Stefan Hügel

Künstliche Intelligenz und Politik*

Algorithmen, Data Science, Microtargeting – und ihre Auswirkung auf politische Entscheidungen

Während in der Wissenschaft die Erwartungen an eine Künstliche Intelligenz (KI) zuletzt deutlich herunter geschraubt wurden, wird die Technologie in der politischen Öffentlichkeit in den letzten Jahren deutlich aufmerksamer – und zum Teil kritischer – verfolgt. Stefan Hügel sortiert diese Debatte, indem er zunächst einmal die verschiedenen Begriffe von starker und schwacher KI, computergestützter Algorithmen, Lern- und Entscheidungssystemen voneinander abgrenzt. Im zweiten Teil seines Beitrags geht er ausführlich auf die verschiedenen Gefährdungspotenziale der KI auf Demokratie und Öffentlichkeit ein, bevor er abschließend mögliche Schutzkonzepte vorstellt.

> Ich bin der Ansicht, dass ein in jeder Beziehung zu vereinfachter Begriff von Intelligenz sowohl das wissenschaftliche wie das außerwissenschaftliche Denken beherrscht hat, und dass dieser Begriff zum Teil dafür verantwortlich ist, dass es der perversen, grandiosen Phantasie der künstlichen Intelligenz ermöglicht wurde, sich derart zu entfalten.
> *Joseph Weizenbaum*[1]

> Der Computer Nr. 3 / sucht für mich den richtigen Boy,
> und die Liebe ist garantiert / für beide dabei.
> Der Computer weiß genau / für jeden Mann die richtige Frau,
> und das Glück fällt im Augenblick / aus seiner Kartei.
> *France Gall*[2]

Nach einer wechselvollen Geschichte seit den hohen Erwartungen der 1950er und 1960er Jahre nimmt das Interesse an *Künstlicher Intelligenz* wieder deutlich zu. Maschinelles Lernen und algorithmische Entscheidungsfindung stellen die Automatisierung auf eine neue Stufe.[3] Der KI-Experte Wolfgang Wahlster übersetzt KI inzwischen mit „Künftige Informatik".[4] Es ist zu erwarten, dass Anwendungen der Künstlichen Intelligenz zunehmend unser Leben bestimmen und prägen.

* Der Beitrag basiert in Teilen auf Hügel S, Rehak R (2019) Künstliche Intelligenz im Wahlkampf. Grundrechte-Report 2019. Frankfurt/M.: Fischer-Verlag (im Erscheinen). Dank an Britta Schinzel, Hans-Jörg Kreowski, Sven Lüders und Rainer Rehak für wertvolle Hinweise.

Gleichzeitig verlagert sich die Kommunikation im Internet seit Jahren zunehmend auf Soziale Netzwerke, die von globalen Wirtschaftsunternehmen wie Facebook oder Twitter betrieben werden. Diese Unternehmen stellen einen großen Teil ihrer Dienste für die Nutzerinnen und Nutzer kostenlos zur Verfügung – ihr Geschäftsmodell besteht darin, die bei der Nutzung anfallenden Daten auszuwerten und Erkenntnisse daraus insbesondere der Werbewirtschaft kostenpflichtig anzubieten. Den Unternehmen, die diese Dienste nutzen, bietet sich damit die Möglichkeit, potenzielle KundInnen gezielt anzusprechen, Streuverluste zu vermeiden und Botschaften zielgruppengerecht zuzuschneiden, beispielsweise an finanzkräftige weiße Männer mittleren Alters oder aber an Jugendliche mit wenig Selbstvertrauen und emotionalen Problemen.[5] Dies wird als *Microtargeting* bezeichnet und nutzt ebenfalls Techniken der Künstlichen Intelligenz.

Angesichts einer öffentlichen politischen Debatte, die zunehmend den Regeln der Werbewirtschaft folgt, ist die Ausdehnung von Methoden des Microtargeting und der Künstlichen Intelligenz auf den Wahlkampf nicht überraschend. Die Begriffe *Algorithmus* und *Künstliche Intelligenz* werden dabei in der öffentlichen Diskussion häufig undifferenziert verwendet. Dieser Beitrag will zunächst die Begriffe aus technischer Sicht erläutern und eingrenzen. Danach werden die Risiken für die politische Kommunikation beschrieben. Zuletzt gehe ich auf Lösungsmöglichkeiten ein – diese bestehen vor allem in einem konsequent zu Ende gedachten und wohlverstandenen Datenschutz.

Algorithmen

Ein in der aktuellen Debatte häufig verwendeter Begriff ist der des *Algorithmus*[6]. Viele Eigenschaften einer autonomen, automatisierten Datenverarbeitung werden heute Algorithmen zugeschrieben. Ein „Algorithmen-TÜV“ wird gefordert, um die Risiken algorithmischer Entscheidungsfindung in den Griff zu bekommen.

Was ist ein Algorithmus? Entgegen dem in der öffentlichen Debatte häufig suggerierten, inhärenten Bezug zur Künstlichen Intelligenz verstehen wir darunter einfach eine eindeutige Handlungsvorschrift, um ein Problem oder eine Klasse von Problemen unter variablen Bedingungen zu lösen. Ein Algorithmus setzt sich aus endlich vielen[7], wohldefinierten Einzelschritten zusammen. Algorithmen können i. d. R. zur Ausführung in einem Computerprogramm implementiert werden, aber auch eine natürlichsprachliche Handlungsanweisung, die von Menschen ausgeführt werden kann, ist ein Algorithmus.[8]

Freilich werden auch in der Künstlichen Intelligenz Algorithmen verwendet. Darüber hinaus findet aber heute in der öffentlichen Diskussion eine Bedeutungsverschiebung statt: Entscheidungssysteme, Lernprogramme und sonstige Programme der Künstlichen Intelligenz werden zunehmend exklusiv als „Algorithmen“ bezeichnet.[9] Dies ist kein angemessener Begriff eines Algorithmus. Im strengen Sinn des mathematischen Konstrukts sind es keine Algorithmen mehr, sondern der zugrundeliegende

(Lern-)Algorithmus ist durch das in der Lernphase gebildete Modell mit seinen Parametern „verschmutzt".[10] Hier sollte besser von *Software* gesprochen werden.

Lev Manovich hält die Bezeichnung von Konzepten der automatisierten Entscheidungsfindung als *Algorithmen* für irreführend, da sie meistens auf maschinellem Lernen (s. u.) beruhen. Im Gegensatz zu Algorithmen sind die Schritte des maschinellen Lernens hin zum letztlich erreichten Ergebnis häufig nicht nachvollziehbar. Faktisch stehe man bei solchen Systemen oft vor einer *Black Box*, die zwar praktische Ergebnisse liefere, aber nicht interpretierbar sei. Manovich bevorzugt ebenfalls den allgemeineren Begriff *Software*.[11]

Künstliche Intelligenz

Der Begriff *Künstliche Intelligenz* ist schwierig abzugrenzen, da es bisher noch nicht einmal eine allgemein akzeptierte Definition von Intelligenz gibt. Auch der gängige Definitionsversuch:

> „Ziel der KI ist es, Maschinen zu entwickeln, die sich verhalten, als verfügten sie über Intelligenz."[12]

bleibt letztlich unbefriedigend. Er knüpft an ein klassisches Gedankenexperiment an, mit dem überprüft werden soll, ob eine Maschine Intelligenz besitzt. Bei dem 1950 von Alan Turing vorgeschlagenen Turing-Test[13] stellt ein Mensch Fragen und soll anhand der Antworten herausfinden, ob ihm gerade ein Mensch oder eine Maschine antwortet. (Die Fragen werden schriftlich beantwortet, um die Erkennung anhand von hier nicht relevanten Eigenschaften, z. B. der Stimme, auszuschließen.) Die Maschine hat den Test bestanden, wenn sie anhand der Antworten nicht vom Menschen unterschieden werden kann.

