

Künstliche Intelligenz: Perspektiven des Kinder- und Jugendmedienschutzes

Schwerpunkt 

Künstliche „Intelligenz“ – Einführung in eine Schlüsseltechnologie

Dominique Facciorusso, Deborah Woldemichael

Künstliche Intelligenz, kurz KI, bestimmt viele Bereiche unseres Lebens. Es ist eine aufregende und höchst dynamische Technologie, bei der wir gerade mal am Beginn der Entwicklung stehen. Und es steht fest: diese Entwicklung ist unumkehrbar. Besonders die jüngsten Fortschritte im Bereich der generativen KI haben gezeigt, welche Wucht in den neuen Möglichkeiten der Mensch-Maschine-Interaktion steckt. Inhalte erzeugende Tools wie ChatGPT, DALL-E und Co. machen künstliche Intelligenz plötzlich massentauglich erlebbar. Damit rückte das Thema „KI“, das zuvor eher in Fachkreisen diskutiert wurde, im Eiltempo in das öffentliche Bewusstsein.

Neben den Potenzialen und Erwartungen, die mit dieser Technologie verbunden sind, findet eine massive Verstärkung bereits vorhandener Gefährdungen und Risiken im digitalen Raum statt. Kein Wunder also, dass sich auch die gesellschaftliche Debatte zusehends darüber entfacht, wo und unter welchen Bedingungen die Reise dieser Schlüsseltechnologie hingeht bzw. hingehen sollte. Flankiert wird die Auseinandersetzung vor allem durch zwei gegensätzliche Lager. Die einen befürchten durch KI-Technologien einen unwiderruflichen Kontrollverlust (Stichwort: Maschinenübernahme) und die anderen sehen in ihr die Lösung ubiquitärer Probleme

(Stichwort: Solutionismus). Diese Dichotomie sorgt aber dafür, dass eine lösungsorientierte Auseinandersetzung mit den jetzt schon bestehenden Herausforderungen und Risiken durch KI erschwert wird. Denn Mystifizierung schafft Distanz.

KI-Technologien haben kein Bewusstsein und sind per se weder „gut“ noch „böse“. Es sind Werkzeuge, die von Menschen gemacht, genutzt und durch Nutzerinteraktion geformt werden. Genau wie bei anderen technischen Errungenschaften kommt es auch hier darauf an, was wir als Gesamtgesellschaft damit machen. Dafür müssen wir zunächst grundsätzlich verstehen, was es mit dieser Technologie auf sich hat. Erst dann können wir mit KI-Technologien souverän und kritisch umgehen und uns an gesellschaftlichen Debatten und Kontroversen kompetent und gestalterisch beteiligen.

Das Streben nach einer künstlichen Kopie des Menschen ist nicht neu

Seit der Antike erschaffen wir Menschen Maschinen, die uns körperlich entlasten oder gar übertreffen. Denn unsere physische Kraft ist, obgleich individuell unterschiedlich ausgeprägt, begrenzt. Nicht anders

verhält es sich mit der Idee, etwas zu erschaffen, das uns in unserer Denkleistung unterstützt bzw. uns diese abnimmt. Denn auch hier erfahren wir Anstrengungen und Limitationen hinsichtlich unserer Geschwindigkeit bei der Informationsverarbeitung und Speicherkapazität unseres Gehirns.

Lange vor der Erfindung des digitalen Computers entstand die Idee von intelligenten selbstständigen Systemen bzw. Geschöpfen. Beispiele sind die Kreaturen Homunculus oder der aus Lehm geformte Golem, die mechanischen Apparate, die Erzählung Frankensteins bis hin zu den ersten Entwürfen einer mechanischen Rechenmaschine. Ab der ersten Hälfte des 20. Jahrhunderts trat vermehrt das Konzept künstlich intelligenter Roboter in der Science-Fiction auf. Beispiele sind der humanoide Roboter in Metropolis oder die Robotergeschichten von Isaac Asimov.

