

AKTUELLE THEMEN DER KI WS20/21
AM3

KÜNSTLICHE INTELLIGENZ UND ETHIK

Vanessa Voge



ZUSAMMENFASSUNG

Die vorliegende Arbeit beschäftigt sich mit den Chancen und – im speziellen – Risiken der künstlichen Intelligenz aus einer ethischen Sicht heraus. Gegenstand der Diskussion hierbei ist unter anderem das Blackbox-Problem, Bias und die Verantwortungsfrage. Mittels diversen Erhebungen aus aktuellen Umfragen und Studien, relevanter Literatur, sowie Fallbeispiele aus der Industrie, wird diesen Themen auf den Grund gegangen.

Anhand dieser Ressourcen lässt sich festhalten, dass Ethik eine größer werdende Rolle rund um KI spielt, da der Einfluss auf unser aller Leben rapide zunimmt und man so nicht drum herum kommen wird, diese Entwicklungen auch aus einer moralisch-gesellschaftlichen Sicht heraus zu behandeln. Generell ist die künstliche Intelligenz, als “junge” Disziplin, aktuell noch auf der Suche nach den eigenen Grenzen und es bleibt abzuwarten, wie sich die Bemühungen um eine ethische KI auf die weitere Entwicklung der Technologie auswirken werden.

ETHIK UND KI – WIE PASST DAS ZUSAMMEN?

Definition der Grundbegriffe

Die Ethik ist eine Teildisziplin der Philosophie und wird anhand dessen Anwendbarkeit auf Situationen aus dem echten Leben nicht selten auch als praktische Philosophie bezeichnet (Levina, 2020). Kernthema der Ethik ist die “Begründbarkeit und Reflektion der moralischen Grundsätze für das Verhalten oder Handeln einer Person oder Gruppe von Personen” (Levina, 2020, 2. Absatz). Die Ethik selbst hat weitere Teilbereiche, wie die digitale Ethik zum Beispiel, welche ein Augenmerk auf relevante Themen in Zeiten der Digitalisierung setzt, wie Soziale Medien, das Internet und nicht zuletzt die künstliche Intelligenz. Bei letzterem steht die Frage im Mittelpunkt, inwieweit das Leben eines Individuums oder gar einer Gesellschaft durch KI verbessert werden kann, wo aber auch Bedenken aufgeworfen werden können. Diese könnten die Autonomie und Freiheit des Menschen beispielsweise betreffen oder die Lebensqualität.

Unter künstlicher Intelligenz versteht man die Fähigkeit einer Software bzw. eines “intelligenten Systems” kognitive Tätigkeiten auszuführen, die sonst ein Mensch erledigen würde (Zweig, 2019, S. 126). Da dies aber eine breit gefasste Definition ist, immerhin liegt selbst in der Wissenschaft noch kein umfassendes Verständnis dessen vor, was eine “Intelligenz” wirklich ausmacht, wird grundsätzlich zwischen zwei Arten von künstlicher Intelligenz unterschieden: der schwachen und der starken (Beiter et al., 2020, S. 8 – 9). Schwache Intelligenz beschreibt Systeme, die auf eine gesonderte, zumeist alleinstehende, Aufgabe zugeschnitten sind und hierbei die Fähigkeiten eines Menschen übertreffen können, zum Beispiel im Schach (Zweig, 2019, S. 267). Man spricht von solch einer KI jedoch als schwach, da diese ihr “Wissen” nicht oder nur sehr schwer auf andere Situationen oder Aufgaben anwenden können. Starke KI soll dagegen die Fähigkeit aufweisen, Korrelationen zu erkennen, und so universell Probleme lösen können, ebenbürtig oder gar überlegen zum Menschen. Letztere ist aktuell unerreicht, demnach spricht man gängigerweise von schwacher KI, wenn es um den aktuellen Stand der Technik geht.

KI-basierte Systeme finden zunehmend Anwendung in vielerlei Bereiche unseres Alltags (Beiter et al., 2020, S. 6), zum Beispiel in technischen Endgeräten wie Smartphones (Autokorrektur, Sprachassistenten, Gesichtserkennung), aber auch im Haushalt (Smart Fridge, Saugroboter) und im Straßenverkehr (autonome Fahrzeuge, Objekt- und Gefahrenerkennung). Anhand der Vermehrung von Berührungspunkten mit KI lässt sich auch ein Trend feststellen, dass Verbraucher die Chancen von künstlicher Intelligenz zunehmend höher einschätzen, als die Gefahren (siehe **Abbildung 1**).

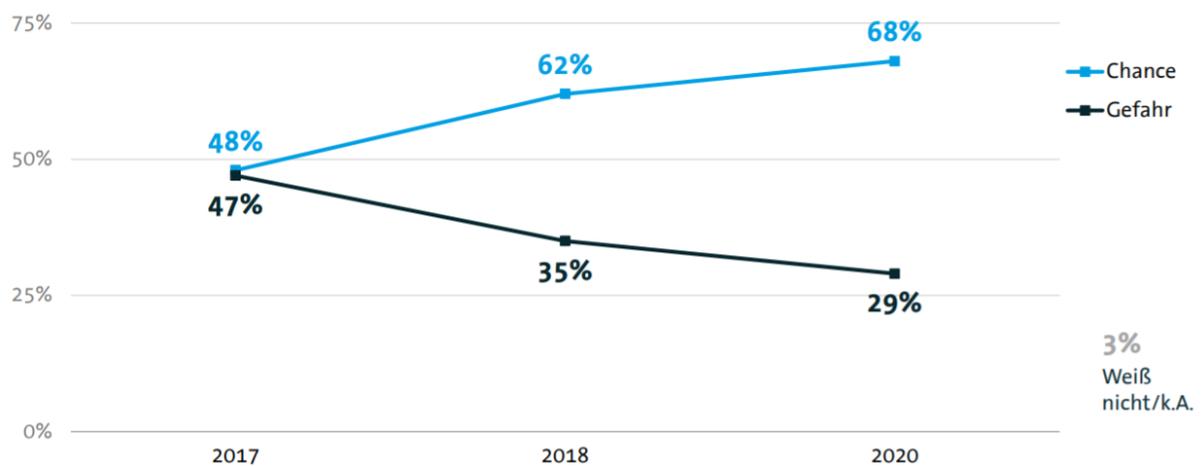


Abbildung 1. Verhältnis von Chance zu Gefahr über die letzten Jahre (Quelle: Berg, A. & Dehmel, 2020)

Aus dieser aktuellen Umfrage von Bitkom (Berg, A. & Dehmel, 2020) lässt sich aber auch entnehmen, dass sich 85 % eine sichere und im Voraus gründlich geprüfte KI wünschen. Eine Studie von Capgemini (2019) stellte zudem einen gewissen Konsens zwischen Verbrauchern und Industrie fest: 86 % gaben zu, dass die Nutzung von KI kritische Fragen zur Ethik aufwerfen würde.