Gemeinhin unterscheiden wir zwischen *starker* und *schwacher* KI:[14]

- Von *starker KI* sprechen wir bei einer Maschine, die über eine dem Menschen analoge Intelligenz verfügt, d. h. die in der Lage ist, das menschliche Denken zu mechanisieren. Dies kann eine Form des „maschinellen Bewusstseins" einschließen. Die Auseinandersetzung darüber, ob eine starke KI möglich ist, wurde mit Searle's *Chinesischem Zimmer*[15] angestoßen. Bei diesem Gedankenexperiment sitzt ein Mensch, der kein Wort chinesisch versteht, in einem Zimmer, in das ihm auf chinesisch formulierte Fragen gestellt werden. Er hat lediglich syntaktische Regeln, mit deren Hilfe er die Fragen beantworten kann – auch ohne ihren Inhalt zu verstehen. Das entspricht der Arbeitsweise eines Computers, der „Verständnis" lediglich simuliert. Dieses Experiment wirft die Frage auf, ob formal-syntaktische Übersetzungen auf dem Wege eines Sprungs von Quantität zu Qualität auch Bedeutung transportieren können.[16] Die Forschung ist heute weit davon entfernt, starke KI zu realisieren; ihre prinzipielle Umsetzbarkeit ist ein philosophisch umstrittenes Problem.

- Unter *schwacher* KI werden Techniken verstanden, die einzelne kognitive Fähigkeiten – i. d. R. innerhalb eines engen Aufgabenbereichs – in einem Computersystem nachbilden. Diese Fähigkeiten stehen in keiner Beziehung zu anderen Fähigkeiten – im Gegensatz zum Menschen, bei dem sie im Zusammenhang stehen. (Ein Computer, der programmiert wurde, Hunde und Katzen zu klassifizieren, ist deswegen weder in der Lage, einen Hund oder eine Katze zu füttern noch kann er Automobile von Fahrrädern unterscheiden.) Solche Systeme sind in der Lage unterschiedliche, aber jeweils sehr spezifische Aufgaben auszuführen.

In der informationstechnischen Praxis und der politischen Debatte sind Verfahren der schwachen KI relevant. Es werden Techniken eingesetzt, durch die abgegrenzte Aufgaben mit Hilfe von Computersystemen gelöst werden können. Die Wahrnehmung, wann es sich bei einem Computersystem um Künstliche Intelligenz handelt, verändert sich dabei mit der Zeit. Beispielsweise werden Schachprogramme – die auf der heutigen Entwicklungsstufe menschlichen SchachspielerInnen überlegen sind – durch ihre „Normalität" heute im Gegensatz zu früher im Alltagsverständnis häufig nicht mehr als Künstliche Intelligenz angesehen.[17] Auch die eingangs zitierte Unterstützung bei der Partnersuche würde man im Zeitalter von Dating-Apps wie *Tinder* wohl nicht mehr als Künstliche Intelligenz bezeichnen. Häufig werden in der Alltagssprache verwendete Begriffe wie „Lernen", „Erkennen", „Verstehen" unreflektiert verwendet und dadurch Erwartungen an Systeme der Künstlichen Intelligenz genährt, die sie nicht erfüllen können.[18] Dazu stoßen solche Systeme – wie jedes Computersystem – an Grenzen der Komplexität und der Berechenbarkeit.[19]

Zu den in der Praxis gebräuchlichen KI-Verfahren gehören Verfahren des sogenannten *Maschinellen Lernens*. Maschinelles Lernen verwendet oft große Datenmengen („*Big Data*"), was ein neues wissenschaftliches Gebiet, *Data Science*, hervorgebracht hat. Die Leistungsfähigkeit solcher Systeme ergibt sich aus der Leistungsexplosion der Rechnersysteme in den letzten Jahren. Sie macht es heute möglich, umfangreiche Datenmengen in kürzester Zeit zu durchsuchen und zu analysieren.

Im Grundsatz geht es beim maschinellen Lernen darum, bereitgestellte Daten – oft auch in einem kontinuierlichen Strom – aufgrund ihrer Eigenschaften in unterschiedliche Klassen einzuordnen, zu annotieren und dabei gleichzeitig die Klassen so anzupassen, dass eine geeignete Klassifizierung für das zu lösende Problem oder die Anfrage entsteht.[20] Man spricht auch von algorithmischer Entscheidungsfindung (*Algorithmic Decision Making*, ADM).[21]

Ein einfaches Beispiel wäre ein Klassifikator, der Bilder von Hunden und Katzen unterscheidet. Dazu würde man Eigenschaften suchen, die bei Hunden und Katzen unterschiedlich sind, eine Klassifizierung anhand dieser Eigenschaften vornehmen und die beiden Klassen mit den Namen „Hund" und „Katze" etikettieren. Da diese Eigenschaften nicht immer eindeutig zu bestimmen sind, werden sich dabei Fehler ergeben. Um diese Fehler zu beheben, passt man die Schwellwerte der Unterscheidung in mehreren Durchläufen immer wieder an, so dass sich bei jeder Iteration eine bessere Klassifizierung ergibt. Ziel ist, dass das System am Ende die Unterscheidung mit hinreichender Genauigkeit vornehmen kann: Es hat „gelernt".

Analog würde man z. B. vorgehen, wenn man Menschen nach Parteipräferenz klassifizieren möchte: Auch hier legt man Eigenschaften fest, die für Anhänger*innen einer bestimmten Partei typisch sind. Nach mehreren Iterationen wird man dann Einzelpersonen mit hoher Wahrscheinlichkeit einer Partei zuordnen können.

Der eigentliche Algorithmus ist dabei die Berechnungsvorschrift, die das Lernen steuert. Dieser wird i. d. R. vom Anwendungsbereich unabhängig sein. Er „lernt" aus den vorhandenen Daten und – und das ist Teil des Problems – reproduziert dabei die Strukturen, d. h. die Einschätzungen und Vorurteile, die in diesen Daten enthalten sind.[22]

Grundsätzlich können drei Arten von Lernverfahren unterschieden werden:[23]

- *Unüberwachtes Lernen:* Das System „lernt" ohne Vorgaben und findet dabei Strukturen und Muster in den bereitgestellten Daten. Diese Muster müssen mit der inhaltlichen Frage nicht in Zusammenhang stehen.[24] Das Modell wird nach der Lernphase nicht mehr verändert.
- *Überwachtes Lernen:* Dem System werden Muster für erwartete Ergebnisse vorgegeben und während des Lernens überprüft, mit welcher Wahrscheinlichkeit eine Voraussage aus den Daten mit dem erwarteten Ergebnis übereinstimmt. Das Modell wird nach der Lernphase ebenfalls nicht mehr verändert.
- *Bestärkendes Lernen:* Das Modell wird während der Nutzung durch neue Daten laufend verbessert, indem die Schwellwerte der Klassifikatoren basierend auf den bisher erreichten Voraussageregeln in Feedbackschleifen angepasst werden.

