ausgegeben wird.

Neuronale Netze können azyklische oder vorwärts gerichtete neuronale Netze (**feedforward neural networks)** sein, bei denen das Signal nur in eine Richtung verläuft, und wiederkehrende neuronale Netze (**recurrent neural networks)** mit Rückkopplungssignalen und Kurzzeitgedächtnissen früherer Inputs.
Deep Learning bedeutet das Erlernen einer langen Kette von Kausalitäten auf der Grundlage neuronaler Netze, während sich das verwandte Konzept des **maschinellen Lernens (ML)** auf das Gedächtnis (Erfahrung) konzentriert, indem Computeralgorithmen entwickelt werden, die sich durch Erfahrung automatisch verbessern. Die **Fuzzy logic** konzentriert sich auf die Manipulation von Informationen, die oft unscharf (fuzzy) sind, z.B. "Setzen Sie es etwas höher", wo der Algorithmus hilft, es in eine genauere Information umzuwandeln.
Die Verarbeitung natürlicher Sprache im **Natural language processing** umfasst Algorithmen zum Verständnis der menschlichen Sprache durch systematische Analyse der Sprachelemente und ihrer Beziehungen. Ein verwandter Bereich ist die Sprachverarbeitung, das **voice processing**.

Ein neuer KI-Bereich sind bioinspirierte Berechnungsmethoden (**Bio-Inspired Computation Methods**), die Sammlungen intelligenter Algorithmen und Methoden verwenden, die bioinspirierte Verhaltensweisen und Eigenschaften wie genetische Algorithmen (GA = Mutation, Rekombination und Auswahl von Algorithmen),

Evolutionsstrategien (ES) und Ameisenkolonien-Optimierungsverfahren (ACO), Partikelschwarmoptimierung (PSO) und künstliche Immunsysteme (artificial immune systems AIS)[4].

Edge Computing ist eine Schicht verteilter Computer zwischen Clouds und Benutzern, die Berechnung und Datenspeicherung näher an den Ort bringt, an dem sie benötigt werden, um die Antwortzeiten zu verbessern.

Die Kernidee der Verbindung von **KI und Robotik** versucht, die Autonomie der Roboter durch Lernen zu optimieren, um die Fähigkeit zur Manipulation, Navigation und Zusammenarbeit zu verbessern. Roboter können die Umgebung durch integrierte Sensoren oder Computersehen, was ein weiteres Feld der KI ist[5]. In der Praxis kann ein Anstieg von Co-Bots (co-worker robots) beobachtet werden, die Menschen unterstützen, z.B. durch Übernahme sich wiederholender Tätigkeiten wie Sortieren oder Tragen von Gegenständen, Raumdesinfektion usw.[6].

Ursprünglich waren KI, maschinelles Lernen, Mustererkennung, Robotik usw. relativ unabhängige Forschungsbereiche, aber mittlerweile fließen sie zunehmend ineinander, sodass ein breiteres Verständnis der KI diese Bereiche in die Diskussion einbezieht. Das moderne Konzept automatisierter Systeme umfasst somit die ursprünglich getrennten, sich jetzt jedoch überlappenden Konzepte von Autonomie, Robotik und KI[7].

1.2.4 Der Einfluss auf Konstruktionsprozesse

1.2.4.1 Computer und Maschinen

Derzeit besteht der typische Konstruktionsprozess größerer Maschinen darin, verschiedene Computerelemente einzubetten und zur Steuerung der Maschine miteinander zu verbinden. Ein *Eurofighter*-Kampfjet hat mehr als 80 Computer und 100 Kilometer Verkabelung[8]. Diese Konstruktion führt jedoch zu einer sehr komplexen Computerumgebung mit vielen Schnittstellen, was das Risiko für Kommunikations- und Kompatibilitätsprobleme sowie Softwareprobleme erhöht, es zudem schwierig macht, alle Systeme auf dem neuesten Stand zu halten, und viele Schwachstellen für Cyberangriffe bietet.

Ein NATO-Staat hat einen Kampfjet zerlegt, um sämtliche Komponenten gegen Cyberattacken zu härten und baute den Jet anschließend wieder zusammen, aber die Kosten der Maßnahme führten zu der Überlegung, dass die Komponentensicherheit stattdessen von den Lieferanten garantiert werden sollte[9]. Das würde jedoch bedeuten, sich auf die Sicherheitsanstrengungen zahlreicher Anbieter verlassen zu müssen, d.h. es ist schwierig, die Cybersicherheit zu delegieren. Ähnliche Prüfungen bei Autohacks zeigten, dass die Vorstellung des **walled garden**-Konzepts, dass man die vielen Komponenten von außen

[4] vgl. Truong/Diep/Celinka 2020, S.24
[5] vgl. Perez et al. 2019, S.24
[6] vgl. Jung 2020, S.70-71
[7] vgl. Hoadley/Sayler 2019, S.4
[8] vgl. Köpke/Demmer 2016, S.2
[9] vgl. Leithäuser 2016, S.8

ganzheitlich schützen könnte, Eindringtesten nicht standhielt, d.h. jede Komponente muss einzeln gesichert werden[10].

Der Trend geht nun dahin, zuerst ein vollständig integriertes Computersystem mit eingebetteten KI-Elementen zu schaffen und die Maschinenumgebung darauf auszurichten und anzupassen, wie z.B. in den neuesten *Tesla*-Automodellen[11].
Dies ermöglicht eine signifikante Vereinfachung der IT-Umgebung in Kombination mit größeren Datenflüssen und kann eine Option für andere Maschinen aber auch z.B. militärische Maschinen und Flugzeuge, die inzwischen mit komplexen Computer-Elementen (über)beladen sind.

1.2.4.2 Computer und Biologische Systeme

Einbettung von Computern ist auch für biologische Organismen relevant. Eine engere Definition spricht von **Cyborgs** (kybernetischen Organismen), wenn biologische und computersteuerbare maschinelle Bestandteile physisch integriert sind. Retina- und Cochleaimplantate erfüllen auch die strikte Definition. Es ist wesentlich, dass die Entwicklung von Cyborgs viel langsamer verläuft als erwartet, da dieser Ansatz ein sehr begrenztes Potenzial hat. Unter anderem sind die Schnittstellen zwischen lebenden und Computerabschnitten eine Herausforderung. Ein weiteres Problem ist die Energieversorgung der Maschinenteile, da Hitze oder Strahlung das umgebende Gewebe beschädigen können. Das Immunsystem und das umgebende Gewebe neigen dazu, mit Entzündungen, Abstoßungen und Fibrosen gegen die Implantate zu reagieren. Wartungs- und Reparaturanforderungen werden bereits als Hintertüren für Cyberangriffe verwendet. Zusammenfassend scheint die Menge an Maschinenteilen, die ein Organismus tragen kann, ziemlich begrenzt zu sein.
Im Vergleich dazu scheinen **autonome Biohybride**, das sind freie Kombinationen von biologischen und synthetischen Materialien ein viel größeres Potenzial zu haben. Hier wird maßgeschneidertes biologisches Material mit computersteuerbaren maschinellen Bestandteilen kombiniert, und die künstliche Intelligenz könnte die Autonomie dieses Systems gewährleisten.
Im Jahr 2016 wurde ein Schwimmroboter gebaut, der einen Rochen nachahmt und der aus einem feinen Goldskelett und einem Gewebe aus 200.000 genetisch veränderten Rattenherzmuskelzellen bestand[12]. Die Zellen wurden genetisch verändert, so dass die Geschwindigkeit und die Richtung durch Veränderung von Licht gesteuert werden konnte. Der Biohybrid blieb jedoch von der Anwesenheit einer physiologischen Kochsalzlösung umgebungsabhängig.

Derzeit werden drei Schlüsseltechnologien entwickelt, die möglicherweise fortgeschrittene Biohybride ermöglichen: **künstliche Zellen, Organoide** und **synthetische/künstliche Genome**.

[10] vgl. Mahaffey 2016, S.V6
[11] vgl. Floemer 2020
[12] vgl. Park et al. 2016

Seit 2010 wird an der Entwicklung einer Zelle mit **minimalem Genom** gearbeitet, d.h. dem kleinstmöglichen Genom, das autonomes Leben und Replikation ermöglicht[13]. 2016 wurde eine neue Zelle namens *Syn 3*.0 erschaffen, indem das Genom von *Mycoplasma capricolum* durch das Genom von *Mycoplasma mycoides* ersetzt und nicht benötigte DNA entfernt wurde[14]. Nachdem festgestellt wurde, dass ein etwas größeres Genom als das Kleinstmögliche zu einem verbesserten Zellwachstum führt, wurde eine modifizierte Minimalzelle erzeugt, die es im Jahr 2019 ermöglichte, die Anzahl der Gene mit unbekannter Funktion auf 30 zu reduzieren[15]. Wenn die Funktion dieser 30 Gene geklärt werden könnte, würden die grundlegenden Mechanismen lebender Zellen identifiziert und könnten dann verwendet werden, um **frei designbare künstliche Zellen** zu erzeugen.
Auch die Kontrolle der Zelldifferenzierung hat erhebliche Fortschritte gemacht: Organoide sind kleine **künstliche Organe**, die durch gezielte Anwendung von Wachstumsfaktoren und Hormonen auf Stammzellen mit vielen Funktionen des ursprünglichen Organs versehen sind, z.B. Lungen und Atemwege[16] für Studien zu Coronavirus-Infektionen, aber auch andere Organoide wie kleine Gehirne.
Das andere Thema sind **synthetische Genome**[17]. Der schnelle technische Fortschritt der DNA-Synthese ermöglicht inzwischen die Synthese künstlicher Hefechromosomen (*S. cerevisiae*).
Zusammen mit designbaren Zellen kann diese Technologie eine groß angelegte genomische Variation und Optimierung ermöglichen.

2. KI-Strategien

2.1 Einführung

Die USA und China konkurrieren um die Technologieführerschaft in der KI, gefolgt von Europa als drittgrößtem Akteur.
Wie bei anderen fortschrittlichen Technologien wird die Forschung von drei Gruppen durchgeführt, d.h. staatlichen Einrichtungen, privaten Unternehmen und akademischer Forschung. In komplexen Projekten kooperieren diese Gruppen miteinander und der Staat versucht, die KI-Projekte von höchstem strategischem Wert zu koordinieren und zu finanzieren. In den Sicherheitssektoren meint dies die Anwendungen mit dem größten Einfluss auf die militärischen und nachrichtendienstlichen Fähigkeiten.
Die zentrale strategische Herausforderung besteht darin, diese strategischen KI-Anwendungen zu identifizieren und die Koordination für die schnelle Entwicklung und Bereitstellung sicherzustellen.

