



15.05.2018

Künstliche Intelligenz – was ist der Kern der Revolution?

Anlass

Künstliche Intelligenz (kurz KI) ist der Hype des Jahres 2018. Tech-Unternehmen erzielen Milliardengewinne, Forschung verlagert sich von den Universitäten und Forschungszentren in die Zentralen von Alphabet Inc, Facebook oder Amazon, der Wettbewerb um Talente intensiviert sich. Daten sind der Rohstoff dieser Entwicklung – die allerdings vor allem in den USA und China systematisch gesammelt werden. Europa scheint den Anschluss zu verlieren.

Das ruft die Politik auf den Plan: Cédric Villani, charismatischer Mathematiker und Abgeordneter in der Nationalversammlung Frankreichs, präsentiert im Auftrage des Präsidenten einen Plan für eine „sinnvolle“ KI in Frankreich und Europa, das Oberhaus in London veröffentlicht kurz darauf eine Übersicht über Chancen und Herausforderungen der KI für Großbritannien. Forscher fordern unterdessen in einem offenen Brief ein „European Lab for Learning and Intelligent Systems, ELLIS“. Deutschland bietet zu diesem Zeitpunkt das „Weißbuch Arbeit 4.0“, es geht auf mögliche Folgen für Arbeiter und Angestellte ein. Ein umfassendes Konzept für den Umgang mit KI fehlt hierzulande bislang. Das will die deutsche Politik nun nachholen: Forschungs- und Bildungsministerin Anja Karliczek will mit Forschern und Vertretern der Wirtschaft bis zum Sommer 2018 erste konkrete Schritte abstimmen.

Doch was kann KI? Was nicht? Was sind die technischen Probleme, und worauf muss sich die Gesellschaft einstellen? Dazu soll dieses Fact Sheet einen knappen Einblick bieten. Vorweg noch ein Wort zum Begriff: Kern der derzeit boomenden KI ist das so genannte maschinelle Lernen. Da künstliche Intelligenz aber deutlich mehr ist als nur maschinelles Lernen, die heute diskutierten Fälle aber fast ausschließlich dem Bereich des maschinellen Lernens angehören, bleiben wir im Folgenden beim genaueren Begriff "maschinelles Lernen".

Übersicht

| | |
|---------------------------------------------------------------|---|
| Warum maschinelles Lernen?..... | 2 |
| Wie funktioniert das Lernen beim maschinellen Lernen? | 2 |
| Welche Lernmethoden gibt es? | 3 |
| Welche Lernstrukturen gibt es?..... | 4 |
| Wo liegen derzeit die technischen Probleme? | 5 |
| Folgen des maschinellen Lernens – was wird sich ändern? | 6 |



| | |
|--------------------------------------------|---|
| Literaturstellen, die zitiert wurden | 8 |
| Weitere Recherchequellen | 9 |

Warum maschinelles Lernen?

Maschinelles Lernen ist eine Teildisziplin der Forschung an einer menschenähnlichen Künstlichen Intelligenz. Schließlich soll ein Computer, der wie ein Mensch denken kann, auch lernen können. Seit den neunziger Jahren setzen Forscher ihre Programme zunehmend auch kommerziell ein [1, S.21]. Wo immer es um große Datenmengen geht, die schnell durchsucht werden müssen, um neue oder bekannte Muster zu entdecken, setzen Programmierer heute auf maschinelles Lernen:

- ▶ Prognosen oder Mustererkennung für Diagnosen, Spracherkennung oder Bilderkennungen sind in vielen Fällen zu komplex und vielfältig, um quasi von Hand alle möglichen Interpretationen zu programmieren, zum Beispiel als wenn-dann Regeln.
- ▶ Wenn in Daten neue Muster erkannt werden sollen – zum Beispiel in Kundenverhalten oder für Empfehlungen (Recommender-Systeme) – ist es erst recht sehr schwer, mit konventionellen Programmen danach in Datensätzen zu suchen.
- ▶ Einer der Vorteile von maschinellem Lernen ist, dass solche Algorithmen auch Probleme lösen können, bei denen man nicht genau weiß, welchen Regeln das System unterliegt – zum Beispiel in der Forschung an Proteinen oder auch in der Hirnforschung [2].
- ▶ Statt jeden einzelnen Schritt auf dem Weg zur Lösung eines Problems explizit zu beschreiben, geben Programmierer dem Rechner beim maschinellen Lernen einen aufbereiteten Datensatz, eine Lerntechnik und eine Aufgabe vor.
- ▶ Vom Ziel, ein selbstdenkendes System zu entwickeln, sind die Forscher jedoch noch sehr weit entfernt [3, S.2, 5].