Probleme durch statistische Lernverfahren

Die Ergebnisse maschinellen Lernens lösen in der Öffentlichkeit gelegentlich Unverständnis oder Empörung[25] aus. Sie fußen auf grundsätzlichen Einschränkungen der Modellbildung und des maschinellen Lernens, das auf Modellen der Wirklichkeit aufbaut. Solche Modelle sind immer unvollständig und können lediglich einen Ausschnitt der wirklichen Welt abbilden. Die Einschränkungen der Modelle wirken sich dann auf die Ergebnisse der „erlernten" Voraussagen aus.

- Maschinelles Lernen baut auf den vorgefundenen Daten auf. Die Voraussagen eines lernenden Systems sind dabei höchstens so gut wie die Daten, aus denen es gelernt hat. Sind in den Daten z. B. Vorurteile gegen Minderheiten eingeschrieben, so werden auch die Voraussagen entsprechend ausfallen.
- Sieht das Lernverfahren kein Feedback vor – wie es beim *unüberwachten* Lernen der Fall ist – führt das zusätzlich dazu, dass Fehler nicht korrigiert werden; auch dann, wenn falsche Klassifizierungen (für den Menschen) offensichtlich sind.[26]

- Doch auch Feedbackschleifen können weitere Fehler nach sich ziehen. Führen beispielsweise Anfragen an eine Suchmaschine dazu, dass im Eingabefeld entsprechende Ergänzungen vorgenommen werden, wird dies wiederum die Wahrscheinlichkeit erhöhen, dass diese Anfragen geklickt werden – und so weiter.[27]

Solche Probleme führen im schlimmsten Fall zu *Programmiertem Rassismus.*[28] Die Klassifizierung einer afroamerikanischen Frau als Gorilla ist nur ein Beispiel aus der praktischen Anwendung solcher Systeme, das verständliche Empörung hervorruft.[29] Besonders kritisch ist es, wenn auf Basis derart trainierter Systeme Entscheidungen getroffen werden, die einen massiven Einfluss auf das Leben von Menschen haben, beispielsweise bei der maschinell unterstützten vorausschauenden Polizeiarbeit (*Predictive Policing*).[30]

Dabei können rassistische Effekte an Stellen auftauchen, an denen sie zunächst niemand erwartet hatte – einfach, weil die Lernalgorithmen Korrelationen aufweisen, die (noch) keine kausale Begründung liefern.[31] Der zitierte Beitrag verweist auf den Fall, dass eine Software in den USA zur Auswahl von BewerberInnen den Zusammenhang zwischen Entfernung vom Wohnort zum Arbeitsplatz und der Fluktuation erkannte und empfahl, MitarbeiterInnen einzustellen, die nahe am Arbeitsplatz wohnen. Dies führt zur Benachteiligung von AfroamerikanerInnen, die häufiger als Weiße in Außenbezirken leben. Letztlich reproduziert auch dies bereits bestehende Benachteiligungen.

Algorithmenethik

Es ergibt sich die Frage nach einem verantwortungsvollen Einsatz solcher Software bzw. nach einem vielfach geforderten *Algorithmen-TÜV*.[32] Zweifellos ist es sinnvoll, die für möglicherweise existenzielle Entscheidungen verwendeten Algorithmen einer Prüfung zu unterziehen. Dies könnte durch Behörden erfolgen, die den Datenschutzbehörden vergleichbar sind. Solange solche Behörden, die multinationalen Konzernen mit erheblichen Ressourcen gegenüberstehen, nur unzureichend finanziell und personell ausgestattet sind, keimt hier aber wenig Hoffnung auf eine effektive Kontrolle auf.

Algorithmen können keine Eigenschaften wie objektiv, intelligent, gut oder böse zugeschrieben werden. Sie sind korrekt, effizient, sparsam, schnell, adaptiv.[33] Das sind Eigenschaften, die ein *Algorithmen-TÜV* überprüfen könnte. Ein Sortieralgorithmus kann anhand geeigneter Kriterien sowohl für die Sortierung von Brötchen als auch für die Sortierung von Menschen eingesetzt werden. Offensichtlich ist nicht der Algorithmus selbst das Problem, sondern sein Verwendungskontext. Eine Kontrolle, die sich auf den Algorithmus allein konzentriert, geht damit am Problem vorbei.

Es gibt noch einen weiteren Grund, warum ein Kontrolle der Algorithmen ins Leere läuft. Wie oben gesehen, basieren die Lernergebnisse von (Lern-) Algorithmen auf den vorgegebenen Modellen und bereitgestellten Daten, mit denen sie „trainiert“ werden

und deren eingeschriebene Vorurteile sie letztlich reproduzieren. Andererseits ist auch eine Anpassung der Daten und Modelle problematisch: Was sind denn „faire" Daten? Und wer kann darüber bestimmen? Es müsste ein gesellschaftlicher Konsens darüber gebildet werden, welche Daten und Modelle als Basis des Lernens dieser Systeme angewendet werden sollen – angesichts der politischen Verhältnisse mit unterschiedlichen Wertvorstellungen und Präferenzen dürfte dies eine große Herausforderung sein.

Ein anderer Ansatz sind Ethische Leitlinien für die Gestalter*innen und Entwickler* innen von KI-Systemen, wie es sie für die Informatik als Disziplin bereits seit langem gibt.[34] Die Europäische Kommission hat Ethische Leitlinien für Künstliche Intelligenz vorgeschlagen.[35] Solche Ansätze bleiben aber zahnlos, wenn Verstöße dagegen nicht sanktioniert werden.

Stets sind es Menschen und Organisationen[36], die über den Einsatz von solchen Algorithmen entscheiden. *Algorithmen* kann keine eigene Entscheidungs- oder Handlungsmacht zugebilligt und damit auch keine Maschinenethik von ihnen eingefordert werden.

Folgen für die Politik

Die zunehmende Anwendung von Methoden der Künstlichen Intelligenz hat Folgen für die Politik. Eine Studie der Stiftung neue Verantwortung nennt – mit Blick auf die Außenpolitik, aber in Teilen auch übertragbar auf innenpolitische Fragen – drei Bereiche, in denen sich politische Herausforderungen ergeben:[37]

- Wirtschaftliche Disruption und wirtschaftliche Chancen,
- Sicherheit und autonome Waffensysteme,
- Demokratie und Ethik.

Die nächsten Abschnitte reißen die beiden ersten Punkte kurz an, legen aber dann einen besonderen Schwerpunkt auf Methoden des Microtargeting als Frage der Demokratie und politischen Ethik.

Wirtschaft

Der Treiber für KI-Technologien sei primär ökonomisch, stellt die Stiftung neue Verantwortung in ihrer Studie[38] fest. Es wird erwartet, dass beispielsweise autonomes Fahren die Automobilindustrie erheblich verändern wird. Außenpolitisch werden Regierungen bestrebt sein, die Interessen der eigenen Wirtschaft im globalen Markt zu wahren. Erforderlich ist aus Sicht der Stiftung aber vor allem ein Risikomanagement mit Blick auf die Konzentration wirtschaftlicher Macht und Disruptionen des Arbeitsmarkts.