2.2 Die KI-Strategie der Vereinigten Staaten

Die Exekutiv-Verordnung *Presidential Executive Order on Maintaining American Leadership in AI*[18] vom 11.02.2019 betonte die Bedeutung einer fortgesetzten

[13] vgl. Kastilan 2010
[14] vgl. Danchin/Fang 2016
[15] vgl. Lachance et al. 2019
[16] vgl. Elbadawi/Efferth 2020, Heide/Huttner/Mora-Bermudez 2018
[17] vgl. Wang/Zhang 2019, S.23
[18] vgl. Trump 2019

amerikanischen Führung in der KI für ihre wirtschaftliche und nationale Sicherheit und für die Gestaltung der globalen Entwicklung der KI in einer Weise, die den Werten, Prinzipien und Prioritäten der USA entspricht. Gleichzeitig veröffentlichte das US-Verteidigungsministerium eine nicht klassifizierte Zusammenfassung seiner KI-Strategie mit einem klaren Fokus auf das *Joint Artificial Intelligence Center (JAIC)* für die Strategieumsetzung[19].

Die primäre strategische Ausrichtung für die Zukunft liegt in der Zusammenarbeit mit den Nachrichtendiensten der *Five Eyes*-Gruppe (US, UK, CDN, AUS, NZ) und dann sekundär auch mit der NATO[20].
Das *White House Office of Science and Technology Policy's National Science and Technology Council* publizierte im Juni 2019 den National *AI R&D Strategic Plan,* der die zentralen Kriterien für Forschungs- und Entwicklungsausgaben der Regierung im Bereich der KI definierte [21].
Die Vereinigten Staaten haben den institutionellen Rahmen für KI-Forschung und -Finanzierung systematisch erweitert[22].

[19] vgl. DoD 2018, S.9
[20] vgl. NSCAI 2020, S.4
[21] vgl. OSTP 2020, S.6
[22] vgl. Hoadley/Sayler 2019, S.9-10, RAND 2019, DoD 2018, OSTP 2020, NSCAI 2020

Sektor/Administration	Institution	KI-Relevanz
Militär		
Department of Defense DoD = Verteidigungs-Ministerium	Joint Artificial Intelligence Center (JAIC) seit 2019	koordiniert die Bemühungen, Technologien für künstliche Intelligenz zu entwickeln, auszureifen und in den praktischen Einsatz zu überführen
	National Security Commission on Artificial Intelligence (NSCAI) seit 2019	Bewertung militärisch relevanter KI-Technologien und Empfehlungen
	Defense Advanced Research Projects Agency (DARPA) seit 60 Jahren	Zurzeit über 20 KI-Programme
	Defense Innovation Unit DIU seit 2016	Die DIU arbeitet mit Unternehmen zusammen, um kommerzielle Lösungen für DoD-Probleme zu entwickeln. Aufträge werden in der Regel in weniger als 90 Tagen vergeben
Nachrichtendienste		
Office of the Director of National Intelligence ODNI Büro des Nationalen Geheimdienst-koordinators	Intelligence Advanced Research Projects Agency (IARPA) seit 2007, integrierte Vorläufer-agenturen aus der NSA, NGA und CIA	Ähnliches Ziel wie DARPA, jedoch mit Schwerpunkt auf Nachrichtendienste. Initiierte das funktionsübergreifende Team für algorithmische Kriegsführung *Algorithmic Warfare Cross-Functional Team (Project Maven)*, das an JAIC übergeben wird. *Project Maven:* seit 2017 zur Automatisierung der Intelligenzverarbeitung mit Computersehen und Algorithmen für maschinelles Lernen zur Zielidentifikation aus Drohnen-Daten. Andere KI-Programme umfassen die Entwicklung von Algorithmen für die mehrsprachige Spracherkennung und -übersetzung in verrauschten Umgebungen, die geografische Lokalisierung von Bildern ohne die zugehörigen Metadaten, das Zusammenführen von 2D-Bildern zur Erstellung von 3D-Modellen und Analysewerkzeuge, um die Funktion eines Gebäudes vom Nutzungsmuster abzuleiten
Central Intelligence Agency CIA	[hat eigene Firma In-Q-Tel für Kooperation mit Start-ups]	Rund 140 KI-Projekte, z.B. zur Bilderkennung und prädiktiven Analyse
Zivile Behörden		
Department of Energy DOE Energieministerium	Artificial Intelligence and Technology Office	die KI-Fähigkeiten von DOE zu beschleunigen und die nationale und wirtschaftliche Sicherheit zu gewährleisten
Regierung		
National Science and Technology Council NSTC	The Select Committee on AI seit 2018	Besteht aus Abteilungsleitern und Agenturen, die hauptsächlich für die KI-Forschung und -Entwicklung (Forschung und Entwicklung) der Regierung verantwortlich sind unterhalb des Information Technology R&D (NITRD) Subcommittees
	The Machine Learning and Artificial	The MLAI Subcommittee überwacht den Stand der Technik im Bereich des maschinellen Lernens (ML) und der künstlichen Intelligenz (AI) und berichtet an die

	Intelligence (MLAI) Subcommittee	NSTC Committee on Technology and the Select Committee on AI
	The AI R&D Interagency Working Group	dem NSTC NITRD Subcommittee nachgeordnet und besteht aus Forschungsprogramm-Managern und technischen Experten aus der ganzen Regierung und berichtet an die MLAI and NITRD Subcommittees

2.3 Die KI-Strategie Chinas

Gemäß dem KI-Entwicklungsplan von 2017 *New Generation AI Development Plan*, strebt China an, weltweit führend in der KI zu werden und bis 2030 einen inländischen KI-Markt im Wert von 150 Mrd. USD zu entwickeln.[23]. Die chinesische Regierung betrachtet KI als eine Gelegenheit, die Vereinigten Staaten zu „überspringen", indem sie sich auf KI konzentriert, um Entscheidungen auf dem Schlachtfeld zu beschleunigen sowie die Cyber-Fähigkeiten, Marschflugkörper und autonome Fahrzeuge in allen militärischen Bereichen zu verbessern[24].
2017 demonstrierte eine zivile chinesische Universität auf einer Flugshow einen KI-fähigen Schwarm von 1.000 unbewohnten Luftfahrzeugen. Um den Transfer von KI-Technologie von kommerziellen Unternehmen und Forschungseinrichtungen zum Militär als *zivil-militärische Integration (CMI)* zu beschleunigen, hat die chinesische Regierung 2017 eine militärisch-zivile *Military-Civil Fusion Development Commission* eingerichtet[25].

Das Konzept, wie es im Verteidigungsweißbuch (*Defense White* Paper DWP) von 2019 dargelegt wurde, führt die Entwicklung der Kriegsführung von der Mechanisierung zur Informationstechnologie und jetzt mit der KI zur „Intelligentisierung". Für die chinesische Armee PLA ist die KI daher für die „**intelligente Kriegsführung**" von wesentlicher Bedeutung.[26]. Der praktische strategische Ansatz besteht darin, Anweisungen und Ressourcen zentral bereitzustellen, diese jedoch lokal umzusetzen, damit der Wettbewerb zwischen chinesischen Städten und Regionen um KI-Forschung aktiviert wird. Um die akademischen Fähigkeiten zu stärken, wurden Hunderte neuer KI-Professuren eingerichtet. Der Forschungsschwerpunkt der militärischen KI liegt auf Command and Control-Systemen sowie auf einem breiten Spektrum unbemannter Fahrzeuge.

China investiert weiter in US-Unternehmen, die an militärisch relevanten KI-Anwendungen arbeiten, und erhält so möglicherweise einen rechtmäßigen Zugang zu Technologie und geistigem Eigentum. Die USA sind jedoch weiterhin besorgt, dass auch Industrie- und Cyberspionage betrieben werden könnte[27].
Das derzeit größte KI-Projekt ist das zivile **China Social Score System**, bei dem Gesundheitsdaten, Finanzdaten (einschließlich Konsumgewohnheiten), digitale Daten, mobile Daten und Überwachungskamerabilder kombiniert werden, um Verhaltens-,

[23] vgl. Hoadley/Sayler 2019, S.1, NATO 2019, S.10
[24] vgl. NATO 2019, S.10
[25] vgl. Hoadley/Sayler 2019, S.20-22
[26] vgl. Bommakanti 2020, S.3-4
[27] vgl. Hoadley/Sayler 2019, S.22-23

Bewegungs- und Inhaltsprofile zu erstellen. Basierend auf dem Output werden niedrigere Zinssätze, einfachere Reisen und andere Vorteile (Beförderungen, Stellenangebote, bessere Positionen auf Dating-Plattformen, wodurch die Chance auf Reproduktion verbessert wird) für Personen mit guter Punktzahl gewährt, mit entsprechenden Nachteilen für Personen mit niedriger Punktzahl. Die Idee ist das automatisierte Management einer großen Gesellschaft[28].

2.4 Die Verflechtung der USA und Chinas

Beide Staaten sind in Bezug auf personelle und technische Ressourcen miteinander verbunden. Eine im Stil des Kalten Krieges denkbare Aufteilung in zwei getrennte Cyber- und KI-Welten könnte sowohl für beide Staaten als auch für den Fortschritt der KI erhebliche Probleme verursachen[29].
Derzeit arbeiten viele chinesische Top-Forscher, die auf KI-Konferenzen herausragende Publikationen geliefert haben, in den USA anstelle von China, selbst wenn sie ihren ersten akademischen Abschluss in China gemacht haben. China versucht, KI-Forscher mit sehr guten Stellenangeboten zu gewinnen, da viele chinesische Forscher auch nach der Promotion länger in den USA bleiben, anstatt nach China zurückzukehren.
Das zentrale KI-Projekt des US-Verteidigungsministeriums *Project Maven* wurde mit Hilfe eines Dutzends von *Google* Ingenieuren entwickelt, viele von ihnen chinesische Staatsbürger. Die Projektaufsicht oblag dem Stanford-Professor Dr. Fei-Fei Li. Das Pentagon sagte, dass sie nur mit nicht klassifizierten Daten arbeiteten und dafür am besten qualifiziert waren[30].