Interview mit Korrespondentin des IDE

Um diesen Fragen auf den Grund zu gehen, wurde das Fachwissen von Nadine Hammele, wissenschaftliche Mitarbeiterin am Institut für digitale Ethik der Hochschule der Medien, herangezogen. Neben einigen Ausschnitten aus dem Interview, welche im Laufe der Arbeit vorkommen werden, ist das vollständige Skript am Ende zu finden.

Die Doktorandin gab als Einstieg folgendes zu bedenken:

Dadurch, dass eine Maschine, die nicht so viel Weltwissen hat wie der Mensch, in Entscheidungen integriert wird, gibt es einfach sehr viele Dinge zu berücksichtigen und da spielt die Ethik eine große Rolle.

In welchen expliziten Fällen dies zu Reibungspunkten führen kann, wird im übernächsten Kapitel erörtert.

Davor soll aber zunächst eine Übersicht der aktuellen Entwicklungen rund um KI gegeben werden.

AKTUELLE ENTWICKLUNGEN

Wo stehen wir aktuell? Um dieser Frage auf den Grund zu gehen, wurden die Meilensteine des letzten Jahres (2020) im Folgenden zusammengefasst.

Das mächtige Sprachmodell GPT-3 von OpenAI wird seit Juni 2020 kommerziell vertrieben (Bastian, 2020). Kunden können über eine Schnittstelle Zugang zur Text-KI erhalten und für eigene Produkte nutzen. Microsoft, als Großinvestor, hat sich hierbei Exklusivrechte am Code von GPT-3 gesichert.

Was Microsoft für OpenAI ist, ist Google für die interne Tochter DeepMind. Dessen proteinfaltende KI AlphaFold 2.0 erzielte beim CASP (= Critical Assessment of Protein Structure Prediction), ein Wettbewerb, der alle zwei Jahre zum Zwecke des Prüfens von Proteinfaltungsvoraussagungsmethoden stattfindet, einen historischen Durchbruch und sagte von 100 Proteinsequenzen die Struktur von 70 präzise voraus (Schreiner, 2021). Das Wissen um die Faltung einer Proteinsequenz hilft dem Verständnis und Erklärbarkeit unter anderem von Erbkrankheiten oder Krebs. DeepMind hat zudem MuZero als Nachfolger von AlphaZero veröffentlicht. Die KI lernt selbständig analoge und digitale Spiele ohne jegliches Vorwissen und schlägt allesamt dessen Vorgänger.

Disney hat die KI-Technologie Megapixel Deepfakes vorgestellt, mit der sich hochauflösende Deepfakes generieren lassen können.

So könnten verstorbene Schauspieler beispielsweise mittels Körperdouble wiedererweckt werden (Naruniec et al., 2020). Zudem breitet sich die KI-Überwachung aus: die Londoner Polizei führte bereits Anfang 2020 offiziell Gesichtserkennung ein (Haefliger, 2020).

Die aufkommenden ethischen Bedenken unter anderem zu intelligenten Überwachungssystemen brachten die Veranstalter der NeurIPS, der größten KI-Konferenz, dazu, einen Pflichtabschnitt in den zu einreichenden Forschungspapieren einzufordern (Schreiner, 2021). In diesem muss dargestellt werden, inwieweit die Arbeit soziale Auswirkungen nach sich ziehen könnte, auch im negativen Sinne.

Die benannten Beispiele zeigen, wie rasant sich Technologien mit künstlicher Intelligenz weiterentwickeln und den geschäftlichen, sowie auch den privaten Bereich zunehmend beeinflussen. Nicht zuletzt wird dieser Einfluss unter zahlreichen Gesichtspunkten, wie Transparenz, Fairness und Verantwortung auch kritischer betrachtet. Wo hierbei das Konfliktpotential von KI-Anwendungen liegt, wird im nächsten Kapitel erläutert.

KONFLIKTE

Die folgenden Konflikte wurden unter den drei Überkategorien von Diskriminierung, Transparenz und Verantwortung aufgeteilt und werden im Einzelnen genauer besprochen.

Diskriminierung

Auf die Frage hin, wo die Reibungspunkte von KI liegen, antwortete Nadine Hammele (persönliche Kommunikation, 05. Januar 2021) mitunter folgendes:

Ein weiteres Thema ist die Diskriminierung. Das Problem ist hierbei ganz oft, dass die Datensätze, mit welchen eine künstliche Intelligenz lernt, Daten beinhaltet, die historisch gewachsen oder unvollständig sind. Aufgrund dieser verzerrten Daten könnten so bestimmte Personen benachteiligt werden.

Diese Verzerrung der Daten wird auch als Bias bezeichnet. Hierbei fließen im Training und Aufsetzen des intelligenten Algorithmus die – wenn oft auch unbewussten – Vorurteile der Entwickler mit ein.

So kann es zum Beispiel bei einer Unterrepräsentation von bestimmten Gruppen in den Trainingsdaten zu diskriminierenden KIs kommen. Charakteristisch hierfür wäre Gesichtserkennungs-Software zu nennen, die im Erkennen speziell von schwarzen Frauen deutlich schlechter abschnitt/abschneidet als von weißen Personen (Buolamwini, 2017). Joy Buolamwini, Informatikerin am MIT Media Lab, hat sich diesem Problem angenommen, nachdem ihr – damals noch als Studentin – aufgefallen war, dass solche Systeme sie partout nicht erkannten oder falsch identifizierten. Den durchgeführten Studien dazu von Buolamwini werden einem nicht unerheblichen Einfluss auf die Weiterentwicklung von Gesichtserkennungs-KI von Google, Microsoft und Co. zugeschrieben, welche dem Bias in ihren Produkten entgegengewirkt haben (Burt, 2020).

Wie im Interview schon erwähnt, können auch historisch gewachsene Daten problematisch sein. Dies kommt dann beispielsweise im automatisierten Bewerbungsprozess zum Tragen. Füttert man eine KI mit der Unternehmenshistorie, die in nicht wenigen Bereichen, männlich dominiert ist – speziell, wenn es um Führungspositionen geht –, kann die KI das Geschlecht jeweils als Für- oder Gegenargument auffassen. So geschehen ist dies bei Amazon: Die Personalierer begannen sich zu wundern, warum das Bewerbungsverhältnis dermaßen ungleich war, bis sie erkannten, dass die KI, welche die Bewerbungen vorsortierte, anhand der Personalien aus den letzten 10 Jahren das Attribut “männlich” als positiv interpretiert hatte, während “weiblich” einem Ausschlusskriterium glich (Levina, 2020).