Wie funktioniert das Lernen beim maschinellen Lernen?

Per Definition ist maschinelles Lernen ein Programm, das sein Ergebnis selbstständig verbessert. Computerlernen hat dabei jedoch bisher wenig bis gar nichts mit menschlichem Lernen zu tun, eher schon mit Üben oder Training. Was vergleichbar ist, ist das häufige Wiederholen einer Aufgabe und die dabei auftretende schrittweise Verbesserung der Ergebnisse [6]. Kern sind dabei **mathematische und statistische Regeln und Verfahren**. Grundsätzlich funktionieren diese auch vollkommen ohne Computer, wie Donald Michi 1961 gezeigt hat: Er entwarf eine Streichholzschachtelmaschine (MENACE), die quasi „lernen“ kann, Tic-Tac-Toe zu spielen [4] [vgl. 5]. Sie ist ein gutes Beispiel, um das Prinzip maschinellen Lernens zu verstehen:

- ▶ Im ersten Schritt muss das Problem analysiert und für den Rechner strukturiert werden. Michi zum Beispiel stellte nach einer Analyse der möglichen Spielzüge Schachteln auf, die er nach Spielstand und möglichen Zügen ordnete und beschriftete. Die Schachtel für einen bestimmten Brettzustand, bei dem die „Maschine“ am Zug war, enthielt Informationen darüber, welcher nächste Spielzug am besten wäre. Diese Information veränderte sich mit der „Spielerfahrung“.
- ▶ Eine Lernmethode muss gefunden werden, mit der der Rechner agieren kann, und die sich von Programmdurchlauf zu Programmdurchlauf verbessert. Michi wählte ein Verfahren, das er **reinforcement process** (heute reinforcement learning, s.u.) bezeichnete [4, S. 232.]: Ein erfolgreicher



Programmdurchlauf erhält ein positives Feedback in allen Stufen, ein erfolgloser ein negatives [4, S.232] vgl. [6] [7].

- ▶ Diese Lernmethode braucht einen mathematischen Prozess, der zu besseren Ergebnissen führt. Michi setzte dafür auf die Wahrscheinlichkeit:
 - Für jedes der neun Spielfelder legte er eine Farbe fest. Da eine Schachtel ja für einen Brettzustand steht, war klar, welche Felder noch frei sind. In diese Schachtel füllte er entsprechend farbige Perlen – wobei es möglich war, für ein Feld mehrere Perlen einzufüllen.
 - Er öffnete die Schachtel, die dem Spielstand entsprach, nahm ohne hinzusehen – also zufällig – eine farbige Perle heraus und setzte das Kreuz ins entsprechende Feld. Das Fach ließ er offen, die Perle davor liegen.
- ▶ Hatte die Maschine gewonnen, füllte er drei zusätzliche Perlen der Farben in die geöffneten Schachteln, die davor lagen. Er gab damit positives Feedback und erhöhte die Wahrscheinlichkeit, beim nächsten Spiel die erfolgreichen Züge zu wiederholen.
- ▶ Hatte die Maschine verloren, legte er die Perlen nicht wieder in die Fächer zurück, gab also negatives Feedback. Er reduzierte die Wahrscheinlichkeit, erfolglose Züge zu wiederholen.
- ▶ Die Maschine kann jedoch nicht so „lernen“, dass sie stets gewinnt, weil jeder Spieler anders spielt.
- ▶ Nach diesem Prinzip, einer starren Struktur, in deren Inneren mathematische Prozesse bei jedem Durchlauf zu neuen, besseren Ergebnisse führen, funktioniert maschinelles Lernen. Die Maschine selbst nimmt dabei nichts wahr, gewinnt keine Erkenntnisse, schließt nichts.

Welche Lernmethoden gibt es?

Programmierer und Forscher unterscheiden beim maschinellen Lernen heute im Wesentlichen drei Methoden, die sich zum Teil auch kombinieren lassen: supervised learning, unsupervised learning und reinforcement learning.