Sowohl die USA als auch China sind wichtige Cyber-Mächte: China ist der wichtigste Produzent von physischer Elektronik in Computern und Smartphones, selbst US-Firmen lagern ihre Produktion oft nach China aus. Das ist sinnvoll, da China der Haupteigentümer von computerrelevanten Metallen ist. Digitaltechnologien, wie Handys und Computer benötigen spezielle Industriemetalle (Seltene Erden) wie Niob, Germanium, Indium, Palladium, Kobalt und Tantal. Die USA haben im Jahr 2019 35 Rohstoffe als kritisch identifiziert, aber weisen bei 14 dieser Rohstoffe keine eigene Produktion auf. Bei den seltenen Erden hat China im Jahr 2019 71% Marktanteil und 37% der Reserven.[31] Eine Knappheit hätte erhebliche Auswirkungen, da diese Metalle bisher nicht hinreichend wirtschaftlich recycelt können, also die Verluste durch Recycling nicht ausgeglichen werden können. Chinas sehr großer Anteil an seltenen Erden, die für die IT-Industrie unersetzlich sind ist daher strategisch bedeutsam.
Auf der anderen Seite dominieren die USA das Infrastrukturniveau der zentralen Server und der Tiefseekabel. In der physischen Welt ist das Internet immer noch an ein physisches Netzwerk mit einem signifikanten Zentralisierungsgrad gebunden. Das US-amerikanische Unternehmen *Equinix* steuert laut Firmenwebseite mit eigenen IXPs und Co-Location von Client-Computern in ihren Rechenzentren rund 90%(!) der Datenübertragung des Internets.

[28] vgl. Westerheide 2020
[29] vgl. Mozur/Metz 2020
[30] vgl. Mozur/Metz 2020
[31] vgl. FAZ 2019, S.17

China hat den Eindruck, dass die USA den Cyberspace dominieren, während sich die USA durch Chinas Aktionen im Cyberspace bedroht fühlen, siehe den Streit um 5G und *Huawei* im Jahr 2019[32].
Das NSCAI ist aber der Ansicht, dass die USA immer noch keine glaubwürdige Alternative zum chinesischen Anbieter *Huawei* für 5G haben[33]. Dies ist ein großes Sicherheitsproblem, da 5G-Netzwerke eine Art „Bindegewebe" zwischen den KI-Anwendungen darstellen[34].

2.5 Die Balance zwischen Cyber- und physischen Fähigkeiten

Computer und KI können menschliche Aktivitäten unterstützen und ersetzen und dadurch die nachrichtendienstlichen und militärischen Fähigkeiten eines Landes erhöhen. Diese Methode ermöglicht es High-Tech-Nationen mit großen Volkswirtschaften, ihre Macht zu konsolidieren und zu erweitern.

2017 hat jedoch das Pentagon, genauer gesagt, das *Strategic Studies Institute (SSI) des U.S. Army War College*, eine Studie aufgelegt, die von dem sog. **Post Primacy-Szenario** ausgeht [35], in dem die USA zwar immer noch die größte Wirtschafts- und Militärmacht sind, sie jedoch aufgrund der stärker werdenden Konkurrenten wie China nicht mehr imstande sind, die globale Weltordnung maßgeblich zu gestalten, so dass Geostrategie nun neu in einer instabilen, multipolaren und nicht mehr unbedingt von westlichen Werten dominierten Welt gedacht werden muss.

Eine Analyse australischer Militärfachleute zu den Fähigkeiten der USA hat gezeigt,[36] dass die Fähigkeit der USA zur Durchsetzung der liberalen Ordnung zurückgegangen ist, da die USA und ihre Verbündeten 1995 noch 80% der weltweiten Verteidigungsausgaben abdeckten, aber mittlerweile nur noch 52%.[37]
Die militärische Ausrüstung ist überlastet und überaltert und mit zunehmenden Unfällen behaftet, was auf nahezu andauernde Kämpfe im Nahen und Mittleren Osten sowie auf instabile Finanzplanungen aufgrund von Schuldenkrise und Parlamentsstreitigkeiten sowie

---

[32] Von westlichen Ländern wurden Sicherheitsbedenken gegen das chinesische Unternehmen *Huawei* geäußert, da dieses mittlerweile einer der größten globalen Smartphone-Hersteller und auch einer der größten Infrastrukturanbieter, insbesondere von Funkmasten für Smartphones und anderen Datenverkehr ist. Die nächste Internet-Kommunikationsgeneration 5G kommt, die erstmals eine breite Umsetzung des Internets der Dinge und intelligenter Home- und Smart City-Lösungen, insbesondere durch deutlich höhere Datenströme, Echtzeitübertragung, massiv reduzierte Latenzzeiten (Übertragungsverzögerungen) unter 1 Millisekunde und einem reduzierten Energiebedarf für die Übertragung pro Bit ermöglichen wird, vgl. Giesen/Mascolo/Tanriverdi 2018
[33] vgl. NSCAI 2020, S.54
[34] vgl. NSCAI 2020, S.55
[35] Lovelace 2017 schreibt im Vorwort: *"The U.S. Department of Defense (DoD) faces persistent fundamental change in its strategic and operating environments. This report suggests this reality is the product of the United States entering or being in the midst of a new, more competitive, post-U.S. primacy environment. Post-primacy conditions promise far-reaching impacts on U.S. national security and defense strategy. Consequently, there is an urgent requirement for DoD to examine and adapt how it develops strategy and describes, identifies, assesses, and communicates corporate-level risk"*
[36] United States Studies Centre 2019
[37] United States Studies Centre 2019, S.11

auf Trainingskürzungen zurückzuführen sind[38]. Es gibt ein wachsendes Missverhältnis zwischen Strategie und Ressourcen.
Die Schlussfolgerung ist, dass dies (eigene Übersetzung, danach Originaltext): *„…harte strategische Entscheidungen erfordert, die die Vereinigten Staaten möglicherweise nicht treffen wollen oder können. In einer Zeit knapper Budgets und multiplizierender geopolitischer Brennpunkte bedeutet die Priorisierung des Großmachtwettbewerbs mit China, dass die amerikanischen Streitkräfte anderweitige globale Aufgaben reduzieren müssen. Eine wachsende Anzahl von Verteidigungsplanern versteht die Notwendigkeit eines Kompromisses. Aber die politischen Führer und ein Großteil des außenpolitischen Establishments sind nach wie vor einem Supermacht-Danken verhaftet, das Amerikas Rolle in der Welt in der Verteidigung einer expansiven liberalen Ordnung sieht."*
[Original]*"…requires hard strategic choices which the United States may be unwilling or unable to make. In an era of constrained budgets and multiplying geopolitical flashpoints, prioritizing great power competition with China means America's armed forces must scale back other global responsibilities. A growing number of defense planners understand this trade-off. But political leaders and much of the foreign policy establishment remain wedded to a superpower mindset that regards America's role in the world as defending an expansive liberal order."* [39]
'Kompromiss' bedeutet hier, die Belastung beim Umgang mit mehreren sekundären Prioritäten zu verringern, um das primäre Ziel zu erreichen.

Zusammenfassend lässt sich sagen, dass der Fokus auf Cyber- und KI-Aktivitäten die Macht eines Staates nur erweitern wird, wenn auch die physischen Fähigkeiten erhalten und aufeinander abgestimmt werden. Andernfalls ist die Handlungsfreiheit trotz verbesserter Kenntnisse und Technologien gefährdet.
Eine aktuelle Diskussion zur Digitalisierung der Spionage kam zu dem Schluss, dass digitale Spionage letztlich die bisherige Arbeit nur ergänzen, aber keinesfalls den Agenten vor Ort ersetzen kann.

2.6 Die KI-Strategie der Europäischen Union

Die Europäische Kommission hat kürzlich ein Weißbuch über künstliche Intelligenz (*White Paper on Artificial Intelligence)* veröffentlicht und unterstützt einen regulatorischen und investitionsorientierten Ansatz mit dem Ziel, die KI zu fördern und die damit verbundenen Risiken vor dem Hintergrund des „harten globalen Wettbewerbs" (original: *"a background of fierce global competition"*) anzugehen.[40]
Ziel ist es, ein weltweit führender Anbieter von Innovationen in der Datenwirtschaft und ihren Anwendungen zu werden, jedoch mit einem regulatorischen Ökosystem des Vertrauens (**ecosystem of trust**) in diese sich schnell entwickelnden Technologien.
Um dies zu erreichen, richtete die Kommission eine hochrangige Expertengruppe (*High-Level Expert Group)* ein, die im April 2019 Leitlinien für vertrauenswürdige KI mit sieben

[38] United States Studies Centre 2019, z.B. unter anderem auf S.47-48
[39] United States Studies Centre 2019, S.9
[40] vgl. EC 2020

Hauptanforderungen veröffentlichte: menschliche Handlungsfähigkeit und Aufsicht, technische Robustheit und Sicherheit, Datenschutz und Datenverwaltung, Transparenz, Vielfalt, Nichtdiskriminierung und Fairness, gesellschaftliches und ökologisches Wohlbefinden sowie Rechenschaftspflicht. Ferner wurde ein Bericht über die Auswirkungen der künstlichen Intelligenz, des Internet der Dinge und der Robotik auf Sicherheit und Haftung (*Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics)* erstellt. Die EU hat jedoch bisher keine klare Strategie für die militärische Dimension der KI[41].
Die Europäische Union verbessert laufend die Finanzierung, betont jedoch die Notwendigkeit, die Anstrengungen zu verstärken, da 2016 in Europa rund 3,2 Mrd. EUR in KI investiert wurden, verglichen mit rund 12,1 Mrd. EUR in Nordamerika und 6,5 Mrd. EUR in Asien[42].

3. Militärische Aspekte

3.1 Eine einführende Fallstudie: Das Eurosur-Projekt

Dieses Projekt war nicht für militärische Zwecke gedacht, zeigt jedoch sehr deutlich die Vision vollständig integrierter autonomer Kontrollsysteme.
In der Europäischen Union sind verschiedene Forschungsprojekte im Gange, bei denen die Steuerung von Drohnen im Alltagsbetrieb nicht mehr von Menschen, sondern von Computern übernommen werden soll. Relevante Projekte sind das zur inneren Sicherheit zählende INDECT-Projekt seit 2009[43] und verschiedene weitere als Teil der Sicherung der europäischen Außengrenzen als *European Border Surveillance System (EUROSUR)* von 2008 bis 2012.

Zu den *Eurosur*-Projekten gehörten hier[44]:

- OPARUS (*Open Architecture for UAV-based Surveillance Systems*) zur Grenzüberwachung aus der Luft, bei dem es auch um die Eingliederung der Drohnen in den zivilen Luftraum ging
- TALOS (*Transportable autonomous patrol for land border surveillance*) mit Patrouillenmaschinen
- WIMAAS (*Wide Maritime area airborne surveillance*) zur Nutzung von Drohnen zur Seeüberwachung

[41] vgl. Franke 2019
[42] vgl. EC 2020, S.4
[43] vgl. Welchering 2013a, S. T6. Die Forschung zur automatischen Erkennung von Bedrohungssituationen richtet sich auf Szenarien wie das folgende: Falls eine Kamera ein verdächtiges Verhalten feststellt, soll die Kombination aus automatisch aktivierten Beobachtungsdrohnen, Richtmikrofonen und automatisierter Gesichtserkennung die Identifikation der Zielperson und ggf. ihrer Absichten ermöglichen. Falls nötig, sollen auch Daten aus Facebook, Twitter, Google plus, Kreditkartendaten usw. genutzt werden, um gefährliche Handlungen zu erkennen.
[44] vgl. Oparus 2010, SEC 2011, S.7, Talos Cooperation 2012.