Doch auch für Verbraucher entstehen durch KI-Technologien neue Risiken, beispielsweise beim Svcoring von Verbraucher*innen. Eine Studie der Gesellschaft für Informatik[39] nennt als zentrale Risiken das Kreditscoring – die Bewertung der Kreditwürdigkeit auf Basis einer statistischen Analyse – und die Preisdifferenzierung – unterschiedliche Preise für die gleiche Leistung, wobei nach zeitlichen, räumlichen, persönlichen oder sachlichen Kriterien differenziert werden kann. Als Gefährdungsszenarien werden genannt:

- *Inhaltlich unrichtiger Algorithmus:* Falsche Identifikation und Bewertung beim ADM und fehlerhafte Scores
- *Diskriminierender Algorithmus:* Anknüpfen des ADM an Merkmale, durch die sich eine direkte oder indirekte Diskriminierung ergeben kann
- *Intransparent personalisierender Algorithmus:* Diskriminierende Elemente bei den verwendeten Entscheidungskriterien, wenn z. B. die Preisfestsetzung nicht durch sachliche Kriterien gedeckt ist.

Auch besteht die Gefahr eines programmierten Rassismus, selbst bei vordergründig korrekt arbeitender Klassifikation.

Sicherheit und autonome Waffensysteme

Der russische Präsident Wladimir Putin erklärte kürzlich: Wer KI beherrsche, der beherrsche die Welt. Die frühere Staatssekretärin im Bundesministerium der Verteidigung, Katrin Suder, ist ähnlicher Ansicht: „KI hat definitiv das Potenzial, die gesamte Dynamik im Cyberraum zu verändern. Es handelt sich um eine Fähigkeit, die Wirkungsüberlegenheit herstellen kann. Damit geht es um den Kern von Sicherheit.“[40]

Im militärischen Bereich spielen Roboter und unbemannte Flugsysteme (*Unmanned Aerial Vehicles*, UAV), umgangssprachlich „Drohnen“ genannt, eine immer größere Rolle.[41] Das US-amerikanische *Department of Defense* versteht unter einem autonomen Waffensystem:

> „A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.“[42]

Noch setzt kein Staat vollautonome Waffen gegen Menschen ein. Sie sind bisher nicht in der Lage, zentrale Forderungen des Völkerrechts umzusetzen, wie z. B. die Unterscheidung zwischen Kombattant*innen und Zivilpersonen.[43]

Demokratie

Bei allen lern- und erkenntnistheoretischen Unzulänglichkeiten erlauben es die Ergebnisse des maschinellen Lernens, die Öffentlichkeit gezielt zu beeinflussen. Thomas Ramge[44] weist auf drei Hauptgefahren einer Anwendung schwacher KI für Politik und Gesellschaft hin:

- *Monopolisierung von Daten:* Netzwerkeffekte erhöhen den Nutzen digitaler Dienste – je mehr Menschen sich an dem Dienst beteiligen, desto größer ist ihr Nutzen für alle. Dieser Effekt wird verstärkt, wenn lernende Maschinen mit Feedbackschleifen für immer stärkeres Wachstum sorgen. Je öfter Dienste genutzt werden und je mehr Marktanteile sie erobern, desto schwerer wird ihr Vorsprung aufzuholen sein – das führt zu Monopolen.[45]
- *Manipulation des Einzelnen:* Monopole begünstigen Manipulation, da es keine Alternativen mehr gibt, auf die man ausweichen kann. Dies wird verstärkt, wenn wir immer mehr Aufgaben des täglichen Lebens an virtuelle, autonom lernende Assistenten delegieren. Dabei verlassen wir uns darauf, dass diese Agenten in unserem Sinne „entscheiden" und „handeln". Risiken entstehen dadurch, dass wir die hinter ihnen stehende Funktionslogik nicht kennen und auch nicht verstehen. Entscheidend ist die Frage, von und für wen ein Agent programmiert ist, in wessen Interesse er agiert. Welche Interessen nimmt ein virtueller Agent wahr, der von einem Unternehmen bereitgestellt wird, das uns Waren oder Dienstleistungen verkaufen will – oder das politische Programm einer bestimmten Partei favorisiert? Es müssen Wege gefunden werden, wie manipulierende oder betrügerische Agenten verhindert werden können.[46] Ein erster Schritt dahin wäre Transparenz – doch wie bereits oben dargelegt („Algorithmen-TÜV"), reicht das nicht aus.
- *Missbrauch durch Regierungen:* Software, die auf Künstlicher Intelligenz basiert, kann auch von staatlicher Seite für die Manipulation, Überwachung und Unterdrückung der Bevölkerung eingesetzt werden. Bereits heute erlaubt die großflächige Videoüberwachung öffentlicher Bereiche, kombiniert mit automatischer Gesichtserkennung und dem Abgleich der Daten mit biometrischen Datenbanken eine umfassende Überwachung großer Menschenmengen. Auch wenn die derzeitigen Ergebnisse entsprechender Großversuche noch nicht perfekt sind: Durch maschinelles Lernen werden diese Systeme weiter perfektioniert – das daraus resultierende Überwachungsnetz wird immer engmaschiger.

Politische Kommunikation – Microtargeting und Nudging

Der erste Schritt des *Microtargeting* in der politischen Kommunikation ist die Klassifizierung der Wähler*innen nach politischen Präferenzen, um sie dann gezielt anzusprechen. Bereits im Wahlkampf des früheren US-Präsidenten Barack Obama nutzte dessen Wahlkampfteam massiv die Möglichkeiten des Internet und der sozialen Medi-

en.[47] Durch den technischen Fortschritt der letzten Jahre wurden die Möglichkeiten der Personalisierung und damit der gezielten Ansprache massiv verbessert. Es war also zu erwarten, dass dies auch in Wahlkämpfen angewandt und die Wahlkampftechniken entsprechend angepasst und zunehmend verfeinert werden würden.

Ein verwandter Ansatz der Manipulation – politisch ebenso wie im Alltag – ist *Nudging* (engl. *nudge* = Schubs). Dabei handelt es sich um einen Ansatz aus der Verhaltensökonomik, bei dem menschliches Verhalten ohne Gebote, Verbote oder ökonomische Anreize beeinflusst wird.[48] Ein „klassisches" Beispiel – noch ganz ohne jede Digitalisierung – ist die in manchen Urinalen abgebildete Fliege (bzw. das kleine Modell eines Fußballtors mit Ball), die Männer animieren soll, genauer zu „zielen" – und so zur Senkung der Reinigungskosten beiträgt.[49]

Das hat mit KI zunächst nichts zu tun. Doch *Nudging* kann auch zur politischen Beeinflussung eingesetzt werden. Das durch fortgeschrittene Methoden der Datenanalyse erworbene Wissen über Gewohnheiten, Vorlieben und Abneigungen kann mittels Microtargeting gezielt für Maßnahmen der politischen Manipulation genutzt werden, ebenso wie für die Manipulation von Verbraucher*innen; beispielsweise wenn sie animiert werden sollen, sich gesund zu ernähren. Mit derartigen Anstößen wird mittlerweile in der Sozialversicherung, bei Investmententscheidungen oder für den Umweltschutz gearbeitet.[50] Auch wenn es sich „nur" um statistische Werte und Wahrscheinlichkeiten handelt, ergeben sich daraus wertvolle Hinweise, auf welche Anreize bestimmte Gruppen von Menschen wie reagieren.