- ▶ Das **reinforcement learning** funktioniert im Prinzip immer noch mit positiven und negativen Feedback, wie Michi es beschrieben hat. Es eignet sich besonders gut für Programme, die mit anderen Akteuren interagieren sollen – zum Beispiel in Spielen (GO [8], Poker) – aber auch zur Programmierung von Robotern.
- ▶ Beim **supervised learning** (überwachten Lernen) lernt der Computer zunächst anhand entsprechend gekennzeichnete Beispiel-Datensätze Muster zu erkennen, bevor er sie in einem unbekanntem Datenbestand finden soll. Wenn er zum Beispiel zwischen Hunden und Katzen auf Bildern unterscheiden soll, lassen ihn Programmierer zuerst anhand von Beispielbildern, die von Experten bereits als Hunde- und Katzenbilder gekennzeichnet (gelabelt) sind, lernen zu unterscheiden. Erst dann durchsucht das Programm einen ungekennzeichneten Datensatz.
- ▶ Das überwachte Lernen eignet sich, um **Vorhersagen** zu treffen [6] [7], also: Ist eine E-Mail Spam? Wer oder was ist auf diesem Bild zu sehen? Hat ein Patient eine bestimmte Krankheit?
- ▶ Beim **unsupervised learning** (unüberwachten Lernen) sortiert oder durchsucht das Programm Daten ohne konkrete Vorgaben (ohne Label). Dabei gibt es keinen Weg, das einzelne Ergebnis auf Richtigkeit zu prüfen. Das Ziel hierbei ist weniger eine bestimmte Repräsentation der Daten, sondern eine möglichst sinnvolle Repräsentation.
- ▶ Es eignet sich zum Beispiel für die Untersuchung von Kundenverhalten aus Daten, die beim Besuch von Webseiten hinterlassen wurden. Diese Aufgabe wurde und wird gelegentlich auch als **Data Mining** [9 Kapitel 9] bezeichnet.



- ▶ Beim Training mit überwachtem und unüberwachtem Lernen kommt es darauf an, die Daten für den Prozess sorgfältig auszuwählen, um Probleme und Risiken wie zum Beispiel Fehlklassifikationen zu minimieren (s.u.).

Das Training selbst kann permanent erfolgen – etwa bei Computer Vision. Das Programm verbessert seine Ergebnisse im Betrieb weiter. Oder es ist irgendwann abgeschlossen. Dann wird das Programm bis zu einer gewünschten Qualität der Ergebnisse trainiert und danach eingesetzt.

Welche Lernstrukturen gibt es?

Donald Michi wählte als Struktur für seine Aufgabe eine ausgeklügelte Ansammlung von Schachteln, die er mit Perlen füllte. Strukturen, in denen sich das Lernen in Form eines Entwickelns statistischer Korrelationen vollzieht, verwenden Programmierer heute noch immer. Zu den wichtigsten Architekturen beim maschinellen Lernen, die teilweise kombiniert werden, zählen:

- ▶ **Neuronale Netze:** Ihr Aufbau folgt vereinfacht gesagt den neuronalen Vernetzungen im menschlichen Gehirn [7], sie wurden schon in den 60er Jahren konzipiert [10]. Die „Neuronen“ sind dabei Punkte, an denen Werte der Daten mit zunächst zufälligen Werten multipliziert werden. Sie sind in Schichten angeordnet, jeder Punkt einer Schicht ist mit mehreren oder allen in der folgenden verbunden.
- ▶ Neuronale Netze „lernen“, indem die Werte, die an den Punkten mit den Datenwerten multipliziert werden, Schritt für Schritt so lange verändert werden, bis das Programm das gewünschte Ergebnis liefert.
- ▶ Sie sind in vielen Bereichen bereits gängige Werkzeuge für die Datenverarbeitung, insbesondere bei riesigen Datenmengen. Sie eignen sich für supervised und unsupervised learning. In der Forschung hoffen Wissenschaftler, mit neuronalen Netzen der Funktion und den Eigenschaften von Neuronen im Gehirn näher zu kommen.
- ▶ **Deep Neural Networks / Deep Learning:** Neuronale Netze mit besonders vielen Schichten. Sie sind die Technologie, die den KI-Trend derzeit am stärksten vorantreiben [6] [10] und sind zum Beispiel im Bereich der Mustererkennung und Spracherkennung sehr erfolgreich [6] [11]. Sie werden seit Mitte der 2000er erfolgreich eingesetzt und sind in der Lage, Probleme zu lösen, für die mit wenigen Schichten nur schlechte Ergebnisse erzielt werden konnten [9 Kapitel 2]. Seitdem spezielle Grafik-Chips für die Berechnungen eingesetzt werden, hat sich die Leistung noch einmal stark erhöht. Deep Learning eignet sich für supervised und unsupervised learning sowie eine Kombination aus beiden [6] [7] [9 Kapitel 9], die Methode ist besonders stark bei unsupervised learning.
- ▶ **Support Vector Machine:** eine unanschauliche Methode, um Daten in Klassen aufzuteilen. Die mathematische Basis wurde schon in den dreißiger Jahren geschaffen. Als Architektur für die Klassifizierung von Daten hat sich diese Technik seit Mitte der 90er Jahre verbreitet. Sie eignet sich für supervised learning und wird zum Beispiel für die Gesichtserkennung eingesetzt, aber auch in der Erforschung komplexer Prozesse – etwa zum Klassifizieren von Proteinstrukturen.
- ▶ **Entscheidungsbäume:** ähneln tatsächlich ein bisschen einem Baum; ausgehend von der Wurzel werden Daten an jeder Astgabel nach einer Eigenschaft bewertet. Sie eignen sich für supervised learning und werden für Prognosen genutzt, etwa um die Kreditwürdigkeit von Bankkunden zu klassifizieren oder ihre Kaufkraft vorherzusagen [12].
- ▶ **Bayes'sches Netz:** Eine Methode, um aus unterschiedlichen Wahrscheinlichkeiten auf mögliche Ursachen oder Entwicklungen zu schließen. Geht zurück auf den britischen Mathematiker Thomas Bayes (1702-1761). Bayes'sche Netze werden sowohl in der Forschung angewandt (zum Beispiel für die Analyse von Proteinstrukturen oder Gen-Expression), als auch in der Ingenieurwissenschaft oder für Risiko-Analysen.