Die Idee, die Alltagsüberwachung von einem Computer steuern zu lassen, dem *Unmanned Units Command Center UUCC*, war ein Teil dieser Projekte, aber aus einer Cyberwar-Perspektive wäre das die entscheidende Schwachstelle, so dass höchste Anforderungen für die Cybersicherheit und –stabilität gestellt werden müssten.

3.2 Praktische Anwendungen

3.2.1 Unmanned Aerial Vehicles (UAVs, Drohnen)

Drohnen, auch bekannt als unbemannte Luftfahrzeuge (**Unmanned Aerial Vehicles UAVs**), sind mittlerweile fortschrittliche Waffen mit wachsender Systemautonomie. Andererseits hat auch die Verteidigung gegen Drohnen erhebliche Fortschritte gemacht.
Die **Drohnen** dienen nicht mehr nur der Aufklärung, sondern können auch in Gefechten eingesetzt werden. Drohnen eignen sich generell für alle Arten von Operationen, die „dull, dirty, dangerous or difficult" sind[45].
Drohnen ermöglichen die Beobachtung und/oder gezielte Tötung von Gegnern als *Lethal Autonomous Weapons Systems (LAWS)*[46]. Der technische Fortschritt ermöglicht immer umfangreichere **Assistenzfunktionen**, d.h. die menschliche Entscheidung immer weitgehender von Computern unterstützt und beeinflusst[47]. In diesem Zusammenhang kam bereits die Frage einer **Haftbarkeit von Maschinen** auf[48]. Jeder Schritt in Richtung vollautomatisierter Drohnen würde jedenfalls deutlich verstärkte Anstrengungen im Bereich der Cyber-Sicherheit erfordern, um zu verhindern, dass die Maschinen von gegnerischen Hackern übernommen werden [49].
Autonome Drohnen können ihre Entdeckung durch Halten von Funkstille vermeiden, so dass die Autonomie Teil eines Tarnkappendrohnenkonzepts ist wie bei der 2013 von China getesteten **Lijan-**Drohne[50].

Das *Drone Databook* von 2019 fasst die Verfügbarkeit und Forschung von Drohnen in 101 Ländern zusammen und verwendet die Klassifizierung des *NATO Standardization Agreement 4670* von I bis III, die weitgehend auf ihrem maximalen Startgewicht basiert: Class I (weniger als 150 Kilogramm, typischerweise *Micro, Mini, and Small Drones*), Class II (150 bis 600 Kilogramm, typischerweise "taktische" UAVs), und Class III (über 600 Kilogramm as *"medium-altitude long-endurance" (MALE)* or *"high-altitude long-endurance" (HALE)* UAVs) [51].

---

[45] vgl. Jahn 2011, S.26: also alles, was „langweilig, schmutzig, gefährlich, schwierig oder anders" ist
[46] vgl. Thiel 2012, S. Z2
[47] Eine mögliche Zukunft mit vollautomatisierten Tötungen bleibt jedoch Spekulation. Die Erforschung von autonomen Kampfrobotern (**lethal autonomous robots LARs**) macht Fortschritte, vgl. Klüver 2013, S.2.
[48] Im zivilen Sektor wird dies in den USA für selbstfahrende Autos (also Autos mit Autopilot-Funktionen) diskutiert, Kalifornien plante entsprechende Reglungen schon für 2015, vgl. Burianski 2012, S.21
[49] Die größten Drohnen sind mittlerweile in der Lage, konventionelle Flugzeuge zu ersetzen, so dass ein gegnerisches Eindringen ein erhebliches Sicherheitsrisiko darstellt. Das europäische Drohnenprojekt **Neuron** ist ein unbemanntes Kampfflugzeug (*unmanned aerial combat vehicle UACV*) mit Tarnkappen (Stealth)-Technologie, welches zu größeren Schlägen aus der Luft als bisherige Drohnen fähig sein soll (vgl. Bittner/Ladurner 2012, S.3; Hanke 2012, S.14).
[50] vgl. TAZ online 2013
[51] vgl. Gettinger 2019, S.IV

Am wichtigsten ist, dass derzeit mindestens 24 Länder neue unbemannte Militärflugzeuge entwickeln (10 Systeme der Klasse I, 12 Systeme der Klasse II und 36 Systeme der Klasse III). Mindestens sieben Länder erforschen Drohnen der nächsten Generation, darunter auch Stealth-Flugkörper (US, China, Russland und Frankreich), sehr hoch fliegende high-altitude pseudo-satellites (US, China, UK), Schwärme (US, China, UK), and teaming systems aus bemannten und unbemannten Systemen (Australien, Japan, UK, China und US) [52].
Schwärme sind KI-basierte Drohnen, die autonom sind (nicht unter zentraler Kontrolle) und in der Lage sind, ihre lokale Umgebung und andere in der Nähe befindliche Schwarmteilnehmer zu erfassen, lokal mit anderen im Schwarm zu kommunizieren und bei der Ausführung einer bestimmten Aufgabe zusammenzuarbeiten[53].

Chinas Drohnenentwicklung konzentriert sich auf eine breite Palette von Klasse III-Drohnen[54]. Drei aktuelle US-Projekte für KI-Drohnen sind *Valkyrie, Skyborg* und *Gremlins*[55].

- XQ-58A *Valkyrie* ist ein Class III UAV mit Jetantrieb des Air Force *Low-Cost Attritable Strike Demonstrator (LCASD)* bzw *Loyal Wingman* -Projekts, das bemannte Flugzeuge in den Kampf begleiten und z.B. feindliche Luftverteidigungen angreifen kann. Der erste Flug fand 2019 statt.
- *Skyborg* ist ein Air Force-Konzept für eine autonome preiswerte Kampfdrohne, die als Mittel zum Testen verschiedener Technologien für künstliche Intelligenz dienen kann, die komplexe, autonome Operationen ermöglichen würden. Ein zukünftige *Skyborg* Drohne könnte mit *Valkyrie* kooperieren, Testluftkämpfe gegen bemannte Jets sind für 2021 geplant.
- *Gremlins* ist ein DARPA-Programm zur Entwicklung preiswerter wiederverwendbarer Drohnenschwärme die z.B. für Aufklärungsmissionen oder elektronische Kampfführung verwendet werden könnten.

Das Funktionieren autonomer Maschinen ist von der zugrunde liegenden Programmierung abhängig, was jedoch zu ethischen und praktischen Dilemmata führen kann[56]. Falls das programmierte Verhalten bekannt ist, könnten Drohnen (wie Autos) durch Vortäuschung von bestimmten Situationen oder Objekten absichtlich irregeführt, abgefangen oder zerstört werden.

Die wichtigsten Möglichkeiten, Drohnen anzugreifen, sind:

- **Drone hacking**: Mit der **Battle Management Language** werden Befehle auf vordefinierten Frequenzen gesendet. Die begrenzten Kosten und Anstrengungen, die für solche Angriffe erforderlich sind, sind ein wichtiges Sicherheitsrisiko für Militärs[57].

[52] vgl. Gettinger 2019, S.XV
[53] vgl. Hoadley/Sayler 2019, S.14
[54] vgl. Gettinger 2019, S.16
[55] vgl. Gettinger 2019, S.245
[56] vgl. Hevelke/Nida-Rümelin 2015, S.82
[57] vgl. Welchering 2017

- **GPS-Spoofing** von Drohnen: Das Senden falscher Koordinaten an die Drohnen kann sie irreführen oder sogar zwingen, eine Notlandung zu machen
- **Jamming:** Überschwemmungen mit elektromagnetischen Signalen können eine Notlandung hervorrufen, die die Zerstörung oder sogar eine Sicherstellung der angegriffenen Drohnen ermöglicht.
- **Physische Angriffe:** Das Abschießen von Drohnen, aber auch die Erfassung von Drohnen, auch durch speziell ausgebildete Tiere, bilden einen wachsenden Markt für Sicherheitsfirmen. Auch die Drohnen-Abwehr durch Laserwaffen befindet sich in der Entwicklung.
- **Kommunikationsverlust:** Die *EuroHawk*-Drohne kombinierte die Drohnentechnologie der *Global Hawk*-Drohne von *Northrop Grumann* mit dem neuartigen hochentwickelten Aufklärungssystem *ISIS (Integrated Signal Intelligence System)* der EADS-Tochter *Cassidian*. Während eines Überführungsfluges nach Europa riss der Kontakt für einige wenige Minuten ab. Da solche Zeitfenster potentielle Gelegenheiten für (Cyber-)Angriffe sein können, ist die Cybersicherheit für zukünftige Entwicklungen besonders wichtig.

Irakische Aufständische konnten mit einer Software in die Videosysteme unbemannter US-Drohnen eindringen und so die Videos dieser Drohnen mit ansehen konnten[58]. 2011 wurde berichtet, dass die Computer der *Creech Air Force Base* in Nevada, die als Steuerzentrale für *Predator*- und *Reaper*- Drohnen dient, von einem Computervirus befallen wurden; laut US Air Force hatte dies jedoch keinen Einfluss auf die Einsatzfähigkeit der Drohnen[59]. Der Iran brachte 2011 eine US-Drohne vom Typ RQ-170 in seinen Besitz[60].
Die Verwundbarkeit von Drohnen ist aber auch typabhängig, da diese mit unterschiedlichen Kontrollmethoden und verschieden großer Systemautonomie gesteuert werden[61].

Die Drohnentechnologie leidet unter bestimmten Schwachstellen, die sich im Verlust einer relevanten Zahl von Drohnen widerspiegelt. Meistens wurden diese Verluste durch Bedienungsfehler und konventionelle technische Probleme verursacht.
Eine systematische Untersuchung der *Washington Pos*t fand 418 Drohnenabstürze im Zeitraum von 2001 bis 2014, wesentliche Ursachen waren beschränkte Möglichkeiten von Kameras und Sensoren zur Kollisionsvermeidung, Pilotenfehler, mechanische Defekte und unzuverlässige Kommunikationsverbindungen[62].