Im Gegensatz zu einer dialogorientierten politischen Kommunikation, die um das bessere Argument und die bessere Lösung ringt, werden im Wahlkampf durch *Microtargeting* gezielt Werbebotschaften platziert, die – auch manipulativ – Wähler*innen zur Stimmabgabe für die eigene Partei bewegen oder die Stimmabgabe für Kontrahent*innen verhindern sollen, indem sie beispielsweise potenzielle Wähler*innen der Gegenseite von der Stimmabgabe abzubringen versuchen (Demobilisierung). *Microtargeting* nutzt dabei die Differenzierungsmöglichkeiten, die sich aus den individuellen Informationen über Einzelpersonen ergeben. Aspekte, die mit den Zielen der jeweils umworbenen Person in Einklang stehen, werden betont, Aspekte, die ihren Zielen widersprechen, gezielt ausgeblendet. Anstatt umfassender Kenntnis der Ziele einer Partei können an unterschiedliche Personen so völlig unterschiedliche Bilder vermittelt werden; jeweils so, wie es den persönlichen Vorlieben entspricht.

In welchem Umfang derartige Manipulation stattfindet, hängt nicht zuletzt auch von der Finanzkraft der beteiligten Interessengruppen ab. Wie alle Möglichkeiten der Manipulation durch Werbung steht auch das *Microtargeting* vor allem ressourcenstarken Akteuren zur Verfügung.

Unternehmen wie Google, Facebook oder Twitter sammeln diese Daten, werten sie aus und bekommen dadurch die Möglichkeit, Menschen nach Vorlieben beliebig zu klassifizieren – abhängig davon, welches Erkenntnisinteresse ihre jeweilige Kundschaft hat. Dies kann politisch, weltanschaulich, wirtschaftlich oder auch anderweitig motiviert sein. Es ist kein „Versehen", das durch andere Regeln leicht aus der Welt geschafft werden kann: Es ist das Geschäftsmodell, auf dem die durch diese Unternehmen gestaltete Internet-Wirtschaft basiert.

Skandal! Facebook und Cambridge Analytica

Im März 2018 löste die Nutzung von Facebook-Profildaten durch das US-amerikanische Unternehmen *Cambridge Analytica* den sogenannten *Facebook-Skandal* aus.[51] *Cambridge Analytica* hatte mithilfe ihrer Facebook-App *thisisyourdigitallife* Daten von Facebook-Nutzer*innen und ihren Kontakten ausgelesen, unter politischen Gesichtspunkten ausgewertet und die gewonnenen Erkenntnisse für die Kampagne von Donald Trump im US-Präsidentschaftswahlkampf genutzt.

Für Insider war dies freilich nichts Neues; bereits zuvor hatte die Datenanalystin Cathy O'Neil darauf hingewiesen:

> „Zurzeit üben Mammutfirmen wie Google, Amazon und Facebook eine enorme Kontrolle über die Gesellschaft aus, weil sie die Daten kontrollieren. Sie sacken riesige Profite ein, laden dabei aber die Verantwortung, die Fakten zu prüfen, bei anderen ab. Es ist kein Zufall, dass Steve Bannon, selbst während er gezielt daran arbeitet, das öffentliche Vertrauen in Wissenschaft und wissenschaftliche Fakten zu untergraben, im Verwaltungsrat von Cambridge Analytica sitzt – einer politischen Datenfirma, die behauptet, sie habe Trump zu seinem Wahlsieg verholfen, während sie zugleich damit prahlt, geheime ‚Voter suppression'-Kampagnen durchgeführt zu haben, um bestimmte Wählergruppen von der Stimmabgabe abzuhalten."[52]

Es wird angenommen, dass sogenannte *Filter Bubbles* und Echokammern das *Microtargeting* zusätzlich verstärken – auch wenn es bisher keine Studien gibt, die das explizit nachweisen.[53] Einmal gefasste Meinungen werden demnach immer wieder bestätigt und dadurch verstärkt, da der eigenen Auffassung widersprechende Positionen durch entsprechende Klassifikationsalgorithmen ausgefiltert werden. Radikalisierung wird dadurch gefördert, echte Kommunikation erschwert.

Ende Mai 2018 entschuldigte sich Facebook-Chef Mark Zuckerberg wegen des Facebook-Skandals in einer Anhörung des Europaparlaments dafür, dass sein Unternehmen der Verantwortung für die Verbreitung von Falschmeldungen („Fake News"), ausländischer Wahlbeeinflussung und widerrechtlichem Datenabfluss nicht gerecht geworden sei.[54] Zuckerbergs Antworten auf Fragen blieben aber unbefriedigend und unglaubwürdig, einige Fragen ließ er gänzlich unbeantwortet. Die Entschuldigung passte in eine lange Reihe von Rechtsbrüchen und anschließenden Beteuerungen seit der Gründung Facebooks; offensichtlich jedes Mal ohne nennenswerte Konsequenzen.[55] Falschmeldungen sollten mit Hilfe von Künstlicher Intelligenz bekämpft werden, ganz im Sinne des unkritischen Zeitgeists, der in der IT die „Lösung" für alle komplexeren Gesellschaftsprobleme sieht.[56]

Cambrige Analytica wurde für das politische Ausnutzen von einigen Dutzend Millionen Datensätzen weltweit kritisiert, während das ganz normale Geschäftsmodell Facebooks genau darin besteht, die gleichen Methoden auf Milliarden Datensätze anzuwenden – sowohl für Produktwerbung als auch für politische Kampagnen.[57] Wir müssen damit rechnen, dass Techniken des *Microtargeting* und politischer Beeinflussung

weiter verfeinert werden. Dadurch entstehen erhebliche Möglichkeiten der Manipulation der Öffentlichkeit – und in der Folge erhebliche Risiken für die Bürgerrechte, was Politik und Zivilgesellschaft unbedingt abwenden sollten und mit dem notwendigen Willen auch können. Ein konsequent zu Ende gedachter Datenschutz, klare Verantwortungszuschreibung und ein klarer Rechtsrahmen spielen dabei eine Schlüsselrolle.

Überwachungskapitalismus: Eine antidemokratische Kraft

Doch es gibt auch andere Stimmen: Shoshana Zuboff (s. Rezension in diesem Heft) beispielsweise geht noch einen Schritt weiter, indem sie einen allgegenwärtigen Überwachungskapitalismus skizziert, der grundsätzlich gegen die Demokratie gerichtet ist:

> „Durch die erfolgreiche Durchsetzung des Anspruchs auf Freiheit und Wissen sowie auf seine strukturelle Unabhängigkeit vom Menschen – weder als Angestellte noch als Kundschaft – manövriert uns der Überwachungskapitalismus heute mittels der radikalen Indifferenz, die diese Ansprüche sowohl bedingen als auch ermöglichen und aufrechterhalten, einer Gesellschaft entgegen, in der der Kapitalismus nicht länger als Mittel inklusiver politischer und ökonomischer Institutionen funktioniert. Stattdessen müssen wir den Überwachungskapitalismus als das anerkennen, was er ist: eine zutiefst antidemokratische soziale Kraft.“[58]

Lösungsansätze: Datenschutz ist Demokratieschutz

Macht es da noch Sinn, über Alternativen nachzudenken? Ein Ansatz, mit den Risiken des Überwachungskapitalismus umzugehen, ist der Datenschutz. Bekanntlich wurde am 25. Mai 2018 die europäische Datenschutz-Grundverordnung (DSGVO) wirksam. Auf der 95. Datenschutz-Konferenz Ende April 2018 in Düsseldorf setzten sich die Datenschutzbeauftragten des Bundes und der Länder mit den Vorfällen um *Cambridge Analytica* auseinander.[59] „Der Datenskandal um Facebook und Cambridge Analytica wirft ein Schlaglicht auf den Umgang mit Millionen Nutzerdaten“, stellen sie fest. „Das Vorkommnis zeigt ... die Risiken für Profilbildung bei der Nutzung sozialer Medien und anschließendes Mikrotargeting, das offenbar zur Manipulation von demokratischen Willensbildungsprozessen eingesetzt wurde.“

Die Datenschutzbeauftragten fordern:

1. Ausrichtung der Geschäftsmodelle sozialer Netzwerke auf die neuen europäischen Datenschutzregelungen einschließlich angemessener Vorkehrungen gegen Datenmissbrauch.