- ▶ **Clustering:** Dabei versucht der Algorithmus Gruppen von ähnlichen Datenpunkten zu finden, ohne eine konkrete Vorgabe zu haben, wonach er sortieren soll [6]. Beim Clustering ist das Ziel, Gruppen zu erstellen, deren Angehörige sich möglichst ähnlich sind und die sich untereinander möglichst stark unterscheiden.
- ▶ Darüber hinaus gibt es noch eine Reihe von weiteren Verfahren, Untervarianten oder auch Kombinationen, um mit Hilfe von maschinellem Lernen aus Daten Erkenntnisse zu gewinnen.

Wo liegen derzeit die technischen Probleme?

Die meisten Fragen können weder von Menschen noch von Maschinen perfekt gelöst werden. Dies liegt zum Beispiel daran, dass es nicht für alle Entscheidungen Regeln mit 100-prozentiger Treffergenauigkeit gibt. Nicht alle Programme erzielen zudem nach dem Training die gewünschte Qualität. Wichtige Messgrößen für die Qualität eines trainierten Programms sind **Sensibilität** und **Spezifität**. Die **Sensibilität** gibt an, wie oft das Programm beispielsweise Bilder von Tumoren wirklich als Tumor erkennt. Die **Spezifität** gibt an, wie viele Bilder, die *keinen* Tumor enthalten, korrekt vom Programm als „ohne Tumor“ bewertet werden. Die Parameter stammen aus der Medizin, wo sie als Indikatoren für die Qualität von diagnostischen Tests dienen. Probleme, die diese beiden sowie weitere Kenngrößen beeinträchtigen können, und an denen Forscher arbeiten, sind die folgenden:

- ▶ **Nachvollziehbarkeit:** Besonders bei Neuronalen Netzwerken kann man nicht nachvollziehen, wie sie zu ihrer Lösung kommen. Folglich kann man auch keine Rückschlüsse auf die zugrundeliegenden Regeln ziehen [7]. Das ist ein Problem, wenn das Ergebnis überprüfbar sein muss, wie zum Beispiel für die medizinische Diagnose oder bei Prozessen vor Gericht, in denen die Ergebnisse dieser Programme eine wichtige Rolle spielen. Derzeit wird ein Forschungsbereich immer wichtiger, der sich damit beschäftigt zu verstehen, wie die Algorithmen zu ihren Ergebnissen und Vorhersagen kommen.
- ▶ **Kausalität:** Selbst in Fällen, in denen bekannt ist, worauf der Algorithmus seine Entscheidung begründet, sagt eine statistische Verbindung, die der Algorithmus zieht, zunächst nichts über eine mögliche Kausalität aus. Nur weil ein Merkmal statistisch als ein guter Indikator identifiziert wird, ist nicht gegeben, dass das Merkmal auch ursächlich mit dem zu lösenden Problem zusammenhängt.
- ▶ **Verzerrungen / Vorurteile:** Sie entstehen vor allem durch Lernen mit verzerrten Daten, etwa aufgrund der Erhebungsmethode – zum Beispiel beim Labeln von Bildern (Weißer Mann = Mann, Farbiger Mann = Farbiger). Fehler in der Datenbasis wird das Programm reproduzieren. So können sich Algorithmen auch menschliche Fehler und Vorurteile aneignen [13].
- ▶ **Zu kleiner Datensatz:** In einigen Fällen ist der Datensatz für das Training nicht groß genug, sodass einige Korrelationen überbewertet werden, die sich in einer größeren Stichprobe als weniger gute Indikatoren erweisen [14]. Wurden bei einem Bilderkennungsprogramm zum Beispiel zufällig alle Motive, die das Programm erkennen soll, bei schlechtem Wetter gemacht, kann es sein, dass das Programm das schlechte Wetter erkennt und nicht das gewünschte Motiv.
- ▶ **Auswendig lernen:** Manchmal findet der Algorithmus zwar einen Weg, die Trainingsdaten mit dem gewünschten Output in Verbindung zu setzen, für neue Daten ergibt der zugehörige Output aber keinen Sinn. Die Lösung ist also nicht verallgemeinerbar. Außerdem kann es passieren, dass Algorithmen nicht anhand von zugrundeliegenden Korrelationen zuordnen, sondern den Trainingsdatensatz „auswendig lernen“ (overfitting). Auch dann ist die Lösung nicht verallgemeinerbar.
- ▶ Verzerrungen, Datensatzumfang und auswendig lernen sind Gründe, warum maschinelles Lernen in der Regel nur mit großen Datensätzen gut funktioniert. Je mehr Daten es gibt, desto größer ist die Wahrscheinlichkeit, dass die Verteilung der Merkmale im Datensatz der realen Verteilung entspricht. Die Notwendigkeit von großen Datenmengen, insbesondere personenbezogener Daten, sorgt jedoch häufig für Datenschutzbedenken [6].