Tests in New Mexico im Jahre 2012 haben die Anfälligkeit von Drohnen für falsche GPS-Signale (**GPS spoofing**) nachgewiesen. Dies galt auch für die neue Flugüberwachung durch *Automatic Dependent Surveillance Broadcast Systems (ADS-B)*. Auch hat man

[58] vgl. Ladurner/Pham 2010, S.12
[59] vgl. Los Angeles Times 13 October 2011
[60] vgl. Bittner/Ladurner 2012, S.3. Als Eindringmethode wurde die Verwendung eines manipulierten GPS-Signals (GPS spoofing) diskutiert, aber das konnte nicht belegt werden.
[61] vgl. Heider 2006, S.9
[62] vgl. Whitlock 2014

festgestellt, dass Drohnen unbeabsichtigt durch Signale, die an andere Drohnen gerichtet sind, abgelenkt werden können.[63]

3.2.2 Autonome Fahrzeuge

Sowohl die USA als auch China arbeiten daran, KI in halbautonome und autonome Fahrzeuge (**semiautonomous** and **autonomous vehicles)** zu integrieren, in den USA auch Kampfflugzeuge (wie das Projekt *Loyal Wingman*), Drohnen, Bodenfahrzeuge (wie das ferngesteuerte *Multi-Utility Tactical Transport MUTT* des Marine Corps), und für die See den *Anti-Submarine Warfare Continuous Trail Unmanned Vessel*-Prototyp, auch bekannt als *Sea Hunter*[64].

3.2.3 Intelligence, Surveillance, and Reconnaissance (ISR)

Es wird erwartet, dass KI in den Bereichen der Überwachung im Rahmen der **Intelligence, Surveillance, and Reconnaissance** (ISR) besonders nützlich ist, da große Datenmengen für die Analyse wie im oben genannten Projekt *Maven* zur Verfügung stehen. Aber **Imaging Intelligence** ist jedoch mehr als nur Zielidentifikation oder Gesichtserkennung, so überwachen zum Beispiel die amerikanischen Dienste *Defense Intelligence Agency (DIA)* und CIA zugangsbeschränkte Gebäude ihrer Gegner zur Analyse der Aktivitäten[65]. Satelliten zum Beispiel überprüfen täglich die Aktivitäten chinesischer Krankenhäuser, indem sie die Autos auf den umliegenden Parkplätzen genau zählen. In einer kürzlich durchgeführten Studie wurde im Herbst 2019 ein massiver Höhepunkt beobachtet, der möglicherweise ein frühes Anzeichen für die Coronavirus-Pandemie war, da eine Analyse des chinesischen Internets in derselben Studie ergab, dass chinesische Benutzer in Wuhan zunehmend mit *Baidu* nach den Begriffen Husten und Durchfall suchten.

3.2.4 Command and Control-Systeme

Command and Control-Systeme mit KI-Elementen werden in China und den USA erforscht. Die US Air Force entwickelt das *Multi-Domain Command and Control (MDC2)* zur Zentralisierung der Planung und Durchführung von Luft-, Raumfahrt-, Cyberspace-, See- und Landoperationen.[66]

3.2.5 Logistik

KI kann auch die militärische Logistik unterstützen[67], die *Defense Innovation Unit (DIU)* und die *US Air Force* arbeiten mit dem JAIC an **Predictive Maintenance**-Lösungen für zukünftige Wartungsanforderungen an Geräten, anstatt Reparaturen durchzuführen oder sich an standardisierte Wartungspläne zu halten[68]. Für den F-35 Jet, werden Echtzeit-

[63] vgl. Humphreys/Wesson 2014, S.82
[64] vgl. Hoadley/Sayler 2019, S.14
[65] vgl. Folmer/Margolin 2020
[66] vgl. Hoadley/Sayler 2019, S.12
[67] vgl. Hoadley/Sayler 2019, S.10
[68] vgl. DoD 2018, S.11

Sensordaten, die in die Triebwerke des Flugzeugs und andere Bordsysteme eingebettet sind, in einen Vorhersagealgorithmus eingegeben, um zu bestimmen, wann Techniker das Flugzeug inspizieren oder Teile ersetzen müssen[69].

[69] vgl. DoD 2018, Hoadley/Sayler 2019

4. Sicherheitsaspekte

4.1 Kurze Einführung

KI-Systeme können manipuliert, umgangen und irregeführt werden, was tiefgreifende Auswirkungen auf die Sicherheit von Anwendungen wie Netzwerküberwachungstools, Finanzsystemen oder autonomen Fahrzeugen hat[70]. KI hat mit Computern, Hardware und Software zu tun, sodass alle gängigen Bedrohungen für digitale Systeme auch für KI-Systeme gemeinsame Bedrohungen darstellen. Eine vollständige Darstellung der Cyber-Bedrohungen findet sich im Cyberwar-Papier in Abschnitt 7.2.
Darüber hinaus gibt es KI-spezifische Schwachstellen, die detaillierter dargestellt werden müssen. Da die Komplexität von KI-Systemen rasch zunimmt, ist es ungewiss, ob diese Probleme gelöst oder in Zukunft sogar noch verschärft werden könnten. Die Software von KI-Systemen kann gestohlen werden, d.h. Cyberspionage kann den gesamten Vorteil von KI-Systemen beseitigen.
Andererseits kann KI die Cyber-Verteidigung bis hin zur automatisierten Cyber-Verteidigung erheblich verbessern und eine Waffe in der Informationskriegsführung sein.

4.2 Cyber-Attacken

Die Angreifer sind die sogenannten Hacker, die nach Schwachstellen in Programmen und Systemen suchen, um dann mit eigenen Programmen wie den *Viren* oder *Trojanern* die Kontrolle zu übernehmen. Die Nutzer versuchen sie dazu zu bringen, schädliche Anhänge oder Internetseiten zu öffnen, Passwörter und Kontodaten preiszugeben.
Es gibt 4 große Ziele, nämlich die Normalnutzer, die Privatwirtschaft, den Staat mit Politik, Verwaltung und öffentlichen Einrichtungen, und die Kritischen Infrastrukturen, die man zum Leben braucht, wie z.B. Strom- und Wasserversorgung, Krankenhäuser usw.
Am häufigsten greifen Kriminelle an, dann die Geheimdienste, während Terroristen und Cyberarmeen noch kaum in Erscheinung traten.
Kriminelle Hacker stehlen Daten, um diese zu verkaufen oder um das Konto des Opfers zu plündern. Oder sie nutzen *Ransomware* genannte Bildschirmsperren, um Geld für die Entfernung zu fordern. Manchmal benutzen sie die Rechner auch, um damit weitere Opfer anzugreifen oder um digitales Geld zu erschaffen (**bitcoin or crypto mining**).
Geheimdienste haben Hackerteams, sogenannte **Advanced Persistent Threats (APTs)**, die in Politik, Wirtschaft und Technik einschließlich der Sabotage aktiv sind. APTs sind längerfristig agierende Angreifergruppen mit definierten **Techniken, Taktiken und Programmen (TTPs)**. Die letzten Jahre haben jedoch gezeigt, dass eine APT eine Projektgruppe innerhalb eines Nachrichtendienstes ist, die ihre TTPs sowie die Angriffsziele entlang der operativen Vorgaben ihres Dienstes entwickelt und anwendet. APTs bilden sich nicht von selber, sie werden durch Zusammenstellung geeigneter Leute gebildet und ihre Cyberaktivitäten an den Zielvorgaben ausgerichtet.
Fremde Staatsapparate sind immer interessant und stehen ständig unter Druck, während die Normalnutzer weniger im Fokus sind, weil es schwierig ist, aus der Vielzahl etwas Brauchbares herauszufiltern (die Nadel im Heuhaufen).

[70] vgl. NSTC 2020, S.1

Die Hacker der Wirtschaftsspionage plündern Forschungseinrichtungen, High Tech- und Rüstungsfirmen. Sabotage-Hacker attackieren Fabriken und kritische Infrastrukturen, was schon zu Stromausfällen geführt hat. Man kann unter anderem Produktionen stören, Daten löschen und Digitalgeräte beschädigen oder direkt die Computerchips.

4.3 Wichtige Schwachstellen von KI-Systemen

4.3.1 Grundlegende Probleme der KI

Die frühen KI-Systeme waren einfach gebaut und daher leicht zu erklären. Inzwischen sind jedoch **Deep Neural Networks** entstanden, die sehr gute Ergebnisse zeigen, jedoch auf Deep Learning-Modellen basieren, die Lernalgorithmen mit bis zu Hunderten von versteckten „neuronalen" Schichten und Millionen von Parametern kombinieren, wodurch sie zu undurchsichtigen Black-Box-Systemen werden. Dies ist auch als **Explainability Issue** (Erklärbarkeitsproblem) bekannt[71].

Die Arten von KI-Algorithmen mit der höchsten Leistung können ihre Prozesse derzeit nicht erklären. Zum Beispiel hat *Google* ein effektives System zur Identifizierung von Katzen in Filmen geschaffen, aber niemand konnte erklären, welches Element einer Katze die Identifizierung ermöglichte. Dieser Mangel an sogenannter "Erklärbarkeit" ist allen solchen KI-Algorithmen gemeinsam[72]. Es gibt jedoch eine Diskussion darüber, dass Maschinen manchmal gemeinsame Muster oder Strukturen in Objektklassen sehen, die Menschen zuvor einfach nicht bemerkt haben.

Infolgedessen kann niemand vorhersagen, wann und aus welchem Grund ein Fehler auftreten kann, und KI-Systeme sind nur begrenzt vorhersehbar (**predictability issue**).

Systematische Fehler: KI-Systemfehler können ein erhebliches Risiko darstellen, wenn die Systeme in großem Maßstab bereitgestellt werden, d.h. KI-Systeme könnten dann gleichzeitig und auf die gleiche Weise versagen und möglicherweise große oder zerstörerische Auswirkungen haben.

Kommunikationsprobleme: 5G-Netzwerke werden eine Art „Bindegewebe" zwischen KI-Anwendungen sein, was bedeutet, dass jeder, der auf die 5G-Netzwerke zugreifen kann, die Kommunikation beeinflussen (verändern, stören) kann.[73]

Missbrauch der Rechenleistung: Die reine Geschwindigkeit der KI macht die Systeme für den Missbrauch sehr attraktiv, z.B. für das Erschaffen (Mining) von Kryptowährung, die viele Berechnungen erfordert[74].