Die folgende Generation an Wissenschaftlerinnen und Wissenschaftlern war kulturell von diesen Visionen geprägt. In den 1950er Jahren arbeiteten sie daran, eine Maschine zu entwickeln, die menschlich lernen und denken können sollte. Einer von ihnen war der britische Mathematiker Alan Turing. In seinem renommierten Aufsatz „Computing Machinery and Intelligence“ (1950) beschreibt er, wie sich intelligente Maschinen bauen und hinsichtlich ihrer Intelligenz testen lassen (Turing-Test). Turing legte damit den Grundstein für die spätere KI-Forschung und gilt als einer der einflussreichsten Visionäre und Theoretiker (Anyoha 2017).

Eine Konferenz am Dartmouth College (USA) im Jahr 1955 markiert einen weiteren Meilenstein der Forschung. Dort vertraten Wissenschaftlerinnen und Wissenschaftler die Ansicht, dass sich die Eigenschaften menschlicher Intelligenz so beschreiben lassen, dass Maschinen sie lesen und ausführen können. Die Konferenz gilt als Geburtsstunde des Begriffs „künstliche Intelligenz“ (Original: Artificial Intelligence) sowie der KI als akademisches Fachgebiet (Anyoha 2017).

In den 1960er Jahren ließen sich erste Erfolge hinsichtlich lernfähiger Maschinen sowie Programmen verzeichnen, die die Kommunikation zwischen Mensch und Maschine in natürlicher Sprache ermöglichten. Der Mathematiker Joseph Weizenbaum entwickelte ein Sprachprogramm namens ELIZA. In der Rolle eines Psychotherapeuten reagierte das Programm auf Schlüsselworte und konnte einfache Therapiesprache durchführen. ELIZA gilt als frühe Vorläuferin des modernen Chatbots (Anyoha 2017). Weizenbaum stand übrigens der Rolle des Computers äußerst kritisch gegenüber (vgl. Brammer 2023).

Seither hat sich die KI-Forschung im Hinblick auf Hard- und Software mit einigen Unterbrechungen stetig weiterentwickelt. Zu Beginn waren Speicherkapazität und Rechenleistung von Computern noch stark limitiert und die Datenverarbeitung sehr teuer. Nur wenige Unternehmen, renommierte Universitäten und staatliche Stellen konnten sich diese kostspielige Forschung leisten. Viele Jahrzehnte später konnten durch technologische Errungenschaften sowie die Etablierung des World Wide Web und des Smartphones massive Mengen an Daten übertragen und gespeichert werden. Diese Entwicklungen lieferten den Rohstoff für Big Data und sorgten in Kombination mit viel höheren Rechenkapazitäten für einen enormen Aufschwung in der KI-Forschung.

Künstliche „Intelligenz“ und das Problem mit den Definitionen

Wenn man heute über KI spricht, werden die Begriffe „Algorithmus“ und „Künstliche Intelligenz“ oft synonym gebraucht. Jedoch gibt es Unterschiede:

Algorithmen sind mathematische Schritt-für-Schritt-Anleitungen, die es einem Computer-System ermöglichen, eine vorher klar definierte Aufgabe zu erledigen, wenn alle dafür notwendigen Daten zur Verfügung stehen. Sie funktionieren wie eine Art „Backrezept“ und geben dem Computer vor, welche Daten (Zutaten) nach vorher festgelegten mathematischen Regeln (Rezeptur) kombiniert und miteinander verarbeitet werden müssen, sodass am Ende ein spezifisches Ergebnis (Kuchen) herauskommt. Damit der Computer solche mathematischen Anweisungen lesen und verarbeiten kann, müssen sie auf eine bestimmte Art beschrieben werden. Dies geschieht mithilfe der Software.