- ▶ Aktuelle Forschung befasst sich unter anderem daher damit, Modelle – vor allem im Bereich Deep Learning – zu entwickeln, die so wie Menschen auch anhand von wenigen Beispielen lernen können (one shot learning) [14]. Andere Forscher schlagen vor, Daten miteinander zu teilen [16]. Auf diese Weise könnte die Datenbasis über den Rahmen einzelner Unternehmen hinaus erweitert werden. Die Forderung wurde zum Beispiel nach einem Unfall eines Uber-Fahrzeugs erhoben [17]. Allerdings sind elektronisch gespeicherte Daten oft in sehr unterschiedlichen Formaten an vielen Stellen abgelegt. Zumindest was persönliche Daten angeht, könnte hier die EU-Datenschutzregelung übrigens Abhilfe schaffen: Sie verpflichtet Unternehmen, Daten von Nutzern elektronisch strukturiert und in einem üblichen Format unentgeltlich zur Verfügung zu stellen [18, S.30, 69].
- ▶ **Spezialisierung:** In der Regel kann ein System nur das, wofür es trainiert ist – Tumore erkennen, Sprache erkennen, Medikamenten-Mischungen empfehlen. Kaum ein System ist derzeit in der Lage, zwei Aufgaben zu erledigen, etwa Tumorzellen und Hirnzellen sicher zu erkennen. Dies ist ein aktuelles Thema in der Forschung. In der Regel kann sich ein Programm auch nicht von selbst an neue Herausforderungen anpassen. Speziell im Bereich der Robotik arbeiten Forscher daher an einer „Long-Term-Autonomy“, die es den Maschinen ermöglichen soll, über lange Zeit mit verschiedenen Menschen in unterschiedlicher Umgebung zu arbeiten [19].

▶ Was kann maschinelles Lernen nicht erreichen?

Maschinelles Lernen benötigt eine gute Datenbasis, eine Struktur und einen mathematischen Prozess, der sich selbst verbessert. Die damit erzielten Ergebnisse sind beeindruckend und erlauben viele mögliche Anwendungen. Dennoch sind sie noch weit entfernt von einer künstlichen menschenähnlichen Intelligenz.

- ▶ Durch maschinelles Lernen trainierte Systeme können nicht wie ein Mensch ihre Umgebung sehen, hören, riechen, schmecken, fühlen und die dabei gewonnenen Sinneseindrücke simultan, auch neue und unerwartete, unmittelbar verarbeiten und in Entscheidungsprozesse einbinden.
- ▶ Diese Systeme können daher in der Regel eine Aufgabe sehr gut bewältigen. Aber sie können keine neue, zweite Aufgabe lernen.
- ▶ Den Programmen fehlt die Ebene des Verständnisses vom Objekt, mit dem sie sich beschäftigen. Sie können daher auch keine Entscheidungen aufgrund rationellen Abwägens treffen.
- ▶ Einige Forscher [16] [20] sehen in diesen Fertigkeiten auch keine notwendige Voraussetzung für einen Einsatz von Programmen, die auf maschinellem Lernen beruhen. In ihren Augen lenkt die Verwendung des Begriffs „Künstliche Intelligenz“ ab von den Problemen, vor denen maschinelles Lernen nach wie vor steht (s.o.).
- ▶ Andere Forscher [21] [22] [23] dagegen beklagen die Begrenzung des Begriffs KI auf das maschinelle Lernen, weil und wenn dadurch auch eine Begrenzung auf die Probleme maschinellen Lernens erfolgen sollte – wie etwa das Gründen einer europäischen Forschungsinitiative oder eines –zentrums. Sie halten nach wie vor fest an dem Ziel, eine menschenähnliche Künstliche Intelligenz zu entwickeln.