[71] vgl. Arrieta et al. 2020, S.83
[72] vgl. Hoadley/Sayler 2019, S.31
[73] vgl. NSCAI 2020, S.55
[74] vgl. Goddins 2020

4.3.2 Missionsstabilität

Ein spezifisches militärisches KI-Problem ist die **Missionsstabilität**[75]. Autonome militärische Systeme können die Aufklärung und die Informationslage verbessern, die Entscheidungsfindung beschleunigen und schnelle Reaktionen ermöglichen, aber auch militärische Missionen destabilisieren.
Beispiele:

- Eine autonome Drohne kann beschließen, ein relevantes Ziel anzugreifen, auf diese Weise jedoch die militärische Präsenz offenlegen und Spezialeinheiten oder Geheimdienstoperationen gefährden.
- Bei der DARPA *Cyber Challenge 2016* war der beste Computer eine Maschine, die sich auf Kosten von ihr betreuten Verteidigungssysteme selbst verteidigte.
- Ein Computer kann entscheiden, dass ein Kampf an einem bestimmten Ort eine Verschwendung von Ressourcen darstellt, und z.B. einen Drohnenschwarm zurückziehen, aber vielleicht nie verstehen, dass manchmal ein bestimmter Ort einen symbolischen und psychologischen Wert hat oder vielleicht als Ankerpunkt einer neuen Frontlinie vorgesehen ist oder dass der Kampf nur dazu dient, Gegner von wichtigeren Bereichen abzulenken. Die Frage ist: Wird eine fortgeschrittene militärische KI wirklich strategisch oder nur taktisch denken können? Der Kontext wird von den Systemen immer noch sehr schlecht verstanden, d.h. ihnen fehlt der gesunde Menschenverstand[76].
- Missionsautoritätsproblem: In Zivilflugzeugen mussten Piloten bereits gegen defekte Autopiloten kämpfen, die in kritischen Situationen nicht außer Kraft gesetzt werden konnten[77].
- Eine KI kann sich zu schnell entscheiden, zu kämpfen, und so die konventionellen Streitkräfte unvorbereitet zu lassen oder die Tür zu einer friedlichen Lösung zu schließen.
- Ein gehacktes KI-System kann gegen seinen Kontrolleur umgedreht oder als Doppelagent verwendet werden (d.h. es sendet Beobachtungen beider Seiten an beide Seiten).

Schlussfolgerung: Je weiter fortgeschritten eine militärische KI ist, desto höher ist das Risiko einer Missionsinstabilität, die plötzlich in Mikrosekunden auftreten kann.

4.3.3 Daten-Manipulation

- **Manipulierte Bilder** können autonome Systeme verwirren. Kleine Aufkleber auf der Straße reichten aus, um den Autopiloten eines Tesla-Fahrzeugs auf die

[75] vgl. Masuhr 2019, Johnson 2020
[76] vgl. Wright 2020, S.7
[77] Voke 2019 schrieb in seiner Analyse auf Seite 33: [Übersetzung] *„Wenn KI unangemessene Absichten zeigt oder schlecht handelt, muss der Mensch in der Lage sein, die KI außer Kraft zu setzen. Auch wenn das System nicht die erforderliche Leistung erbrachte, muss der Mensch in der Lage sein, die Kontrolle auszuüben, sobald eine gefährliche Situation erkannt wird. Transparenz ist eine Voraussetzung für Kontrolle, und Kontrolle ist eine Voraussetzung für Vertrauen.“ „Moreover, if AI is showing improper intentions or acting poorly, humans must be able to override its behavior. Although the system did not perform as required, the human must be able to exercise control once recognition of a hazardous situation occurs. Transparency is a requirement for control, and control is a requirement for trust.“*

gegenüberliegende Fahrspur zu lenken[78]. Mittlerweile gibt es auf modernen chinesischen Militärfahrzeugen, aber auch auf russischen Hubschraubern Tarnbilder im Pixelstil.
Bereits kleinste - für das menschliche Auge unsichtbare - Änderungen in digitalen Bildern können zu systematischen Fehlinterpretationen durch die KI führen, ein Prozess, der als adverses maschinelles Lernen (**adversarial machine learning)** bezeichnet wird[79].

- **Data poisoning** ('Datenvergiftung'): Maschinen können durch falsch beschriftete Daten systematisch irregeführt werden. Dies kann durch Tapes auf Stoppschildern für den Verkehr geschehen[80], aber möglicherweise könnte der Missbrauch von Militärflaggen und -symbolen eine andere Option sein.
- **Attrappen** könnten sicherlich sogar autonome Kampfdrohnen irreführen.
- **Spoofing:** Irreführung von GPS-gesteuerten Systemen, indem sie ein falsches GPS-Signal senden, das das richtige Signal überlagert, z.B. gegen Drohnen oder Schiffe

4.3.4 Hardware in der KI

Ein fortschrittliches KI-System kann möglicherweise jeden Angriff erkennen und darauf reagieren, hat jedoch keine Chance, sich gegen Hardwarefehler innerhalb der KI zu verteidigen, die es sogar ermöglichen können, das KI-System von außen zu übernehmen:

- **Gefälschte Mikrochips**

Jedoch fürchten die USA selber Hintertüren, z.B. als versteckte Funktionen in Chips, weshalb keine asiatischen Chips mehr in sicherheitsrelevanter US-Technologie verwendet werden sollen. Aus demselben Grunde will das US State Department auch keine chinesischen Computer mehr verwenden. Gleichwohl lässt sich die Nutzung kommerzieller Produkte, englisch **commercial off-the-shelf (COTS) technology**, in sicherheitsrelevanten Bereichen trotz der dadurch erhöhten Anfälligkeit nicht ganz vermeiden. Nicht nur Hersteller, sondern auch die globalen Lieferketten bilden mögliche Angriffspunkte: eine Studie des US-Senats von 2012 berichtete, dass in US-Waffen mehr als eine Million gefälschter Chips installiert wurden, 70% der Chips kamen aus China, aber relevante Mengen stammten auch aus Großbritannien und Kanada[81]. Um diesem Problem entgegenzuwirken, hat die US-Regierung strenge Anforderungen an die in KI-Systemen verwendete Mikroelektronik definiert[82].

- **Veränderte Platinen (motherboards)**

China produziert 75 Prozent der Mobiltelefone und 90 Prozent aller PCs, da selbst US-Unternehmen diesen Produktionsschritt nach China auslagern. Laut einem umstrittenen *Bloomberg*-Bericht könnten Subunternehmer in China von der Hardware-Hacking-Einheit der chinesischen PLA unter Druck gesetzt worden sein, diese zusätzlichen Chips

[78] vgl. FAS 2019, S.21
[79] vgl. Wolff 2020
[80] vgl. Wolff 2020
[81] vgl. Fahrion 2012, S.1
[82] vgl. NSCAI 2020, S.60

einzubauen, die eine totale Hintergrundkontrolle ermöglichen würden[83]. Die Firma *Super Micro* ist ein Anbieter von Server-Motherboards (Platinen). Während einer Evaluation des Software-Unternehmens *Elemental Technologies* durch *Amazon Web Services (AWS)* wurde ein winziger Mikrochip gefunden, ein bisschen größer als ein Reiskorn, und der nicht Teil des ursprünglichen Designs war[84]. *Elemental Technologies*, die seit 2009 Entwicklungspartner der CIA-Firma *In-Q-Tel* ist, stellte Server für die DoD-Rechenzentren, die Drohnenoperationen der CIA und für Kriegsschiffe zur Verfügung.

Fuzzing: Die vielleicht stärkste Cyberwaffe ist das **Fuzzing**, das Verschicken von Zufallscodes an Chips, das militärisch weitreichende Konsequenzen hat: die USA haben um 2007 die Verwendung chinesischer Chips in den Waffensystemen gestoppt, aus Furcht im Gefecht abgeschaltet werden zu können. Weiter oben wurde bereits gezeigt, dass viele Chips störanfällig durch Fuzzing sind. Die Chiphersteller versuchen, die Lücken zu schließen, es werden aber ständig neue entdeckt. So sollten Chips in der existierenden Militärtechnik intensiv getestet werden, damit nicht plötzlich die Lichter ausgehen, wenn sie dem Feind zu nahekommen. Einer dieser Zufallsbefehle trägt den Namen „*halt and catch fire*“ der den Computerchip irreparabel abschaltet. Auch wenn dieser Befehl nur bei bestimmten Chips zur Ausführung gebracht werden konnte und Einzelheiten verständlicherweise geheim blieben, zeigt er, dass ein ‚**digitaler Rettungsschuß**‘ zumindest technisch möglich ist[85].

4.4 Cyberverteidigung

4.4.1 Detektion von Cyberangriffen

Sicherheitsfirmen können Angriffe im Rahmen der **Threat Intelligence** mit Angriffsmuster-Datenbanken abgleichen, aber auch im Rahmen der **Intrusion Detection** den Datenverkehr auf ungewöhnliche Vorgänge und statistische Auffälligkeiten abklopfen. **Threat Intelligence Repositories** vergleichen eingehende Informationen mit bekannten IP-Adressen, Domainnamen, Webseiten und auch mit Listen bekannter bösartiger Attachments. Dies ermöglicht eine sofortige Erkennung und manchmal sogar die Zuordnung eines eingehenden Angriffs.
Die US-Regierung baut im Moment hochentwickelte Sensorsysteme aus[86]: Das *Continuous Diagnostics and Mitigation (CDM)*-Programm kann abnormes Verhalten in Echtzeit erkennen und entsprechende Übersichtsberichte an Administratoren erstellen.
Einstein 3A arbeitet mit Sensoren an Webzugangspunkten, um Bedrohungen aus dem zu schützenden System herauszuhalten, während das CDM Bedrohungen identifizieren soll, wenn sie schon im System sind.

US-Forscher haben **Mustererkennungsalgorithmen** zur Cyberabwehr entwickelt, die im Falle eines erkannten Angriffes die Löschung von Datenpaketen des Angreifers erlauben.

---

[83] vgl. Robertson/Riley 2018
[84] vgl. Robertson/Riley 2018
[85] Man muss aber anmerken, dass in der Fuzzing-Forschung schon früher Befehle auffielen, die die Chipfunktionen störten, wobei dies wohl zunächst eher als lästiges Testhindernis betrachtet wurde.
[86] vgl. Gerstein 2015, S.4-5

Zur Vermeidung von Eskalationen ist jedoch keine automatisierte Vergeltung vorgesehen. China erforscht Simulationen von Cyberattacken.

KI-Methoden zur Erkennung von Cyberangriffen umfassen die Erkennung oder Kategorisierung von Malware, Netzwerkeinbrüchen, Phishing- und Spam-Angriffen, um so *Advanced Persistent Threats (APTs)* entgegenzuwirken. und Domänen identifizieren, die durch Domänengenerierungsalgorithmen (*domain generation algorithms DGAs*) erzeugt wird [87].

4.4.2 Automatisierte Cyberabwehr

Die *DARPA* führte am 04.08.2016 die *Cyber Grand Challenge* in Las Vegas durch, wobei 7 Computer Cyberattacken wahrnahmen und vollautomatisch, d.h. ohne jeden menschlichen Eingriff, darauf reagierten. Dieser Wettbewerb ging über 12 Stunden und 30 Runden. Die Computer und ihre Programmierteams wurden aus hundert Bewerbern ausgewählt[88].
Eine Maschine namens *Mayhem* gewann den Wettbewerb, indem sie die meiste Zeit über passiv blieb, während die anderen sich gegenseitig bekämpften. Eine andere Maschine nahm eine Sicherheitslücke wahr, der von ihr hergestellte ‚Patch' verlangsamte jedoch die Maschine, so dass die Maschine entschied, den Patch besser wieder zu entfernen [89].