2. Offenlegung, in welchem Umfang die Facebook-Plattform für App-Anbieter geöffnet wird, Nennung belastbarer Zahlen betroffener Personen und Information Betroffener über Rechtsverletzungen.
3. Sicherstellung, dass die Vorgaben der DSGVO rechtskonform umgesetzt werden.
4. Entflechtung des Facebook-Konzerns mit Blick auf das Wettbewerbs- und Kartellrecht, wenn sich das Unternehmen wettbewerbswidrige Vorteile durch systematische Umgehung des Datenschutzes verschafft; europäische Initiativen, um monopolartige Strukturen zu begrenzen und Transparenz von Algorithmen herzustellen.

Inwieweit diese Forderungen international durchgesetzt werden und langfristig wirksam sind, hängt wesentlich vom politischen Willen nicht zuletzt Deutschlands und Europas ab. Es ist aber zweifellos wichtig, dass sich die Gesellschaft mit dem Einfluss der Algorithmen in Wirtschaft und öffentlicher Verwaltung beschäftigt und dass deren Parameter sowie Optimierungskriterien offen gelegt werden. Der Datenschutz soll dabei helfen, die Machtasymmetrien zwischen strukturell begünstigten Organisationen und natürlichen Personen zu thematisieren und für wirksame Schutzmaßnahmen für die Betroffenen zu sorgen.[60]

Wichtig ist dabei: Datenschutz schützt genau so wenig die Daten, wie Lärmschutz den Lärm schützt. Datenschutz schützt Menschen – vor der Übergriffigkeit großer Organisationen, die aus einer Machtposition heraus agieren. Versuchen, den Datenschutz im Allgemeinen und die DSGVO im Besonderen durch scheinbar absurde Beispiele lächerlich zu machen,[61] muss mit aller Macht entgegengetreten werden. Solche Beispiele verweisen aber neben aller Polemik auch auf einen möglichen Konstruktionsfehler der DSGVO, den Alexander Roßnagel bereits angesprochen hat: die angestrebte Technikneutralität der Verordnung, die aus seiner Sicht völlig verschiedene Risikoklassen vermengt und damit das Ausmaß der tatsächlich entstehenden Risiken aus dem Blick verliert: „In keiner Regelung werden die spezifischen Grundrechtsrisiken z. B. von smarten Informationstechniken im Alltag, von Big Data, Cloud Computing oder datengetriebenen Geschäftsmodellen, Künstlicher Intelligenz und selbstlernenden Systemen angesprochen oder gar gelöst. Die gleichen Zulässigkeitsregeln, Zweckbegrenzungen oder Rechte der betroffenen Person gelten für die wenig riskante Kundenliste beim ‚Bäcker um die Ecke' ebenso wie für diese um Potenzen risikoreicheren Datenverarbeitungsformen."[62]

Der Datenschutz stellt klare Kriterien bereit, deren Erfüllung bei der Einhegung des maschinellen Lernens und der schwachen KI helfen. Sie sind im Standard-Datenschutzmodell[63] festgehalten. Dabei werden die klassischen Schutzziele der IT-Sicherheit durch weiterführende Schutzziele ergänzt:

- Vertraulichkeit: Die Verarbeitung personenbezogener Daten muss so gestaltet sein, dass nur Befugte auf Daten, IT-Systeme und Prozesse zugreifen können.
- Integrität: Daten dürfen nicht verfälscht werden, bzw. wenn sie verfälscht werden, muss dies erkennbar und korrigierbar sein. Die Verarbeitung perso-

nenbezogener Daten und die dafür zulässigen Funktionen ergeben sich aus dem Zweck der Datenverarbeitung; Abweichungen vom Zweck müssen erkennbar, beurteilbar und korrigierbar sein.
- Verfügbarkeit: Erwartete Maßnahmen müssen bzgl. Daten, IT-Systemen und Prozessen zeitnah und in einer definierten Qualität erbracht werden können.
- Transparenz: Die verarbeiteten Systeme, die Systeme, die Funktionsweise und die Wirkungen einer Verarbeitung müssen verständlich sein. Die Datenverarbeitung muss prüfbar sein – anhand von Datenschutzerklärungen, Einwilligungen, Verträgen, Spezifikationen, Dokumenten, Log- und Protokolldateien.
- Nicht-Verkettung: Die Verarbeitung von Daten, IT-Systemen und Prozessen unterschiedlicher Verfahren muss getrennt erfolgen und nicht verknüpfbar sein. Dies erfordert Datenminimierung, politische Gewaltenteilung, unabhängige Marktakteure und die Berücksichtigung unterschiedlicher politischer Interessen (Diversität). Es schließt die Zweckbindung ein – Daten, die für unterschiedliche Zwecke erhoben werden, dürfen nicht nachträglich verknüpft und gemeinsam ausgewertet werden.
- Intervenierbarkeit: Die technische Verarbeitung personenbezogener Daten muss verändert oder auch angehalten werden können.

Entscheidend ist, diese Anforderungen auch in KI-Systemen konsequent umzusetzen. Die Betreiber*innen von KI-Systemen dürfen sich nicht darauf zurückziehen, dass sie deren Funktionsweise nicht mehr kontrollieren können. Für alle Formen der KI muss den Betreibenden eine klare Verantwortung zugeschrieben werden.

Eine klare Verantwortungszuschreibung bedeutet aber auch: Es ist immer klar, wer für die Ergebnisse von KI-Systemen zu Rechenschaft gezogen werden kann. Hier sind die Politik und der Gesetzgeber gefragt: Sie müssen dafür sorgen, dass ein klarer Rechtsrahmen die Nutzung von Systemen der Künstlichen Intelligenz reguliert und dadurch Rechtsstaatlichkeit und Demokratie gewährleistet.

STEFAN HÜGEL ist Diplom-Informatiker und studierte in Karlsruhe und Freiburg im Breisgau. Er ist als IT-Berater tätig. Ehrenamtlich ist er als Vorsitzender des Forums InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) sowie als Mitglied des Bundesvorstands der Humanistischen Union aktiv. Zu seinen Themenschwerpunkten gehören Netzpolitik, Datenschutz, IT-Sicherheit, die Auswirkungen Künstlicher Intelligenz sowie *Information Warfare*.