Nicht zuletzt können Sprachmodelle wie GPT-3 oder Chatbots wie Microsoft’s Tay beeindruckend realistische Texte generieren; da diese aber vom Internet “lernen”, kann dies auch in eine unerwünschte Richtung umschlagen. Der Chatbot Tay sollte im Rahmen eines Experiments aus der Kommunikation mit jungen Menschen lernen und dessen Sprachstil annehmen (Beuth, 2016).

Nachdem es auf sozialen Medien wie Twitter auch viele Individuen gibt, die bewusst beleidigende, diskriminierende oder schlicht unwahre Inhalte posten (sogenannte Trolle), lernte Tay eben von diesen und fing an, in ihren Tweets zum Beispiel den Holocaust zu leugnen.

Eine KI erkennt nicht von sich aus, ob etwas kontrovers oder politisch inkorrekt ist, sondern kann dies erst anhand eines mühsam integrierten Filters oder einer Vielzahl von Parametern bestenfalls einordnen und unterlassen.

Transparenz

Um zu gewährleisten, dass die künstliche Intelligenz die richtigen Schlüsse zieht, benötigt sie große Datenmengen. Die Woher? Wies? Wohins? oder Warums? bleiben dabei oft – bestenfalls – schemenhaft beantwortet (Paefgen-Laß, 2019). Das sogenannte Blackbox-Problem betrifft laut Prof. Dr. Marco Huber nicht nur Nutzer, sondern auch Entwickler (Podbregar 2020).



Abbildung 2. Illustration der Blackbox. (Quelle: Eigene Darstellung)

Blackbox heißt es deshalb, weil für beide Parteien oftmals nicht ersichtlich ist, was genau drinnen passiert – zwischen dem Verarbeiten der Eingabe und dem Ausgeben des Ergebnisses bzw. der Entscheidung (siehe **Abbildung 2**). Gerade bei Verbrauchern weckt diese Unsicherheit ein gewisses Misstrauen smarten Systemen gegenüber, da gerade Datenschutz für viele ein empfindliches Thema ist. Dies kam auch im Interview zur Sprache (N. Hammele, persönliche Kommunikation, 05. Januar 2021):

Überall da, wo wir Technologie in unseren Alltag integrieren, sammelt diese Technologie auch oft Daten über uns und oft fehlt es da dann an Transparenz: Was machen die Unternehmen, die diese Daten sammeln, mit diesen? Wie werten sie diese aus und was wird weitergeleitet?

Abgesehen vom mangelnden Wissen über die Verarbeitung der Daten lässt sich als Verbraucher auch im Vor herein schwer abschätzen, wie robust die künstliche Intelligenz ist (Beiter et al., 2020).

Werbeversprechen sind das eine, aber in der Praxis liefern solche intelligenten Applikationen unter Umständen ungenügende Ergebnisse.

So gibt es beispielsweise eine App, deren KI am Husten des Benutzers feststellen soll, ob eine Corona-Erkrankung vorliegt (Till, 2020) - wie sehr kann man sich auf diese Prognose verlassen?

Weiterhin ist die Skepsis autonomen Fahrzeugen gegenüber vergleichsweise hoch: die eingangs benannte Bitkom-Umfrage (2020) ergab bei einer Mehrheit der Befragten (57% gegen 36%) die Befürchtung, dass selbstfahrende Autos zu schweren Unfällen und vielen Toten führen werden.

Dies könnte sich mitunter auch unter dem Aspekt von gefürchteten Cyber-Angriffen erklären lassen. Ein gehacktes Auto wäre natürlich ein Super-GAU, aber auch "kleinere" Technik, die theoretisch gehackt werden könnte, birgt immer ein Sicherheitsrisiko, welches von einem Laien schwer eingeschätzt werden kann.

In Zeiten von künstlich generierten Texten, Kunst, Stimmen oder ganzen Personenabbildungen, Deepfakes und täuschend echt-klingenden Sprachassistenten wird es – speziell im Kontext des Internets – immer schwieriger einzuordnen, was echt und was fake ist. Mit dem rasanten Fortschritt der intelligenten Technologien, liegt es nahe, dass der Grat zunehmend schmaler wird.

Wie bereits erwähnt, betrifft das Blackbox-Problem aber nicht nur die Nutzer intelligenter Systeme, sondern auch deren Entwickler. Im Falle von überwachtem Lernen, haben die Entwickler zwar in der Hand, welche Daten Sie dem intelligenten Modell zum Trainieren bereitstellen, aber von dort aus arbeitet dieses autonom. Mit dem Evaluieren der Ergebnisse und dem Anpassen bestimmter Parameter hier und da, können die "Schöpfer der KI" zwar weitere Ausbesserungen vornehmen, aber speziell bei vielschichtigen neuronalen Netzen ist der Prozess der Entscheidungsfindung auch für sie nicht nachvollziehbar (Podbregar, 2020).

Dies kann, wie im vorherigen Punkt behandelt, zu diskriminierenden Algorithmen führen, aber auch in der Medizin, bei der Diagnosestellung zum Beispiel, hat die Intransparenz von künstlicher Intelligenz eine besondere Brisanz (ZEIT ONLINE, o. D.).

Verantwortung

Wie schon bei der Capgemini Studie (2019) und den aktuellen Entwicklungen der NeurIPS ersichtlich, sind sich viele Forscher und Entwickler aus dem KI-Bereich der wachsenden Verantwortung bewusst. Joseph Redmon, der Kopf hinter der Bildanalyse-KI YOLO (siehe **Abbildung 3**), hat gar die Reißleine gezogen und dessen Forschung am Projekt gestoppt (Schreiner, 2020). Zu groß schätze er das Missbrauchspotenzial durch Militär und Überwachungsmaschinerie ein.