Algorithmen bilden die grundlegenden Bausteine von KI-Systemen. Jedoch sind nicht alle Algorithmen für den Einsatz in KI-Systeme geeignet, da sie hierfür über bestimmte Eigenschaften verfügen müssen. Sie sollten beispielsweise in der Lage sein, mit großen Datenmengen umzugehen und dort Muster und Zusammenhänge erkennen. Auf Grundlage dieser Erkenntnisse sollten sie Entscheidungen treffen, aus Erfahrungen lernen und sich ggf. anpassen können. Nicht alle Algorithmen sind dazu im Stande.

Was also versteht man unter „Künstlicher Intelligenz“? Es gibt zwar keine allgemeingültige Definition, aber eine gemeinsame Idee, die sich aus der Etymologie des Begriffs ableiten lässt. „Künstlich“ meint, dass es sich um ein vom Menschen künstlich hergestelltes System, eine Maschine, handelt. Mit

dieser versucht man intelligente Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern.

Was demgegenüber als „echte“ Intelligenz gilt, ist selbst in Fachkreisen nicht eindeutig geklärt. Allgemein versteht man darunter Fähigkeiten wie z. B. Probleme lösen, logisch denken oder sich in neuen Situationen anpassen zu können. Doch es gibt Aufgaben, die Maschinen besser bzw. schneller als Menschen bewältigen können, ohne dass dafür zwingend hohe kognitive Kompetenzen erforderlich sind. Ein Taschenrechner ist dem Menschen z. B. in seiner Rechenfähigkeit weit überlegen, weshalb ihn aber niemand als „intelligent“ bezeichnen würde. Und auch wenn eine KI durch semantische Textanalysen Fragen korrekt beantwortet, so kann sie die tieferliegende Bedeutung der Frage oder die der Antwort nicht „verstehen“. Mithilfe von KI lassen sich schlicht Daten zueinander in Beziehung setzen, um darauf basierend Entscheidungen oder Antworten zu berechnen. Einige Stimmen halten den Begriff „Künstliche Intelligenz“ daher für irreführend, weil er zu falschen Erwartungen führt und Ängste schürt.¹

„Schwache“ KIs haben starken Einfluss auf unsere (digitale) Lebenswelt

Grundsätzlich gibt es nicht „die“ eine KI. Hinter dem Begriff „Künstliche Intelligenz“ verbergen sich unterschiedlichste Methoden und Technologien. KI-Systeme werden für verschiedene Aufgaben und Anforderungen programmiert. Damit können die Funktionen und Ansätze ganz unterschiedlich sein. Mit Blick auf die Fähigkeiten von KI-Technologien, wird oft kategorisiert, wie „mensenähnlich“ bzw. „intelligent“ solche Systeme sind bzw. sein könn(t)en. Daran angelehnt lassen sich vor allem zwei Arten von KI unterscheiden:

Starke KI

Bei der sogenannten „starken“ KI versucht man Computersysteme zu bauen, die die intellektuellen Fähigkeiten des Menschen nachahmen oder diese sogar übertreffen können. Eine solche KI wäre in der Lage, verschiedene komplexe Aufgaben in multiplen Anwendungsbereichen zu verstehen und zu bewältigen. Diese Art der KI existiert bisher nicht. Sie ist

eine Zukunftsvision und momentan nicht Schwerpunkt der Forschung. Es gibt in Fachkreisen auch die Diskussion, ob so ein System technisch überhaupt realisierbar ist.

Schwache KI

Der aktuelle Fokus in der KI-Entwicklung liegt vor allem auf der sogenannten „schwachen“ KI. Also auf Systemen, die für einen klar definierten Aufgabenbereich konzipiert sind, um dort eine spezifische Aufgabe zu lösen, zum Beispiel Bilder zu erkennen oder Texte zu übersetzen. Schwache KI-Systeme besitzen keine Intelligenz und kein umfassendes Verständnis von der Welt. Dennoch können diese Systeme Erstaunliches leisten. Denn aufgrund der hohen Rechenleistung, der Verarbeitung sehr großer Datenmengen und ausgeklügelter Funktionen können KI-Systeme in vielen Bereichen schneller zu Lösungen kommen als Menschen. Mit schwachen KIs kommen wir täglich in Berührung, angefangen bei der Auto-korrektur, personalisierten Empfehlungssystemen oder Übersetzungsprogrammen.