Die DARPA war mit dem Ergebnis zufrieden, da es ein erster Schritt in Richtung vollautomatischer Abwehr- und Reaktionssysteme war[90]. Da die Zahl der Sicherheitslücken inzwischen immens ist[91], könnten automatisierte Systeme unbekannte Lücken wahrnehmen und stoppen.

Während es möglich sein mag, die Routineüberwachung an Maschinen zu übertragen, wird die menschliche Aufsicht unverzichtbar bleiben. Andernfalls könnte eine irregeführte (gespoofte) Maschine sich entschließen, das eigene Netzwerk anzugreifen. Oder ein Angreifer könnte die Maschine davon überzeugen, in den inaktiven Zustand überzugehen oder einen Patch zu konstruieren, der das Verteidigungssystem lahmlegt.

4.5 Informationskrieg

Das Konzept des Informationskrieges ist gut etabliert, z.B. in der psychologischen Kriegsführung, bei der gezielte Informationen oder Propaganda wurde an die freigegeben wurde, um das Verhalten zu beeinflussen. Der moderne Informationskrieg ist etwas anders gelagert, denn dies ist die *kombinierte Manipulation von digitalen Technologien und Informationen*, um Gegner zu beeinflussen.

[87] vgl. Truong/Diep/Celinka 2020, S.24
[88] vgl. DARPA 2016
[89] vgl. Atherton 2016
[90] vgl. DARPA 2016
[91] Eine US-Datenbank hat 75.000 Sicherheitslücken in 2015 gesammelt, vgl. Betschon 2016; in einem Test fand das Pentagon 138 Sicherheitslücken in seinen Systemen, vgl. Die Welt online 2016

Bots (automatische Akteure und Kommunikatoren im Internet) sind mittlerweile weitverbreitet und man kann große Mengen an fake tweets erzeugen und menschliche Kommunikation vortäuschen (**social bots, internet of thingies**)[92].
Eine neue Variante ist sogenannter *fake traffic*. In einem Test konnte eine fake traffic software von einem Computer aus 100,000 Klicks auf eine einzige Website ausführen, aber es so aussehen lassen, als wenn jeder Klick von einem anderen Computer gekommen wäre.

Im Sommer 2017 wurde von der University of Oxford eine Studie über **Computational Propaganda** veröffentlicht. Ein Team von 12 Forschern bewertete die Situation in 9 Ländern[93]. Die Autoren definieren die Computational Propaganda als „den Einsatz von Algorithmen, Automatisierung und menschlicher Bearbeitung, um irreführende Informationen über soziale Medien gezielt zu verteilen" [*„as the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks"*.] Derzeit sind *Facebook* und *Twitter* die wichtigsten Plattformen für diese Aktivitäten. Die EU hat eine *Task Force* gegründet, die gefälschte Nachrichten erkennen, sie korrigieren und auch eine positive Wahrnehmung der EU in den östlichen Staaten unterstützen sollte[94].

Informationen können als politische Waffe eingesetzt werden. In der Vergangenheit wurde dies (unter Bezugnahme auf den russischen Begriff) **Kompromat** genannt, der reale und/oder erfundene Fakten über politische Gegner enthielt, um sie zu schwächen. KI ermöglicht zunehmend realistische Foto-, Audio- und Videofälschungen oder „**deep fakes**".[95]

Eine moderne Version des Kompromates sind **Leaks**, die von staatlichen oder nicht-staatlichen Akteuren stammen können. Während diese ursprünglich typischerweise ins Internet gestellt wurden, während inzwischen zunehmend abgehörte bzw. aufgezeichnete vertrauliche Telefonate von führenden Politikern veröffentlicht.
Künstliche Intelligenz könnte auch verwendet werden, um vollständige digitale Lebensmuster (full digital patterns-of-life) zu erstellen und so ein umfassendes Profil von Servicemitgliedern, mutmaßlichen Geheimdienstmitarbeitern, Regierungsbeamten oder Privatpersonen zu erstellen[96]. Ein bereits etablierter gezielterer Ansatz ist die Erstellung von **User-Profilen**. Im März 2012 hat *Google* bekanntgegeben, dass Profile durch Verknüpfungen von Suchmaschinennutzungen, *YouTube*, *Google plus* und *gmail* erstellt werden[97]. Ähnliche Prozeduren sind auch von Betreiberfirmen sozialer Netzwerke bekannt.

Ein weiterer Ansatz ist die **digital dust analysis**. Wenn in Russland oder China ein neues US-Botschaftsmitglied angekündigt wird, wird nicht nur die Menge, sondern auch die Verbreitung digitaler Informationen überprüft. Wenn der digitale Fußabdruck des

[92] vgl. Graff 2014, S.13, Brundage et al. 2018, S.43-49
[93] vgl. Woolley/Howard 2017
[94] vgl. Stabenow 2017, S.3
[95] vgl. Hoadley/Sayler 2019, S.11-12
[96] vgl. Hoadley/Sayler 2019, S.12
[97] vgl. Spiegel 2013, S.111

Newcomers zu klein ist, wie Social Media-Beiträge, Handy-Telefonate, Kreditkarten-Zahlungen, dann wird der Diplomat als Undercover-CIA-Offizier markiert[98].

5. Ethik und Maschinen-Logik

Es gibt viele Aspekte der KI, die ethische Probleme verursachen können, z.B. im militärischen Bereich, wenn die automatisierte Entscheidungsfindung zur Tötung von Gegnern führen kann. Es gilt als selbstverständlich, dass KI-Systeme eine menschliche Aufsicht oder zumindest eine Notübersteuerung bei offensichtlichen Fehlfunktionen ermöglichen sollten.
Eine weitere Herausforderung ist das Problem der Vorhersehbarkeit (**predictability**) und Erklärbarkeit (**explainability**). Die spezifischen Merkmale vieler KI-Technologien, einschließlich Intransparenz („Black-Box-Effekt"), Komplexität, Unvorhersehbarkeit und teilweise autonomen Verhalten können es schwierig machen, die Einhaltung von Rechtsregeln zum Schutz von Grundrechten zu überprüfen, und so deren wirksame Durchsetzung behindern[99]. Bestimmte KI-Algorithmen können geschlechtsspezifische und rassistische Vorurteile integrieren, z.B. zur Gesichtsanalyse. Menschliche Entscheidungen können auch voreingenommen sein, aber die gleiche Voreingenommenheit in weit verbreiteten KI-Systemen könnte einen viel größeren Effekt haben und viele Menschen betreffen und diskriminieren[100].

Während es möglich ist, dass sich KI-Forscher und ihre Länder ethischen und gesellschaftlichen Werten verpflichtet fühlen, ist es derzeit, wo KI ein begrenztes Verständnis der Situationskontexte hat, sehr schwierig, sich eine KI mit eingebetteten Werten vorzustellen. Zum Beispiel haben Menschen normalerweise eine klare Vorstellung davon, was Würde, Gerechtigkeit und Fairness für sie bedeuten, aber wie könnten diese Begriffe im Programmcode oder in Maschinensprache aussehen?

Ein klassisches Problem der Maschinenethik und -logik ist das **Kollisionsdilemma** autonomer Autos[101]: Ein Fußgänger kann plötzlich die Straße überqueren und das autonome Autosystem kann mit zwei Optionen konfrontiert werden, nämlich Ausweichen mit dem Risiko des Todes des Fahrers oder Weiterfahren mit dem Risiko des Todes des Fußgängers.

Eine starke künstliche Intelligenz, d.h. ein System mit der Fähigkeit, nach dem Sinn zu fragen und mit einem autonomen Selbst (*cogito ergo sum*) wird - basierend auf überlegenem Wissen und Intelligenz - wahrscheinlich nicht eher der menschlichen Logik und Ethik folgen. Im DARPA-Wettbewerb 2016 hat die Maschine gewonnen, die sich selbst gerettet hat, anstatt die Verteidigungssysteme dauerhaft aktiv zu halten.

[98] vgl. Rohde 2016
[99] vgl. EC 2020, S.11-12
[100] vgl. EC 2020, S.11-12
[101] vgl. Hevelke/Nida-Rümelin 2015

6. Abschließende Bemerkungen

In diesem Papier wurden militärische und sicherheitstechnische Aspekte der künstlichen Intelligenz (KI) als neuer Bereich der Sicherheitspolitik vorgestellt. Bereits aktuelle KI-Systeme sind in der Lage, menschliche Aktivitäten in wesentlichen Bereichen des täglichen Lebens, der Kommunikation, des Handels, der Industrie usw. zu unterstützen oder zu ersetzen und alle Arten des Maschinennutzens zu unterstützen oder zu steuern, was das massive Wachstum der KI und ihr enormes Potenzial erklärt. Die USA und China konkurrieren um die Technologieführerschaft in der KI, gefolgt von Europa. Die militärischen Projekte konzentrieren sich auf unbemannte und autonome Fahrzeuge, C2- (Command and Control) und Intelligence-, Surveillance- und Reconnaissance- (ISR) Programme. China und die USA sind in Bezug auf personelle und technische Ressourcen miteinander verbunden. Eine im Stil des Kalten Krieges denkbare Aufteilung in zwei getrennte Cyber- und KI-Welten könnte sowohl für beide Staaten als auch für den Fortschritt der KI erhebliche Probleme verursachen. Es wurde gezeigt, dass der Fokus auf Cyber- und KI-Aktivitäten die Macht eines Staates nur erweitern wird, wenn auch die physischen militärischen Fähigkeiten erhalten bleiben.

KI-Systeme haben ein spezifisches Cybersicherheitsprofil, sie können zur Erkennung von Cyberangriffen und zur automatisierten Cyberabwehr dienen, weisen jedoch auch komplexe Schwachstellen auf, wie gezielten Manipulationen von Input-Daten und Bildern ausgenutzt werden können. Da die Komplexität von KI-Systemen rasch zunimmt, ist es ungewiss, ob diese Probleme gelöst oder sich in Zukunft sogar noch verschärfen könnten. KI-Systeme sind auch für die Informationskriegsführung von wachsender Bedeutung, insbesondere für deep fakes.

Maschinenlogik und -ethik sind ein weiteres herausforderndes Thema. Eine starke künstliche Intelligenz, d.h. ein System mit der Fähigkeit, nach dem Sinn zu fragen und mit einem autonomen Selbst (*cogito ergo sum*) wird - basierend auf überlegenem Wissen und Intelligenz - wahrscheinlich nicht eher der menschlichen Logik und Ethik folgen.