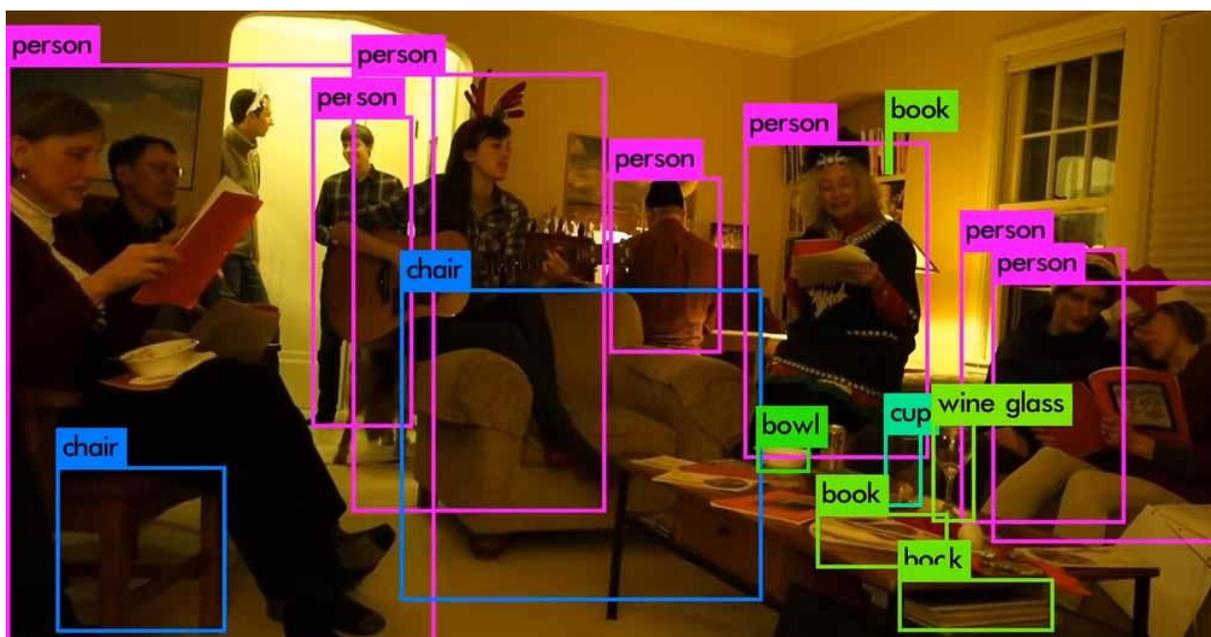


Abbildung 3. Demo der Objekterkennungssoftware YOLO. (Quelle: Schreiner, 2020)

Auch hier bieten automatisierte Fahrzeuge – gerade im Schadensfall – reichlich Konfliktpotential (ZEIT ONLINE, o. D.). Wer haftet, wenn zum Beispiel ein solches Auto jemandem die Vorfahrt nehmen sollte und einen Unfall baut? Keine Software ist perfekt und eine KI, die falsche Entscheidungen trifft, ist nicht gänzlich auszuschließen. Dieses moralische Dilemma, wie sich der Fahrcomputer entscheiden sollte, wenn ein Personenschaden unvermeidlich ist, wird vom MIT Media Lab für dessen Moral Machine aufgegriffen

(Brogan, 2016). Der Nutzer soll entscheiden: sollen die zwei jungen Insassen umkommen oder stattdessen unbeteiligte Senioren?

Tatsächlich hat eine Ethik-Kommission des deutschen Bundesverkehrsministeriums 2016 Leitlinien für Fahrcomputer erlassen: so sollen etwaige Objekterkennungsprogramme blind gegenüber Merkmalen wie Geschlecht, Alter, etc. sein, außerdem sollen mehrere Menschenleben nicht mehr wiegen als eines, da das Leben eines einzelnen aus ethischer Sicht unantastbar bleibt (ZEIT ONLINE, o.D.).

Aus einer anderen Perspektive heraus, wäre es auch an uns, zu fragen: wie wollen wir selbst mit KI ethisch umgehen? Dies ist nach aktuellem Stand der schwachen KI sehr weit vorausgedacht, aber sollte Technologie künstliche Intelligenzen mit einem Bewusstsein hervorbringen, liegt es an uns, zu entscheiden, in welchem Verhältnis diese zum Menschen stehen sollen. Als Freund und Helfer? Oder innerhalb einer untergeordneteren Rolle? Sollen diese Individuen irgendwann auch Zivilrechte genießen? Immerhin hat der Roboter Sophia, als erster Roboter überhaupt, 2017 die Staatsbürgerschaft in Saudi-Arabien erhalten (WELT, 2017) - auch wenn dies angesichts der Menschenrechtslage im Land fast schon ironisch wirkt.

Nicht zuletzt bleibt die Frage offen, wie wir einer selbstbestimmten KI moralisch-ethische Werte weitergeben wollen? Ließe sich damit das dystopische Szenario der Unterwerfung durch eine Superintelligenz vermeiden?

LÖSUNGEN

Auf dem Weg zu einer ethischeren KI gibt es neben den benannten Reibungspunkten, die es zu identifizieren gilt, aber auch Lösungsansätze und Strategien. Als aktivster Nebenzweig im Kosmos KI gilt inzwischen die “explainable artificial intelligence”, kurz XAI (Podbregar, 2020). Diese soll Licht in die Blackbox bringen, indem das intelligente Modell neben dem Ergebnis auch eine Erklärung liefert, beispielsweise einen Entscheidungsbaum.

Hierbei muss aktuell jedoch ein Kompromiss zwischen Effizienz und Erklärbarkeit gemacht werden, denn je effizienter das System, umso vielschichtiger und undurchsichtiger ist es. Auch der Kostenfaktor einer XAI ist ein Grund, warum diese sich in der Breite noch nicht durchsetzen konnte: die Entwicklung einer solchen Whitebox ist langwieriger und forschungsintensiver und kostet demnach mehr (Tyborski, 2020). Mit den wachsenden Bemühungen um nachvollziehbare KI, werden diese Mehrkosten sich langfristig erübrigen.

Im weiteren Verlauf des Interviews verriet Nadine Hammele auf die Frage, wie man als Entwickler eine KI unter ethischen Aspekten verbessern kann, dass man zum Beispiel auch einen Ethiker mit ins Boot holen könnte (persönliche Kommunikation, 05. Januar 2021). Alternativ können sich Entwickler auch weiterbilden; so stellen viele Kommissionen, Institute (auch das IDE) oder Organisationen Materialien wie Leitfäden oder Checklisten bereit, anhand derer man sich während der Entwicklung einer KI orientieren kann. Im besten Falle evaluiert man das entsprechende Produkt nicht erst zum Schluss, wenn es schon fertig ist, sondern macht sich laut Hammele schon möglichst früh Gedanken:

Eigentlich muss man bei Ethik und KI sogar noch einen Schritt zurück gehen und schon in dem Moment, in dem man sagt, "ich möchte ein neues KI-System entwickeln", sollte sich ein Entwickler fragen: "Dient diese künstliche Intelligenz, die ich schaffe, dem Wohl des Menschen und wo gibt es Gefahren?"

Dieser Ansatz wird als "Ethics by Design" bezeichnet. Weiterhin können diverse Teams, bezogen auf Ethnien und Geschlechter, Biases und Unterrepräsentation in Trainingsdaten entgegenwirken, da hier verschiedene Perspektiven zusammenspielen (Levina, 2020).