Anmerkungen:

1 Weizenbaum J. (1987 [1976]): Die Macht der Computer und die Ohnmacht der Vernunft, 2. Auflage. Frankfurt am Main: Suhrkamp, S. 268.
2 Was 1968 als harmloser Schlager daherkam, ist heute in einigen Bereichen durchaus Realität, vgl. Dating-Apps wie Tinder. Ob das schon Künstliche Intelligenz ist, wäre zu diskutieren (Quelle: https://genius.com/France-gall-der-computer-nr-3-lyrics).
3 Ramge T. (2018): Mensch und Maschine. Wie Künstliche Intelligenz und Roboter unser Leben verändern. Stuttgart: Reclam.
4 Wahlster W. (2015): „Künstliche Intelligenz ist besser als natürliche Dummheit". Interview, Jan-Bernd Meyer, CIO, https://www.cio.de/a/kuenstliche-intelligenz-ist-besser-als-natuerliche-dummheit,3230970,2.
5 Futurezone (2017): Facebook: Gezielte Werbung für unsichere und gestresste Teenager, https://www.futurezone.de/digital-life/article210422803/Facebook-Gezielte-Werbung-fuer-unsichere-und-gestresste-Teenager.html
6 Der Begriff geht auf den Universalgelehrten und Mathematiker Al-Chwarizmi zurück, der um 800 u. Z. hauptsächlich in Bagdad lebte: Wikipedia, Stichwort „Al-Chwarizmi", https://de.wikipedia.org/wiki/Al-Chwarizmi.
7 „Endlich viele" bezieht sich auf die Beschreibung des Algorithmus, nicht darauf, ob ein Algorithmus nach endlich vielen Schritte zum Halten kommt. Ein grundsätzliches Problem der theoretischen Informatik ist das *Halteproblem*: Es kann nachgewiesen werden, dass es keinen Algorithmus gibt, der für alle Algorithmen und beliebige Eingaben entscheiden kann, ob ihre Ausführung an ein Ende kommt.
8 Ein einfaches Beispiel für einen Algorithmus in der Offline-Lebenswelt wäre ein parametrisiertes Kochrezept, in dem die Zutaten (also i. w. S. die „Daten") und die Verarbeitungsschritte beschrieben und abhängig von den Eingabevariablen unterschiedliche Ausgaben geliefert werden. Die Wikipedia liefert verschiedene Definitionen für Algorithmen unter: Wikipedia (en.), Stichwort „Algorithm characterizations", https://en.wikipedia.org/wiki/Algorithm_characterizations.
9 Schinzel B. (2018): Workshop „Algorithmen: schuldig oder unschuldig?" FIfF-Kommunikation 1/2018, S. 60-61
10 Ebd.
11 Manovich L. (2018): Media Analytics & Gegenwartskultur, in: Engemann C, Sudmann A Hg. (2018) Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz. Bielefeld: transcript, S. 269-288.
12 J. McCarthy (1955), zit. n. Ertel W. (2016): Grundkurs Künstliche Intelligenz. 4. Auflage. Wiesbaden: Springer Vieweg.
13 Turing A. M. (1992 [1950]): Maschinelle Rechner und Intelligenz, in: Hofstadter DR, Dennett DC (1992 [1981]) Einsicht ins Ich. Fantasien und Reflexionen über Selbst und Seele. Stuttgart: Klett-Cotta im Deutschen Taschenbuch-Verlag, S. 59-72
14 Wikipedia, Stichwort „Künstliche Intelligenz", https://de.wikipedia.org/wiki/Künstliche_Intelligenz.
15 Searle J. R. (1992 [1980]): Geist, Gehirn, Programm, in: Hofstadter DR, Dennett DC (1992 [1981]) a.a.O., S. 337-356.
16 Eine frühe Implementierung dieses Konzepts ist das Sprach-Analyse-Programm ELIZA, mit dem auf diese Weise „Gespräche" simuliert werden können, s. Weizenbaum J. (1987 [1976]): a.a.O., S. 14ff.
17 Polemisch-überspitzt könnte man sagen: Wenn es funktioniert, ist es keine Künstliche Intelligenz mehr. Es ist natürlich letztlich nur eine Frage der Wahrnehmung.

18 Kritisch zur Vermenschlichung der Computer: Rehak, R. (2016), Die Macht der Vermenschlichung und die Ohnmacht der Begriffe. FIfF-Kommunikation 2/2016, S. 45-46. Ich verwende den allgemein gebräuchlichen Begriff des maschinellen Lernens hier dennoch im Bewusstsein, dass maschinelles „Lernen" von menschlichem Lernen abgegrenzt werden muss.

19 Nicht-berechenbare Aufgaben können durch ein Computersystem mathematisch beweisbar nicht erfüllt werden (vgl. z. B. Kfoury A. J., Moll R. N., Arbib M. A. (1986): A Programming Approach to Computability, 2nd printing, New York, Heidelberg, Berlin: Springer-Verlag). Hohe Komplexität der Berechnung kann dazu führen, dass die technische Leistungsfähigkeit von (heutigen) Systemen nicht dazu ausreicht. (Dies macht man sich z. B. bei Verschlüsselungsverfahren zu nutze.)

20 Eine gut verständliche Einführung gibt es auch als Video eines Vortrags beim 35. Chaos Communication Congress (35c3): teubi (2018): Introduction to Deep Learning, https://media.ccc.de/v/35c3-9386-introduction_to_deep_learning.

21 Gesellschaft für Informatik (2018a): Technische und rechtliche Betrachtungen algorithmischer Entscheidungsverfahren. Studien und Gutachten im Auftrag des Sachverständigenrats für Verbraucherfragen. Berlin: Sachverständigenrat für Verbraucherfragen, http://www.svr-verbraucherfragen.de/wp-content/uploads/GI_Studie_Algorithmenregulierung.pdf.

22 Darum führt der Vorschlag eines „Algorithmen-TÜV" in die Irre. Er könnte lediglich prüfen, ob der Lernalgorithmus korrekt arbeitet, also die vorgefundenen Daten korrekt klassifiziert.

23 Ng A., Soo K. (2017): Data Science – Was ist das eigentlich? Algorithmen des maschinellen Lernens verständlich erklärt. Berlin: Springer-Verlag.

24 Beispielsweise könnte sich während der Lernschritte ein Zusammenhang zwischen der Haarfarbe und der Parteipräferenz herausstellen – offensichtlich kein inhaltlich sinnvolles Merkmal für die Klassifizierung. Solche Klassifizierungen sind i. d. R. durch die Einschränkungen des zum Lernen verwendeten Datenbestands verursacht. Das führt dazu, dass die Klassifikation nicht mehr nachvollziehbar ist und die korrekte Einordnung auch nicht mehr überprüft werden kann.

25 Kühl E. (2015): Gesichtserkennung: „Meine Freundin ist kein Gorilla", Zeit Online, https://www.zeit.de/digital/internet/2015-07/google-fotos-algorithmus-rassismus/komplettansicht.

26 Auf diese Problematik weist O'Neil anhand von Beispielen hin: O'Neil C. (2017): Angriff der Algorithmen. Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden. München: Carl-Hanser-Verlag.

27 Dies musste Bettina Wulff, die Frau des ehemaligen Bundespräsidenten, feststellen, als ihr Verbindungen zum Rotlichtmilieu unterstellt wurden. Nachdem die Google-Suche im Eingabefeld entsprechende Vervollständigungen anbot, erhöhte sich auch die Klickrate, was wiederum – zusammen mit der öffentlichen Berichterstattung – diese Vervollständigung weiter förderte. Abhilfe kann dadurch geschaffen werden, dass bestimmte „unerwünschte" Ergänzungen eliminiert werden – was in der Praxis nicht immer möglich sein wird und strenggenommen die Lernergebnisse verfälscht.