Folgen des maschinellen Lernens – was wird sich ändern?

Maschinelles Lernen wurde bis jetzt vorwiegend von Unternehmen eingesetzt, die dem Umfeld des Internets oder der Logistik zuzurechnen sind – Amazon, Apple, Google, Facebook, Alibaba, Baidu oder Tencent. Analysten erwarten nun jedoch, dass auch andere Branchen auf maschinell angelegte Programme setzen werden – etwa Banken, Versicherungen. Experten erwarten, dass eine Reihe von Berufen durch Anwendungen von maschinell angelegten Geräten – seien es Chatbots, automatisch fahrende Busse, Bahnen oder Taxen, Roboter oder auch Bankangestellte – in den kommenden Jahrzehnten verschwinden



könnten [24]. Veränderungen ergeben sich dann sowohl in Branchen und Berufsbildern, als auch im gesellschaftlichen Zusammenleben:

- ▶ **Medizin:** Die Qualität der Diagnose oder Medikamentenwahl wird verbessert, wenn Systeme mit den Daten der besten Krankenhäuser und Studien arbeiten und so regelmäßig auf dem aktuellen Stand der Forschung und Richtlinien gehalten werden kann, wie es ein Mensch nur mit großem Aufwand – wenn überhaupt – erreichen könnte.
- ▶ **Industrie:**
 - Industrie 4.0: Die Verzahnung von Produktion, Handel und unter Umständen auch Wartung macht eine sehr effiziente und ressourcenschonende Produktion wie Logistik möglich [vgl. 25, S.107f.].
 - Roboter können sich dank maschinellem Lernen leichter für neue Aufgaben anlernen lassen. Auch sollen sie lernen, in einem Raum mit Menschen zu arbeiten, ohne diese durch ihre Bewegungen zu verletzen oder gar zu töten.
- ▶ **Verkehr/Mobilität:** Die Automatisierung des Fahrens kann Unfälle vermeiden und die Produktivität von Taxen wie Leihwagen erheblich erhöhen. Das Gleiche gilt für automatische Schiffe oder Flugzeuge. Allerdings müssen auch hier die Folgen bedacht werden, falls mehr Güter per LKW gefahren würden oder Menschen von Bussen und Bahnen auf einzeln fahrende PKW umstiegen, weil das Nachteile für den Klimaschutz nach sich ziehen kann.
- ▶ **Sicherheit:** Anfang 2018 veröffentlichten Forscher den „Malicious AI Report“, in dem sie darlegen, welche Gefahren ein Missbrauch heutiger, auf maschinellem Lernen beruhender KI-Technologien mit sich bringt: Sie beschreiben unter anderem hocheffiziente Hacker- und Drohnenangriffe [26]. Generell lassen maschinell trainierte Programme auch sehr effiziente Waffen möglich erscheinen, sowohl im Netz als auch in der realen Welt.
- ▶ Cédric Villani fordert in seinem Bericht für eine Europäische und Französische KI-Strategie konsequenterweise eine Konzentration der Europäischen Forschungsanstrengungen auf die Bereiche Gesundheit, Umwelt, Transport / Mobilität und Sicherheit / Verteidigung [27, S.8f.].
- ▶ **Gesellschaft**
 - ▶ **Transparenz:** Immer mehr Entscheidungen über Diagnosen, Kreditwürdigkeit, Einladungen zu Bewerbungsgesprächen, Strafmaß für Verbrecher, Beobachtung durch Polizei oder Geheimdienste werden KI-Systemen überlassen. Wie und nach welchen Kriterien diese entscheiden, kann daher für den Betroffenen undurchsichtig und gegebenenfalls überraschend sein. So nutzen Mastercard und Visa zum Beispiel Programme, die ungewöhnliche Kontenbewegungen auf Kreditkarten registrieren, und diese dann sperren. Das House of Lords in London fordert daher zum Beispiel, dass Entscheidungen von KI-Systemen gekennzeichnet werden sollten [18, S.6].
 - ▶ **Lebenslanges Lernen / Zukunft der Arbeit:** Experten erwarten, dass mit dem Wegfall alter Arbeitsplätze unter Umständen neue Jobs entstehen, zum Beispiel als Daten-Manager. Um diese Chancen zu nutzen, müssen allerdings die betroffenen Bürgerinnen und Bürger umgeschult oder weitergebildet werden [28, S.112]. Das wird ein Umdenken des Bildungssystems erfordern [18, S.6]. Es erscheint offen, ob Zahl und Qualität der neuen Jobs denen der fortgefallenen gleichkommt.
 - ▶ **Monopolbildung:** Tesla-Gründer Elon Musk warnte bereits 2017, dass ein Durchbruch im Bereich menschenähnlicher KI dem Besitzer der Technologie einen gefährlichen Machtvorteil verschaffen würde. Er wies darauf hin, dass ein großer Teil des Know-hows und der Daten zurzeit im Besitz von Firmen wie Alibaba, Alphabet, Apple, Amazon, Facebook, Google oder Tencent sind [29]. Diese haben zum Teil bereits heute faktisch ein Monopol auf ihren jeweiligen Gebieten. Das britische Oberhaus und Cédric Villani sehen die gleiche Gefahr und fordern daher konkrete Maßnahmen zum einen für eine effektive Datenkontrolle durch Bürgerinnen und Bürger, zum anderen einen Schutz gegen Monopole und freien Zugang zu Daten [18] [25].