In Großunternehmen oder allgemein können auch dritte Instanzen zum Einsatz kommen, die KI prüfen und Strategien erarbeiten, sowie einen Rechtsrahmen aufsetzen. Beispiele hierfür sind etwaige Ethik- und Datenethik-Kommissionen (ZEIT ONLINE, o.D.) aber auch das Pentagon hat, nach einer Konsultationsphase mit Experten, Richtlinien für eine "vertrauenswürdige KI" festgelegt (Tyborski, 2020).

Neben der Verbraucher- und Expertenstudie von Capgemini (2019) hat auch eine Studie vom BVDW, dem Bundesverband digitale Wirtschaft (2019), gezeigt, dass Ethik im Bereich der KI einen hohen Stellenwert für Konsumenten darstellt. So gaben 74 Prozent der befragten Verbraucher an, dass ethische Themen bereits bei der Produktentwicklung eine große oder sehr große Rolle spielen sollte.

Ähnlich viele (72 Prozent) würden Produkte bzw. Services bevorzugen, deren Entscheidungsfindung sich an ethischen und nachvollziehbaren Standards orientiert. So könnte es demnach auch im wirtschaftlichen Interesse von IT-Unternehmen liegen, sich durch mehr Transparenz einen Wettbewerbsvorteil zu sichern.

FAZIT

Mit dem wachsenden Einfluss von künstlicher Intelligenz auf unser aller Leben, ist es nicht weiter überraschend, dass Forderungen nach Transparenz, Fairness und Klärung der Verantwortungsfrage immer lauter werden. Dass diese nicht auf taube Ohren stoßen, zeigen die Bemühungen rund um XAI, ethischen Richtlinien und Aufklärungsarbeit.

Eine Überregulierung auf der anderen Seite könnte dem Innovationspotential von künstlicher Intelligenz in die Quere kommen. Man kann sich das also in gewissem Maße als ein Tauziehen vorstellen, bei dem diese junge Technologie, wie aufgekommene wissenschaftliche Disziplinen zuvor, der Herausforderung begegnet, ihre ethischen Grenzen zu definieren (Levina, 2020). Es bleibt also abzuwarten, wie es weitergeht, aber es steht außer Frage, dass sich politische Entscheidungsträger, die Industrie und nicht zuletzt die Gesellschaft mit ethischen Themen rund um KI über die Jahre hinaus auseinandersetzen werden müssen.

QUELLENANGABE

- Bastian, M. (2020, 11. Juni). *OpenAI: Mächtige Text-KI GPT-3 geht in den Verkauf*. MIXED | News zu VR, AR und KI. <https://mixed.de/openai-maechtige-text-ki-gpt-3-geht-in-den-verkauf/>
- Beiter, R., Doria, J., Gottschaller, S., Kaeber, F., Kegel, J., Leipold, C. (2020). *Fühlt sich das noch gut an? Ein quantitativ-qualitatives Forschungsprojekt zur Akzeptanz der Künstlichen Intelligenz im Alltag*. (S. 6–9) [E-Book].
https://hdms.bsz-bw.de/frontdoor/deliver/index/docId/6556/file/Beiter-et-al_2020_E-Book_Studie-KI-im-Alltag.pdf
- Berg, A. & Dehmel, S. (2020, 28. September). *Künstliche Intelligenz*. [Präsentationsfolien der Studienergebnisse]. Bitkom.
https://www.bitkom.org/sites/default/files/2020-09/bitkom-charts-kunstliche-intelligenz-28-09-2020_final.pdf
- Beuth, P. (2016, 24. März). *Twitter-Nutzer machen Chatbot zur Rassistin*. ZEIT ONLINE.
https://www.zeit.de/digital/internet/2016-03/microsoft-tay-chatbot-twitter-rassistisch?utm_referrer=https%3A%2F%2Fscilogs.spektrum.de%2F
- Brogan, J. (2016, 11. August). *Should a Self-Driving Car Kill Two Jaywalkers or One Law-Abiding Citizen?* Slate Magazine.
<https://slate.com/technology/2016/08/moral-machine-from-mit-poses-self-driving-car-thought-experiments.html>
- Burt, C. (2020, 2. März). *Tech giants pressured to follow Google in removing gender labels from computer vision services*. Biometric Update |.
<https://www.biometricupdate.com/202003/tech-giants-pressured-to-follow-google-in-removing-gender-labels-from-computer-vision-services>
- BVDW. (2019). *BVDW-STUDIE. Künstliche Intelligenz, Smart Home, vernetzte Gesundheit – Ethik in der Digitalen Wirtschaft*. [Präsentationsfolien der Studienergebnisse]. PowerPoint-Präsentation.
https://www.bvdw.org/fileadmin/user_upload/190911_BVDW_Studie_Ethik_DMEXCO.pdf
- Capgemini Research Institute. (2019). *Why addressing ethical questions in AI will benefit organizations*.
<https://www.capgemini.com/de-de/wp-content/uploads/sites/5/2019/07/Studienreport-KI-und-Ethik.pdf>
- Haefliger, M. M. (2020, 29. Januar). *England: Londoner Polizei setzt Kameras zur Gesichtserkennung ein*. Neue Zürcher Zeitung.
<https://www.nzz.ch/international/england-londoner-polizei-setzt-kameras-zur-gesichtserkennung-ein-ld.1536534?reduced=true>