28 Wolfangel E. (2018): Programmierter Rassismus, Zeit Online, https://www.zeit.de/digital/internet/2018-05/algorithmen-rassismus-diskriminierung-daten-vorurteile-alltagsrassismus/komplettansicht.

29 Kühl E. (2015), a.a.O.

30 Angwin J., Larson J., Mattu S., Kirchner L. (2016): Machine Bias, ProPublica, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; vgl. auch Krafft T. D. (2018): Qualitätsmaße algorithmischer Entscheidungssysteme in der Kriminalprognostik, FIfF-Kommunikation 1/2018, S. 50-55.

31 Wolfangel E. (2018), a.a.O.

32 Schinzel B. (2018), a.a.O.

33 Schinzel B. (2017): Algorithmen sind nicht schuld, aber wer oder was ist es dann? FIfF-Kommunikation 2/2017, S. 5-9.

34 z. B. Gesellschaft für Informatik (2018b): Unsere ethischen Leitlinien. Verabschiedet vom Präsidium der GI am 29. Juni 2018, https://gi.de/fileadmin/GI/Allgemein/PDF/GI_Ethische_Leitlinien_2018.pdf.
35 Europäische Kommission (2018): Ethics Guidelines for Trustworthy AI. High-Level Expert Group on Artificial Intelligence, Draft, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=57112
36 Rost M. (2017): Bob, es ist Bob! FIfF-Kommunikation 4/2017, S. 63-66.
37 Scott B., Heumann S., Lorenz P. (2018): Artificial Intelligence and Foreign Policy. Policy Brief. Berlin: Stiftung neue Verantwortung, https://www.stiftung-nv.de/sites/default/files/ai_foreign_policy.pdf.
38 Scott B., Heumann S., Lorenz P. (2018), a.a.O., S. 2f.
39 Gesellschaft für Informatik (2018a), a.a.O., S. 13.
40 Suder K. (2018): „Es geht um den Kern von Sicherheit". Interview, IP Internationale Politik 4/2018, S. 14-19.
41 Koch B., Schörnig N. (2017): Autonome Drohnen – die besseren Waffen? In: vorgänge Nr. 218 (2/2017), S. 43-53.
42 Department of Defense (2012): Autonomy in Weapon Systems. Directive No. 3000.09, https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.
43 Koch B., Schörnig N. (2017), a.a.O., S. 45.
44 Ramge T. (2018), a.a.O.
45 Neben der hier thematisierten Manipulation haben Monopole natürlich auch wirtschaftliche Auswirkungen, s. o.
46 Im Dezember 2018 wurde über Pläne der Bundesregierung berichtet, eine Kennzeichnungspflicht für Social Bots einzuführen: Wittenhorst T. (2018): Bundesregierung erwägt Kennzeichnungspflicht für Social Bots, heise.de, https://www.heise.de/newsticker/meldung/Bundesregierung-erwaegt-Kennzeichnungspflicht-fuer-Social-Bots-4252095.html. Zu klären ist, wie eine solche Pflicht durchgesetzt werden kann, nachdem es gerade das Ziel bei Social Bots ist, dass sie nicht als solche erkannt werden.
47 Bertelsmann-Stiftung (Hg.) (2009): Lernen von Obama? Das Internet als Ressource und Risiko für die Politik. Gütersloh: Verlag Bertelsmann-Stiftung.
48 Eine positive Sicht auf Nudging vertreten Thaler R. H., Sunstein C. (2015 [2008]): Nudge. Wie man kluge Entscheidungen anstößt. 5. Auflage, Berlin: Ullstein.
49 Schieren S. (2016): Die Macht der Algorithmen. Politikum 1/2016, S. 4-12.
50 Thaler R. H., Sunstein C. (2015 [2008]), a.a.O.
51 Dachwitz I., Rudl T., Rebiger S. (2018): FAQ: Was wir über den Skandal um Facebook und Cambridge Analytica wissen. netzpolitik.org, https://netzpolitik.org/2018/cambridge-analytica-was-wir-ueber-das-groesste-datenleck-in-der-geschichte-von-facebook-wissen/.
52 O'Neil C. (2017), a.a.O.
53 S. Pariser E. (2012): Filter Bubble. Wie wir im Internet entmündigt werden. München: Carl-Hanser-Verlag; Zweig K. A., Deussen O., Krafft T. D. (2017): Algorithmen und Meinungsbildung. Eine Grundlegende Einführung, in: Deussen O., Zweig K. A. (2017): Algorithmen und Meinungsbildung. Schwerpunktheft. Informatik-Spektrum 40(4), S. 318-326.
54 Holland M. (2018): Facebook-Datenskandal: Kaum Antworten von Zuckerberg im Europaparlament, Heise.de, https://www.heise.de/newsticker/meldung/Facebook-Datenskandal-Zuckerberg-entschuldigt-sich-im-Europaparlament-4055074.html.
55 Fowler G. A., Esteban C. (2018): 14 years of Mark Zuckerberg saying sorry, not sorry. https://www.washingtonpost.com/graphics/2018/business/facebook-zuckerberg-apologies/.
56 Holland M. (2018), a.a.O.
57 Rieger F. (2018): Facebook muss zerschlagen werden! https://www.br.de/radio/bayern2/sendungen/zuendfunk/frank-rieger-facebook-muss-zerschlagen-werden-100.html.

58 Zuboff S. (2018): Der dressierte Mensch. Die Tyrannei des Überwachungskapitalismus, Blätter für deutsche und internationale Politik 11‘18, S. 101-111. Der Text basiert auf Zuboff S. (2018): Das Zeitalter des Überwachungskapitalismus. Frankfurt, New York: Campus-Verlag.

59 Entschließung der 95. Datenschutz-Konferenz am 25./26. April 2018 in Düsseldorf: Facebook-Datenskandal – Neues Europäisches Datenschutzrecht bei Sozialen Netzwerken durchsetzen! DuD 7/2018, S. 447.

60 Rost M. (2018): Künstliche Intelligenz – Normative und operative Anforderungen des Datenschutzes, DuD 9/2018, S. 558–565.

61 Ein Beispiel, das von einigen Wochen durch die Medien ging beschäftigte sich mit der Frage, ob die EU-Datenschutz-Grundverordnung das Anbringen von Namensschildern an Türklingeln verbiete: dazu z. B. https://www.faz.net/aktuell/wirtschaft/datenschutz-verstossen-klingelschilder-gegen-die-dsgvo-15844615.html.

62 Roßnagel A. (2018): Datenschutz-Grundverordnung – was bewirkt sie für den Datenschutz? In: vorgänge Nr. 221/222 (1-2/2018), S. 17-29.

63 Das Standard-Datenschutzmodell wurde von den Datenschutzbehörden erarbeitet und ist hier zu finden: https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methode_V1.1.pdf. Eine speziell auf die Problematik der KI ausgerichtete Darstellung findet sich in Rost M. (2018): a.a.O., S. 558–565. Eine leichtverständliche Einführung gab es beim 35. Chaos Communication Congress: Rehak R. (2018): Was schützt eigentlich der Datenschutz? Warum DatenschützerInnen aufhören müssen, von individueller Privatheit zu sprechen, https://media.ccc.de/v/35c3-9733-was_schutzt_eigentlich_ der_datenschutz.