Wenn wir von „KI“ sprechen, ist maschinelles Lernen gemeint

Wenn wir heute von künstlicher Intelligenz sprechen, meinen wir meistens „lernende“ Systeme bzw. das sogenannte „maschinelle Lernen“ (Machine Learning). Der Ansatz des maschinellen Lernens ist ein wichtiger Teilbereich der Künstlichen Intelligenz und zählt aktuell zu der am häufigsten verwendeten KI-Methode. Dabei lernt ein Computersystem durch Beispiele, indem es Muster und Zusammenhänge in Daten erkennt und dies in einem statistischen Modell festhält. Als Ergebnis erhält man eine Funktion, die eine Eingabe entgegennimmt, um eine Ausgabe zu generieren. Man unterscheidet hierbei grundsätzlich zwei Lernansätze: Überwachtes Lernen (Supervised Learning) und unüberwachtes Lernen (Unsupervised Learning). Auf den dritten Ansatz, das sogenannte bestärkende Lernen (Reinforcement Learning), wird hier nicht näher eingegangen.

Zur Funktionsweise der Lernprozesse

Phase 1 beim überwachten Lernen:

Das Modell wird trainiert und bewertet

Ein Computersystem lernt durch sehr viele Beispiele und menschliches Feedback, welche Eingabe zu wel-

¹ Der Wissenschaftler J. Geuter spricht z. B. von „leistungsfähigen Statistiksystemen“ und die Informatikerin L. Pooyan-Weihls von einer „Erweiterung der menschlichen Intelligenz“. Auch Informatikerin S. Spiekermann problematisiert den Vergleich von menschlicher und Künstlicher Intelligenz und der Mathematiker und Philosoph R. Seising beschreibt im Buch „Es denkt nicht“ warum Menschen den KI-Systemen Intelligenz zuschreiben.

cher Ausgabe gehört. Nehmen wir an, wir möchten ein System trainieren, das zuverlässig Hunde auf Bildern erkennt. Die Eingabe wäre also das Bild eines Hundes und die Ausgabe wäre die Feststellung „Das ist ein Hund“. Dafür benötigt das System Trainingsdaten, also eine sehr große Menge an Hundebildern sowie Fotos von anderen Tierarten.

Anhand dieser Trainingsdaten leitet das System relevante Muster und Regeln ab, die in diesen Daten vorliegen. Zum Beispiel: Hunde haben einen Schwanz und ein Fell, es gibt sie in diversen Farben und Größen. Am Ende entsteht auf dieser Basis ein statistisches Modell, also eine mathematische Gleichung. Sie enthält alle relevanten Muster, Zusammenhänge und Entscheidungsregeln. Entwicklerinnen und Entwickler korrigieren Ungenauigkeiten und Fehler im Modell und bewerten die Treffsicherheit.

Phase 2 beim überwachten Lernen:
Auf Basis des Modells werden Prognosen erstellt und Entscheidungen abgeleitet

Nach einem erfolgreich abgeschlossenen Lernprozess kann das Modell verwendet werden, um neue Daten zu verarbeiten und ähnliche Entscheidungen und Vorhersagen zu treffen. Manche Modelle sind dynamisch und können über die Testphase hinaus weiterlernen. Sie ziehen aus den neuen Daten Rückschlüsse und verbessern ihre Leistung und Funktionen. Aber es gibt auch statische Modelle, die sich nicht anpassen. Sie kommen zum Einsatz, wenn sich Daten nur sehr langsam oder gar nicht verändern.