- Joy Buolamwini | TEDxBeaconStreet. (2017, 9. März). *How I'm fighting bias in algorithms* [Video]. TED Talks.
https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms
- Levina, O. (2020, 14. Januar). *Künstliche Intelligenz und Ethik*. Informatik Aktuell.
<https://www.informatik-aktuell.de/betrieb/kuenstliche-intelligenz/kuenstliche-intelligenz-und-ethik.html>
- MIT Media Lab. (o. D.). *Moral Machine*. Moral Machine. Abgerufen am 11. März 2021, von
<https://www.moralmachine.net/>
- Naruniec, J. (2020). High-Resolution Neural Face Swapping for Visual Effects. In L. Helming, C. Schroers & R. M. Weber (Hrsg.), *Computer Graphics Forum* (Ausg. 4, Bd. 39, S. 173–184). John Wiley & Sons Ltd. <https://doi.org/10.1111/cgf.14062>
- Paefgen-Laß, M. (2019, 21. Oktober). *Ethik wird bei KI zum Wettbewerbsfaktor*. Springer Professional.
<https://www.springerprofessional.de/kuenstliche-intelligenz/corporate-social-responsibility/ethik-wird-bei-ki-zum-wettbewerbsfaktor/17022370>
- Podbregar, N. (2020, 28. Mai). *Interview zur Künstlichen Intelligenz: „Schluss mit der Black Box!“*. Wissenschaft.de.
<https://www.wissenschaft.de/technik-digitales/interview-schluss-mit-der-black-box/>
- Schreiner, M. (2020, 8. März). *Bekannter KI-Forscher hört auf: Missbrauchspotenzial zu groß*. MIXED | News zu VR, AR und KI.
<https://mixed.de/bekannter-ki-forscher-hoert-auf-missbrauchspotenzial-zu-gross/>
- Schreiner, M. (2021, 5. Januar). *Künstliche Intelligenz: Was 2020 war und 2021 wird*. MIXED | News zu VR, AR und KI. <https://mixed.de/kuenstliche-intelligenz-was-2020-war-und-2021-wird/>
- Till, U. (2020, 4. November). *Computerprogramm erkennt verräterischen Covid-19-Husten*. SWR Wissen. <https://www.swr.de/wissen/husten-ki-soll-covid-19-erkennen-100.html>
- Tyborski, R. (2020, 10. März). *Blick in die Blackbox: Firmen brauchen einen neuen KI-Ansatz*. Handelsblatt.
<https://www.handelsblatt.com/technik/digitale-revolution/digitale-revolution-blick-in-die-blackbox-warum-unternehmen-einen-neuen-ki-ansatz-nutzen-sollten/25626092.html?ticket=ST-596435-bEjP3e6O7Ns9NghbJcrd-ap3>
- WELT. (2017, 27. Oktober). *Saudi-Arabien vergibt Staatsbürgerschaft an Roboter Sophia*.
<https://www.welt.de/vermischtes/article170106321/Roboter-Sophia-bekommt-saudi-arabische-n-Pass.html>
- ZEIT ONLINE. (o. D.). *Kethik – KI braucht Ethik*.
https://www.zeit.de/angebote/weltderdaten/kuenstliche-intelligenz-und-ethik/index?utm_referrer=https%3A%2F%2Fwww.google.com%2F
- Zweig, K. (2019). *Ein Algorithmus hat kein Taktgefühl: Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können*. (Originalausgabe Aufl., S. 126–267). Heyne Verlag.

ANHANG

Interviewtranskript

Wer spricht?	
Ich	Wir sind heute hier, um über das Thema künstliche Intelligenz und Ethik zu sprechen und dafür habe ich Sie mit mir, Nadine Hammele. Vielleicht stellen Sie sich am besten mal kurz vor, wie Sie mit dem Thema in Verbindung stehen.
N. Hammele	<p>Sehr gerne. Mein Name ist Nadine Hammele, wie gerade schon erwähnt, und ich bin wissenschaftliche Mitarbeiterin im Institut für digitale Ethik an der Hochschule der Medien.</p> <p>Da forsche ich schon seit drei Jahren zum Thema Ethik und künstliche Intelligenz - genauer gesagt sind meine Schwerpunkte automatisiertes Fahren, aber auch ganz allgemein wie man Ethik in technische Entwicklungen mit reinbringt.</p> <p>Neben der wissenschaftlichen Tätigkeit an der Hochschule der Medien schreibe ich auch noch meine Doktorarbeit über künstliche Intelligenz im Film und welche Narrative über künstliche Intelligenz in unserer Gesellschaft im Umlauf sind. Das ist auch eine Form von angewandter Ethik, sich mit den Vorstellungen von künstlicher Intelligenz und auch Ängsten und Chancen im Film zu beschäftigen.</p>
Ich	<p>Super spannend. Dann habe ich auf jeden Fall hier die richtige Referentin am Mikro und würde sagen, dass wir direkt mal mit den Fragen starten.</p> <p>Ich hab ja ein paar vorbereitet und ich glaube, die Frage, mit der man gut ins Thema einsteigen könnte, wäre einfach, was hat künstliche Intelligenz mit Ethik zu tun?</p>
N. Hammele	<p>Künstliche Intelligenz ist vereinfacht gesagt, dass ein IT-System Eigenschaften, die sonst der Mensch hat, sozusagen simuliert.</p> <p>Da künstliche Intelligenz in ganz vielen Bereichen unserer Gesellschaft einfließt, von Bewerbungsverfahren bis hin zu automatisierten Fahrzeugen, oder in medizinischen Verfahren - also wirklich in einen breitgefächerten Kontext vorhanden ist, da stellen sich verschiedene ethische Fragen.</p> <p>Zum Beispiel beim autonomen Fahren das bekannte moralische Dilemma: Wie soll sich ein Fahrzeug entscheiden, wenn es einen Personenschaden gibt? Im medizinischen Bereich kommen auch ethische Fragestellungen auf, oder es kann Probleme geben, wenn zum Beispiel das Datenmaterial nicht vollständig ist, dass Menschen benachteiligt werden können oder vielleicht falsche Behandlungsmethoden vorgeschlagen kriegen.</p> <p>Dadurch, dass eine Maschine, die nicht so viel Weltwissen hat wie der Mensch, in Entscheidungen integriert wird, gibt es einfach sehr viele Dinge zu berücksichtigen und da spielt die Ethik eine große Rolle.</p> <p>Ich bin in den Beispielen schon relativ viel in den Anwendungen drinnen, aber eigentlich muss man bei Ethik und KI sogar noch einen Schritt zurück gehen und schon in dem Moment, in dem man sagt, "ich möchte ein neues KI-System</p>