7. Literatur

7.1 Literaturquellen

Arrieta, A.B. et al. (2020): Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. Information Fusion 58 (2020), p.82–111

Atherton, K.D. (2016): DARPA's Cyber Grand Challenge Ends In Triumph. Popular Science 06 Aug 2016, 2 pages

Betschon, S. (2016): Die Crux mit gefälschten Chips. Neue Zürcher Zeitung 31.08.2016, S.39

Bittner, J., Ladurner, U. (2012): Die Waffe der Überflieger. Die Zeit Nr.50/2012, S.2-3

Bommakanti, K. (2020): A.I. in the Chinese Military: Current Initiatives and the Implications for India Observer Research Foundation (ORF) Occasional Paper 234 February 2020

Brundage, M. et al. (2018): The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. Future of Humanity Institute University of Oxford/Centre for the Study of Existential Risk University of Cambridge/Center for a New American Security/Electronic Frontier Foundation/OpenAI February 2018

Burianski, M. (2012): Maschinen können nicht haften. Frankfurter Allgemeine Zeitung Nr. 272/2012, S.21

Danchin A., Fang, G. (2016): Unknown unknowns: essential genes in quest for function. Microb Biotechnol. 2016 Sep;9(5):530-40. doi: 10.1111/1751-7915.12384. Epub 2016 Jul 20

Die Welt online (2016): Pentagon: Hacker finden bei Test 138 Sicherheitslücken. http://www.welt.de/newsticker/news1/article156330187, 1 S.

DoD (2018): U.S. Department of Defense, Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity

EC (2020): White Paper On Artificial Intelligence - A European approach to excellence and trust Brussels, 19.2.2020 COM(2020) 65 final

Elbadawi M., Efferth T. (2020): Organoids of human airways to study infectivity and cytopathy of SARS-CoV-2. Lancet Respir Med 2020 Published Online May 21, 2020 https://doi.org/10.1016/S2213-2600(20)30238-1

Fahrion, G. (2012): Pfusch am Gewehr. Financial Times Deutschland, 23.05.2012, S.1

FAS (2019): Sicherheitsexperten manipulieren Teslas Autopiloten. Frankfurter Allgemeine Sonntagszeitung Nr. 9, 03.04.2019, S.21

FAZ (2019): Amerika will mehr seltene Erden fördern. Frankfurter Allgemeine Zeitung, Nr.130, S.17

Floemer, A. (2020): Teslas Modell 3 ist VW und Toyota technisch um sechs Jahre voraus. Welt Online 19.02.2020

Folmer, K., Margolin, J. (2020): Satellite data suggest Coronavirus may have hit China earlier: Researchers. ABC News online, 08 June 2020

Franke, U.E. (2019): Not smart enough: The poverty of European military thinking on artificial intelligence – ECFR/311 December 2019

Gerstein, D.M. (2015): Strategies for Defending U.S. Government Networks in Cyberspace. RAND Office of External Affairs Document CT-436 June 2015, 7 pages

Gettinger, D. (2019): The Drone Databook. The Center for the Study of The Drone at Bard College, 353 pages

Giesen, C., Mascolo, G. and Tanriverdi, H. (2018): Hört, hört. Süddeutsche Zeitung 14.12.2018, S.3

Goddins, D. (2020): Machine-learning clusters in Azure hijacked to mine cryptocurrency. Ars Technica, 11 June 2020

Graff, B. (2014): Sie sind da. Süddeutsche Zeitung Nr. 107, 10/11.05.2014, S.13

Hanke, T. (2012): Erfolgreicher Probeflug der europäischen Kampfdrohne. Handelsblatt 03.12.2012, S.14-15

Heide, M., Huttner W.B. and Mora-Bermudez, F. (2018): Brain organoid models for neocortex development and evolution. Current Opinion in Cell Biology 2018, 55:8–1

Heider, D. (2006): Drohnen im zivilen und militärischen Einsatz. University of Münster 01 Feb 2006, 10 pages

Hevelke, A., Nida-Rümelin, J. (2015): Intelligente Autos im Dilemma. Spektrum der Wissenschaft Oktober 2015, S.82-85

Hoadley D.S., Sayler, K.M. (2019): Artificial Intelligence and National Security Congressional Research Service R45178 Version 6 Updated November 21, 2019

Humphreys, T./Wesson, K. (2014): Drohnen auf Abwegen. Spektrum der Wissenschaft (German Edition of Scientific American) März 2014, S.82-86

Jahn, T. (2011): Das Milliardengeschäft mit den Drohnen. Handelsblatt 25.11.2011, S.26

Johnson, J.S. (2020): Artificial Intelligence: A Threat to Strategic Stability. Strategic Studies Quarterly Spring 2020, S.16-39

Jung, A. (2020): Ära der Cobots. Der Spiegel 25/2020, 13.06.2020, S.70-71

Kastilan, S. (2010): Vier Flaschen für ein Heureka. Frankfurter Allgemeine Zeitung 21.05.2010, S.33

Klüver, R. (2013): Automaten des Todes. Süddeutsche Zeitung Nr 187/2013, S.2

Lachance J.C., Rodrigue S., Palsson B.O. (2019): Minimal cells, maximal knowledge. Elife. 2019 Mar 12;8. pii: e45379. doi: 10.7554/eLife.45379.

Ladurner, U., Pham, K. (2010): Iran im Krieg 2.0. Die Zeit Nr.40, S.12

Los Angeles Times (2011): Air Force says drone computer viruses pose ‘no threat’. Los Angeles Times online 13 October 2011, 11:26 am

Lovelace, DC Jr. (2017): in: The Strategic Studies Institute (SSI) and U.S. Army War College Press. At our own peril: DoD risk assessment in a post-primacy world. Principal Author and Project Director: Nathan P. Freier. June 2017

Masuhr, N. (2019): AI in Military Enabling Applications. CSS Analyses in Security Policy No. 251, October 2019

Mozur, P., Metz, C. (2020): A U.S. Secret Weapon in A.I.: Chinese Talent New York Times online 09 June 2020

NATO (2019): Artificial Intelligence: Implications for NATO’s Armed Forces. Science and Technology Committee (STC) - Sub-Committee on Technology Trends and Security (STCTTS) Rapporteur: Matej Tonin (Slovenia) 149 STCTTS 19 E rev. 1 fin Original: English 13 October 2019

NDAA (2019): National Defense Authorization Act (NDAA) United States of America 2019

NSCAI (2020): National Security Commission on Artificial Intelligence First Quarter Recommendations March 2020, 131 pages

NSTC (2020): Artificial Intelligence and Cybersecurity: Opportunities and Challenges Technical Workshop Summary Report - A report by the Networking & Information Technology Research and Development Subcommittee and the Machine Learning & Artificial Intelligence Subcommittee of the National Science & Technology Council March 2020

Oparus (2010): Oparus Overview and Objectives. Website of the OPARUS project, 3 pages, oparus.eu

OSTP (2020): American Artificial Intelligence Initiative: Year One Annual Report. Prepared by The White House Office of Science and Technology Policy February 2020

Park, S.J. et al. (2016): Phototactic guidance of a tissue-engineered soft-robotic ray. Science 08 Jul 2016: Vol. 353, Issue 6295, pp.158-162

Perez J.A., Deligianni, F., Ravi D. and Yan G.Z. (2019): Artificial Intelligence and Robotics. The UK-RAS Network

RAND (2019): The Department of Defense Posture for Artificial Intelligence. Rand Corporation Document RR4229 Santa Monica, USA

Robertson, J., Riley, M. (2018): How China used a tiny chip to infiltrate America’s top companies. Bloomberg Businessweek 04 Oct 2018

Rohde, D. (2016): Is the CIA ready for the age of Cyberwar? The Atlantic online 02 Nov 2016

SEC (2011): Commission Staff Working Paper. Determining the technical and operational framework of the European Border Surveillance System (EUROSUR) and the actions to be taken for its establishment. Brussels, 28 Jan 2011, SEC (2011) 145 final 11 pages

Spiegel (2013): Verdacht statt Vertrauen, Der Spiegel 26/2013, S.111

Stabenow, M. (2017): Warnung in roten Lettern. Frankfurter Allgemeine Zeitung 25.01.2017, S.3

Talos Cooperation (2012): Transportable Autonomous Patrol for Land Border Surveillance D.10.3 4th Workshop 25 May 2012

TAZ online (2013): China testet das "scharfe Schwert". 23.11.2013, 4 Seiten

Thiel, T. (2012): Auf der sicheren Seite. Frankfurter Allgemeine Zeitung No. 281/2012, S.Z1-Z2

Trump, D.J. (2019): Donald J. Trump, Executive Order 13859 on Maintaining American Leadership in Artificial Intelligence, Washington, D.C.: The White House, February 11, 2019.

Truong, T.C., Diep, Q.B. and Zelinka, I: (2020): Artificial Intelligence in the Cyber Domain: Offense and Defense Symmetry 2020, 12, 410; doi:10.3390/sym12030410 www.mdpi.com/journal/symmetry

United States Studies Centre (2019): Townshend A. and Brendan Thomas-Noone with Matilda Steward "Averting crisis: American strategy, military spending and collective defence in the Indo-Pacific," United States Studies Centre at the University of Sydney, August 2019

Voke, M.R. (2019): Artificial Intelligence for Command and Control of Air Power. Wright Flyer Paper No. 72 Air University Press

Wang F., Zhang W. (2019): Synthetic biology: Recent progress, biosafety and biosecurity concerns, and possible solutions. Journal of Biosafety and Biosecurity 1 (2019) 22–30

Welchering, P. (2013): Digitale Überwachungsaugen an jeder Ecke. Frankfurter Allgemeine Zeitung Nr. 110/2013, S.T6

Welchering, P. (2017): Cyberwar in der Luft - Hacker warnen vor Angriffen. Heute online Mai 2017

Westerheide, F. (2020): China – The First Artificial Intelligence Superpower. Forbes Cognitive World Contributor Group online 14 Jan 2020

Whitlock, C. (2014): When drones fall from the sky. Washington Post online from 20 June 2014

Wolff, J. (2020): How to Improve Cybersecurity for Artificial Intelligence. Brookings Report 08 June 2020

Woolley, S.C., Howard, P.N. (2017): Computational Propaganda –worldwide– Executive Summary. Working Paper No. 2017.11 University of Oxford, Project on Computational Propaganda 2017, 15 pages

Wright, N.D. (2019): Artificial Intelligence, China, Russia, and the Global Order Technological, Political, Global, and Creative Perspectives. Air University Press in October 2019

7.2 Literaturempfehlungen

2019 Political Warfare
https://nbn-resolving.org/urn:nbn:de:gbv:700-201909181987

2019 Cyberwar-Grundlagen-Methoden-Beispiele
https://repositorium.ub.uni-osnabrueck.de/handle/urn:nbn:de:gbv:700-201907091700

2019 Attribution of Cyber Attacks – Chapter 13 in: Reuter, C. (Hrsg.): Information Technology for Peace and Security, 279-303. ISBN: 9780128117361 Springer Vieweg. Abstract in www.springerprofessional.de/attribution-of-cyber-attacks/16544512