- **Recht:** Offen ist, wie rechtlich mit Systemen umzugehen ist, die fehlerhafte Entscheidungen fällen und Schaden verursachen. Das britische Oberhaus zum Beispiel fordert den Gesetzgeber auf, Vorbereitungen zu treffen [18, S.30ff], ebenso bezeichneten Experten gegenüber dem Ausschuss Digitale Agenda des Bundestages Haftungs- sowie strafrechtliche Fragen als größte Herausforderung beim Einsatz von KI [25]. Cédric Villani sieht dagegen als grundsätzliches Problem, dass die Gesetzgebung der schnellen Entwicklung von Programmen nicht folgen kann. Er fordert daher, dass Programmierer und Forscher selbst strenge ethische Regeln an ihre Arbeit anlegen (vgl. die Diskussion um einen Hippokratischen Eid für Programmierer [30], oder die Forderung, die Entwicklung von KI streng zu kontrollieren [26]) und eine regelmäßige politische, vor allem aber informierte öffentliche Diskussion über Möglichkeiten und Grenzen der Künstlichen Intelligenz zu führen [25 S.113ff.].

Literaturstellen, die zitiert wurden

- [1] Domingos P (2015): *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*.
- [2] Banino A et al. (2018): *Vector-based navigation using grid-like representations in artificial agents*. Nature Letter, DOI:10.1038/s41586-018-0102-6.
- [3] Rojas R (2017): *Beantwortung der Fragen gem. Fragenkatalog für das Fachgespräch zum Thema „Künstliche Intelligenz“ des Ausschusses Digitale Agenda am 22.03. 2017*. Deutscher Bundestag.
- [4] Michi D (1961): *Experiments on the mechanization of game-learning Part I. Characterization of the model and its parameters*. The Computer Journal 6 (3), 232-236. DOI:10.1093/comjnl/6.3.232.
- [5] Brooks, R (2017): *Machine Learning Explained*. (Blogeintrag in der Serie „Future of Robotics and Artificial Intelligence“).
- [6] Jordan M et al. (2015): *Machine learning: Trends, perspectives, and prospects*. Science 349 (6245), 255-260. DOI:10.1126/science.aaa8415.
- [7] Prieto A et al. (2016): *Neural Networks: An overview of early research, current frameworks and new challenges*. Neurocomputing 214, 242-268. DOI:10.1016/j.neucom.2016.06.014.
- [8] Silver D et al. (2017): *Mastering the game of Go without human knowledge*. Nature 550, 354-359. DOI:10.1038/nature24270.
- [9] Bengio Y (2009): *Learning Deep Architectures for AI*. Foundations and Trends in Machine Learning 2 (1), 1-127. DOI:10.1561/2200000006.
- [10] Pan Y (2016): *Heading toward Artificial Intelligence 2.0*. Engineering 2 (4), 409-413. DOI: 10.1016/J.ENG.2016.04.018.
- [11] Schmidhuber J (2014): *Deep learning in neural networks: An overview*. Neural Networks 61, 85-117. DOI: doi.org/10.1016/j.neunet.2014.09.003.
- [12] Aunkofer B (2017): *maschinelles Lernen mit Entscheidungsbaumverfahren*. Data Science Blog.
- [13] Burgess M (2018): *UK police are using AI to inform custodial decisions – but it could be discriminating against the poor*. Wired.
- [14] Liu W et al. (2016): *A survey of deep neural network architectures and their applications*. Neurocomputing 234, 11-26. DOI: 10.1016/j.neucom.2016.12.038.
- [16] Jordan, M (2018): *Artificial Intelligence – The Revolution Hasn't Happened Yet*. Medium. Vgl. Ders.(2018): *Perspectives and Challenges*. Presentation SysML 2018.
- [17] Williams, E. (2018): *Self-Driven: Uber and Tesla*. Hackaday, 02.04.2018.