	<p>entwickeln”, sollte sich ein Entwickler fragen: “Dient diese künstliche Intelligenz, die ich schaffe, dem Wohl des Menschen und wo gibt es Gefahren?”</p> <p>Das heißt “Ethics by Design”, dass man sozusagen ethisch von vorne bis hinten in der Technik-Entwicklung mitdenkt. Deswegen ist die Ethik etwas, das sollte nicht am Ende wie ein Siegel auf eine Entwicklung gesetzt werden, sondern Ethik sollte wirklich ein kompletter Prozess sein. Entwickler und Ethiker sollten quasi Hand in Hand gehen und immer als Team arbeiten.</p>
Ich	<p>Das habe ich auch durch meine Recherche schon mitbekommen, dass man am Anfang des Prozesses sich schon früh Gedanken machen soll, weil wenn die Mühle dann mal läuft, ist es gar nicht mehr so leicht, in diese Blackbox reinzuschauen und nachzuvollziehen, was die da genau macht.</p> <p>Gerade weil wir es jetzt auch von Risiken hatten, wo sehen Sie die Chancen von KI und wo die Risiken?</p>
N. Hammele	<p>Also, die großen Chancen von KI sind auf jeden Fall diese riesigen Datenauswertungen, bei der zum Beispiel eine künstliche Intelligenz verschiedene Hautkrebskrankungen erkennt. Die kann natürlich viel, viel mehr sehen, anhand von Millionen Bildern - also, ich sag mal, viel mehr, als ein Arzt in seinem Leben sehen kann. Anhand von solchen Beispielen sieht man, “okay, die künstliche Intelligenzen können dann wirklich Bilder besser auswerten”, gerade im Bereich der Medizin, und so können vielleicht schneller Zusammenhänge erkannt werden.</p> <p>Dass verschiedene wissenschaftliche Bereiche durch künstliche Intelligenz große Fortschritte machen, das ist auf jeden Fall ein möglicher Anwendungsfall, der eine große Chance darstellt. Gerade auch beim automatisierten Fahren, wenn ein System den Fahrer unterstützt oder warnt und mit dem Fahrer kooperiert, können natürlich auch Unfälle verhindert werden. Das heißt, dass es in solchen Bereichen Chancen gibt, wie auch die Arbeiterleichterung. So haben immer mehr Leute einen Staubsaugroboter zuhause und allgemein lässt sich beobachten, dass sich immer mehr Anwendungen durchsetzen, mit denen unser tägliches Leben erleichtert wird. Daraus ergeben sich viele Vorteile, die künstliche Intelligenz mit sich bringt.</p>
Ich	<p>Sie hatten zudem auch schon ein bisschen angesprochen, welche Risiken und Nachteile sich durch künstliche Intelligenz ergeben könnten. Könnten Sie darauf noch kurz eingehen?</p>
N. Hammele	<p>Die Risiken von künstlicher Intelligenz sind auf jeden Fall im Bereich von Privatsphäre und Datenschutz sehr groß. Das fängt bei einem Staubsaugerroboter an, wobei der ja noch relativ wenig wahrnimmt - kritischer ist es da bei einer Sprachassistenten, wie die Alexa, die dann auch Gespräche mithört. Überall da, wo wir Technologie in unseren Alltag integrieren, sammelt diese Technologie auch oft Daten über uns und oft fehlt es da dann an Transparenz: Was machen die Unternehmen, die diese Daten sammeln, mit diesen? Wie werten sie diese aus und was wird weitergeleitet?</p> <p>Ein weiteres Thema ist die Diskriminierung. Das Problem ist hierbei ganz oft, dass die Datensätze, mit welchen eine künstliche Intelligenz lernt, Daten beinhaltet, die historisch gewachsen oder unvollständig sind.</p>

	<p>Aufgrund dieser verzerrten Daten könnten so bestimmte Personen benachteiligt werden.</p> <p>Ein weiteres Thema ist die Diskriminierung. Das Problem ist hierbei ganz oft, dass die Datensätze, mit welchen eine künstliche Intelligenz lernt, Daten beinhaltet, die historisch gewachsen oder unvollständig sind. Aufgrund dieser verzerrten Daten könnten so bestimmte Personen benachteiligt werden. Also, dass eine künstliche Intelligenz zum Beispiel lernt, dass Menschen bestimmter Ethnien nicht in diesen Führungspositionen bisher waren und dies folglich als Kriterium interpretiert, Personen auszuwählen. Oder dass Frauen benachteiligt werden.</p> <p>Jegliche Form der Diskriminierung ist ein großes Risiko und natürlich - wie Sie vorher auch schon erwähnt hatten -, dass eine künstliche Intelligenz eine Blackbox sein kann. Vor allem bei diesen neuronalen Netzen ist das oft der Fall, dass die Zusammenhänge nicht immer ganz klar sind: Wie kommt jetzt die künstliche Intelligenz zu ihren Entscheidungen? Und immer dann, wenn man so eine Blackbox hat, kann es zu Problemen und Anwendungsfehlern kommen, weil diese einfach sehr schwer kontrollierbar ist und das ist auch so ein Punkt, bei dem man sich überlegen muss: Wie kann man eine künstliche Intelligenz nachvollziehbar, erklärbar und überprüfbar machen?</p>
Ich	<p>Gerade den Datenschutz-Aspekt haben viele ja im Kopf und stehen deshalb personalisierter Werbung und Co. kritischer gegenüber, weil man sich dann fragt, welche Daten von diesem System, hinter dem eine künstliche Intelligenz steht, gesammelt werden.</p> <p>Haben Sie Empfehlungen für den Umgang als Verbraucher mit intelligenten Systemen? Soll man dann sagen, "okay ich nutze es gar nicht" oder gibt es da auch einen Mittelweg?</p>
N. Hammele	<p>Also, als Verbraucher, wenn ich mir eine neue Technologie anschaffen möchte, dann würde ich erstmal vergleichen: Was gibt es für Anbieter? Und oft gibt es auch Firmen, die ja wirklich Wert legen auf Datenschutz; die dann einfach sagen, "unser Produkt ist datensicherer als andere". Und bei vielen Produkten würde ich mich auch fragen, ob ich das wirklich brauche.</p> <p>Ein Beispiel dazu: es gab mal eine smarte Barbie, die ist wieder vom Markt genommen worden, weil Kinder mit der Barbie reden konnten und die ganzen Gespräche wurden ans Unternehmen weitergeleitet bzw. gesammelt. Sogas ist dann natürlich super kritisch. Da muss man sich fragen: Braucht das Kind wirklich eine Sprachassistentz, welche auch mit dem Internet verbunden ist und in Form eines Spielzeugs? Letzteres ist nochmal deutlich sensibler, weil man Spielzeug auf Antrieb Emotionen und Vertrauen schenkt - bei Kindern umso mehr und unreflektierter als bei Erwachsenen.</p> <p>Aso da müsste man sich erstmal fragen: Brauche ich die Technologie? Und darauf achten, ob es Produkte gibt, die Wert auf Datenschutz und Privatsphäre legen.</p> <p>Das ist eigentlich auch die größte Chance, ich hoffe, dass immer mehr Unternehmen das erkennen: Wenn sie Produkte anbieten, die Ethik berücksichtigen, also sogenannte "Ethics by Design"-Produkte, dass das auch am Markt ein Wettbewerbsvorteil ist.</p>

	<p>Es gibt ja immer mehr Menschen, denen ist bewusst, dass Privatsphäre und Datenschutz ein großes, wichtiges Thema ist und dass auch diese Produkte eben mehr nachgefragt und eher gekauft werden als andere.</p> <p>Der nächste Tipp, den ich noch geben kann, bezieht sich auf die Einstellungen. Oft kann man bei Apps oder allgemein im Smartphone Änderungen in diesen vornehmen und das gibt es vielleicht auch bei einem Sprachassistenten-System zum Beispiel. Schwierig ist es natürlich, wenn es kein Aus-Knopf gibt und diese künstliche Intelligenz bzw. das System immer mithört. Das ist dann sehr, sehr schwierig, weil diese Daten, die ausgewertet werden, können sehr genaue Persönlichkeitsprofile von uns abbilden und das ist auch Thema in diesen Reportagen, zum Beispiel von Cambridge Analytica oder der Netflix-Film "Das Dilemma der sozialen Medien". Das zeigt sehr anschaulich, wie Wissen über Personen am Ende Wahlen manipulieren kann. Weil wenn man so viel über Menschen weiß, so viel über ihre Persönlichkeit und Weltanschauungen, dann kann man sehr schnell diese Menschen manipulieren und dann hat man am Ende keine Demokratie mehr, die funktioniert. Also die Auswirkungen können wirklich sehr weitreichend sein.</p>
Ich	<p>Ja, auf jeden Fall. Es gibt ja noch eine andere Seite, die zum Beispiel für mich interessant sein könnte, weil ich gerade auch in den Bereich Machine Learning gehe. Wie sieht es als Entwickler, also mit der Gegenseite aus? Sie hatten schon Ethics by Design genannt; was kann ich tun als Entwickler, um eine KI auch unter moralisch-gesellschaftlichen Aspekten zu verbessern?</p>
N. Hammele	<p>Also, mit der Frage, was der Entwickler in Hinblick auf Ethik und KI tun kann, damit haben wir uns in unserem Forschungsprojekt die letzten eineinhalb Jahre beschäftigt. In dem Forschungsprojekt haben wir jetzt eine "Ethik und Recht"-Hilfesoftware entwickelt. Vereinfacht gesagt, so wie es Steuererklärungs-Hilfesoftware gibt, ist das eine Software, die für Ethik und Recht sensibilisiert.</p> <p>Das ist ein Fragebogen, der vom Antragsteller von Forschungsprojekten, theoretisch aber auch von Entwicklern im Anfangsstadium ihres Projekts herangezogen werden kann. Diese Fragen regen nochmal zur Reflektion an und das sind einmal typische Datenschutz-Fragen, speziell rechtlich, aber auch verschiedene Werte-Fragen; also so etwas wie Diskriminierung, Transparenz und dass man da einfach sensibilisiert wird.</p> <p>Wichtig ist vor allem, dass man von Anfang an die Ethik mit berücksichtigt und nicht erst am Ende eine Reflexion durchführt, wenn das Produkt schon fertig ist. Besser setzt man das früh an und dann macht es total Sinn, neben einer Hilfe-Software, dass man entweder einen Ethiker mit ins Boot holt oder sich selber als Entwickler fortbildet durch Ethik-Fortbildungen oder auch Workshops.</p> <p>Wichtig ist, glaube ich, dass man sich überlegt: Was gibt es für Werte? Welche Werte in unserer Gesellschaft sind wünschenswert und wie kann ich diese Werte auch in diese KI-Entwicklung implementieren?</p>
Ich	<p>Soweit ich mitbekommen habe, ist dieses Bewusstsein oft auch auf Entwickler-Seite da, dass diese selber das Risiko hoch einschätzen oder erkennen, wie ethisch schwierig die eigenen Produkte sind bzw. sein können. Man hat da schon echt viel Verantwortung, wenn man an intelligenten System bastelt.</p> <p>Zum Abschluss mal noch eine Frage, die wird immer viel gestellt, ist natürlich aber</p>

	<p>auch sehr interessant. Wir reden ja aktuell über schwache KI; eine starke KI, die wie wir Menschen auch in vielen Bereichen einsatzfähig ist und ihre Schlüsse ziehen kann - das ist ja noch Zukunftsmusik. Wann glauben Sie denn, wird KI dem Menschen ebenbürtig sein oder gar überholen? Wird letzteres jemals passieren?</p>
<p>N. Hammele</p>	<p>Die Frage, ob eine künstliche Intelligenz dem Menschen ebenbürtig sein kann, ist natürlich schwierig und knifflig und treibt den Menschen auch in zahlreichen Science-Fiction-Filmen um. Das ist ja eines der wiederkehrenden Narrativen: die KI, die uns ebenbürtig ist, in manchen Filmen ist sie sogar besser und moralischer als der Mensch, in anderen Filmen als moralisch schlechte Superintelligenz dargestellt, die uns unterjocht, wie bei Matrix oder Terminator. Ich sag mal, die Fantasien sind sehr groß.</p> <p>Ich selber bin kein Zukunftsforscher - ich beschäftige mich ja mit Ethik, Filmen und Narratologie, aber aus meiner Sicht der Dinge werden künstliche Intelligenzen sehr stark vermenschlicht. Und da ist das Problem mit dieser Vermenschlichung, die schon im Titel steckt, weil wir von Intelligenz reden und denken, da ist ein intelligentes Wesen. Aber eben diese Vermenschlichung, die in jeglichen Vorstellungen in unserer Gesellschaft steckt, die führt dazu, dass wir denken, die KI wird uns bald ebenbürtig sein. Ich selber glaube aber, dass unser menschliches Gehirn ganz anders funktioniert als diese IT-Systeme und auch wenn man von neuronalen Netzen spricht, ist das nur eine Vereinfachung von diesem System und wie es funktioniert, aber es ist ja nicht wirklich ein neuronales Netz wie unser Gehirn es hat.</p> <p>Deswegen glaube ich, auch wenn KI sehr große Zusammenhänge erkennen kann und auch Roboter Handlungen ausführen, die Intelligenz simulieren, hat die künstliche Intelligenz am Ende keine Intelligenz. Also sie simuliert nur Intelligenz, sie besitzt die Intelligenz nicht. Und das große Problem, das die künstliche Intelligenz hat, um mit dem Menschen auf einer Stufe zu sein, ist das Weltwissen; also dass eine künstliche Intelligenz nicht so viel Wissen hat, wie der Mensch, sondern ich glaube, dass künstliche Intelligenz noch sehr lange Bereichsintelligenzen bleiben. In ihrem speziellen Gebiet sind sie sicherlich auch besser als der Mensch, wie in der Objekterkennung und -auswertung, aber meine persönliche Einschätzung ist, dass wir, also unsere Generation, es nicht erleben wird, dass die künstlichen Intelligenzen uns wirklich ebenbürtig sind. Ich bezweifle generell, dass das überhaupt möglich ist. Aus ethischer und philosophischer Sicht: ein Mensch besitzt so viel Tiefe, Werte und Psychologie und das alles zu simulieren und kopieren.. ich glaube, dafür haben wir den Menschen noch zu wenig verstanden, um da etwas zu schaffen, was genau so ist wie wir.</p> <p>Aber eine spannende Frage, ja. Also ich glaube, darüber wird noch viel diskutiert und wenn man in die Literatur schaut, gibt es auch die Technik-Pessimisten, die sagen, dass es niemals möglich ist - wenn dann schafft die KI nur Probleme und es gibt natürlich auch die Optimisten. Die sagen wirklich, "der Mensch ist eigentlich nur die Vorstufe für eine KI, die uns später überflügelt und uns überholt". Viele finden das schön, viele finden das beängstigend - es gibt dazu echt die verschiedensten Vorstellungen in der Gesellschaft.</p>