- [18] Select Committee on Artificial Intelligence (2018): [AI in the UK: ready, willing and able?](#) Report of Session 2017-2019.
- [19] Docket, T (2018): [AI for Long-Term Autonomy](#). IEEE RA Special Issue October 2018.
- [20] Bach F et al. (2018), [Initiative to establish a European Lab for Learning & Intelligent Systems](#). Open Letter, Online.
- [21] Rodney, B (2018): [The Origins of „Artificial Intelligence“](#). (Blogeintrag in der Serie „Future of Robotics and Artificial Intelligence).
- [22] Armbruster, A (2018): [„Viele europäische Ideen haben amerikanische Firmen umgesetzt“](#). Interview mit Wolfgang Wahlster, DFKI. FAZ, 25.04.2018.
- [23] Armbruster, A (2018): [„Google verfügt über tolle Daten, die wir nicht haben.“](#) FAZ, 02.05.2018.
- [24] Löhr, J (2018): [Digitalisierung zerstört 3,4 Millionen Stellen](#). FAZ, 02.02. 2018.
- [25] Bundestag (2018): [Bundestagsausschuss Digitale Agenda, Anhörung Künstliche Intelligenz, 22. März 2017](#).
- [26] Brundage M et al. (2018): [The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation](#).
- [27] Villani, C (2018): [For a meaningful artificial intelligence. Towards a French and European strategy](#).
- [28] Bundesministerium für Arbeit und Soziales (2017), [Weißbuch Arbeiten 4.0](#).
- [29] Sulleyman A (2017): [AI is highly likely to destroy humans, Elon Musk warns](#). Independent.
- [30] Upchurch, T (2018): [To work for society, data scientists need a Hippocratic oath with teeth](#). Wired 08.04.2018

Weitere Recherchequellen

Eine visuelle Erklärung maschinellen Lernens [anhand von Entscheidungsbäumen](#).

Video-Reihe: Erklärung Neuronaler Netze. [Teil 1](#) und [Teil 2](#).

Menace: [the pile of matchboxes which can learn](#). Video.



fact sheet

Ansprechpartner in der Redaktion

Sophie Rotgeri

Volontärin

Sönke Gäthke

Redakteur

Telefon +49 221 8888 25-0

E-Mail redaktion@sciencemediacenter.de

Disclaimer

Dieses Fact Sheet wird herausgegeben vom Science Media Center Germany. Es bietet Hintergrund-informationen zu wissenschaftlichen Themen, die in den Schlagzeilen deutschsprachiger Medien sind, und soll Journalisten als Recherchehilfe dienen.

SMC-Fact Sheets verstehen sich nicht als letztes Wort zu einem Thema, sondern als eine Zusammenfassung des aktuell verfügbaren Wissens und als ein Hinweis auf Quellen und weiterführende Informationen.

Dieses Fact Sheet wurde von Experten aus der Wissenschaft auf Korrektheit geprüft.

Sie haben Fragen zu diesem Fact Sheet (z. B. nach Primärquellen für einzelne Informationen) oder wünschen Informationen zu anderen Angeboten des Science Media Center Germany? Dann schicken Sie uns gerne eine E-Mail an redaktion@sciencemediacenter.de oder rufen Sie uns an unter +49 221 8888 25-0.

Impressum

Die Science Media Center Germany gGmbH (SMC) liefert Journalisten schnellen Zugang zu Stellungnahmen und Bewertungen von Experten aus der Wissenschaft – vor allem dann, wenn neuartige, ambivalente oder umstrittene Erkenntnisse aus der Wissenschaft Schlagzeilen machen oder wissenschaftliches Wissen helfen kann, aktuelle Ereignisse einzuordnen. Die Gründung geht auf eine Initiative der Wissenschafts-Pressekonferenz e.V. zurück und wurde möglich durch eine Förderzusage der Klaus Tschira Stiftung gGmbH.

Nähere Informationen: www.sciencemediacenter.de

Diensteanbieter im Sinne RStV/TMG

Science Media Center Germany gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg
Amtsgericht Mannheim
HRB 335493

Redaktionssitz

Science Media Center Germany gGmbH
Rosenstr. 42-44
50678 Köln

Vertretungsberechtigte Geschäftsführer

Beate Spiegel, Volker Stollorz

Verantwortlich für das redaktionelle Angebot (Webmaster) im Sinne des §55 Abs.2 RStV

Volker Stollorz