Wenn man dem Computersystem die besagten Hundefotos zeigt und ihm vorgibt, dass es sich dabei um Hunde handelt, dann „labelt“ man die Daten. Beim unüberwachten Lernen hingegen wären die Daten ungelabelt. Man würde dem Computersystem die Eingabe-Ausgabe-Paare nicht vorgeben, sondern anweisen, dass es nach verborgenen Strukturen und Mustern in den Daten suchen soll. Ein Ergebnis könnte dann z. B. sein, dass es drei Kategorien gibt, die einmal Vögel, Katzen und Hunde zeigen. Oder auf unseren Online-Alltag übertragen: Die KIs von Streaming-Diensten und E-Commerce-Plattformen würden anhand unseres Nutzungsverhaltens Muster bzw. Ähnlichkeiten zu anderen Nutzerdaten erkennen und uns darauf basierend Inhalte empfehlen.

Wenn sehr komplexe Muster und Abhängigkeiten in den Daten erfasst werden sollen, kommt das sogenannte Deep Learning (tiefes Lernen) zum Einsatz – eine Methode des maschinellen Lernens. Hierbei wird eine neuronale Struktur erzeugt, um Infor-

mationen besser verarbeiten bzw. exakter einordnen zu können. Für diese Methode benötigt man große Rechenleistung und Datenmengen. Die grundlegende Idee ist inspiriert vom System der Nervenzellen des menschlichen Gehirns.

KI-Generatoren als Booster der öffentlichen Wahrnehmung von KI

Bei der generativen KI werden Modelle darauf trainiert, „neue“ Inhalte synthetisch zu erzeugen. Das können etwa Texte in menschenähnlicher Sprache, realistisch anmutende Bilder und Videos oder auch Audioinhalte wie Stimmen oder Musik sein. Diese Modelle basieren auf dem Ansatz des maschinellen Lernens bzw. Deep Learnings und greifen für ihren Output auf vorhandene Trainingsdaten zurück.

Die jüngsten Fortschritte in diesem Bereich haben dafür gesorgt, dass für Nutzende neue Möglichkeiten entstanden sind, mit Maschinen zu interagieren und Inhalte zu verarbeiten. Zudem nehmen viele, die bislang wenig wissenschaftliche Berührungspunkte mit KI hatten, durch ChatGPT das Thema „Künstliche Intelligenz“ nun bewusster wahr. Manche nennen generative KI daher einen Gamechanger, andere halten die Technologie für einen Hype. Klar ist: Seit der Veröffentlichung von ChatGPT hat noch nie eine Technologie in unserer Geschichte so schnellwachsende Nutzendenzahlen erzielt. Schon nach zwei Monaten haben circa 100 Millionen Menschen pro Monat das Tool aktiv genutzt (o. A. 2023). Klar ist auch, dass uns generative KI-Tools neue Anwendungsmöglichkeiten bieten – sowohl mit Blick auf die Chancen und Potenziale, aber eben auch hinsichtlich der Gefahren.

Deepfakes und Desinformation fordern unsere Urteilsfähigkeit heraus

KI-Generatoren stellen uns vor ernstzunehmende und altbekannte Risiken. Altbekannt deshalb, weil es Phänomene wie digitale (sexualisierte) Gewalt, Betrugsmaschen oder Desinformation schon lange vor KI-Tools gab. Durch generative KI hat sich jedoch die Quantität und Qualität der manipulativen Inhalte erhöht.

Generative Chatbots führen menschlich wirkende Konversationen. Dadurch wird es immer schwieriger zu erkennen, ob man mit echten Menschen oder mit einem KI-System kommuniziert. Manche Dienste, wie z. B. Snapchats Chatbot „My AI“, weisen ihre Nut-

zerinnen und Nutzer nur einmal vor der ersten Nutzung darauf hin. Fraglich ist, ob diese einmalige Information ausreicht, wenn Maschine und Mensch dauerhaft miteinander interagieren. Wie können Dienste künftig eine transparente und sichere Online-Kommunikation für Nutzende sicherstellen?

In Social Media und Messengern erschweren zudem Deepfakes und Filter, dass wir die Identität unseres Online-Gegenübers sicher feststellen können, etwa mit Blick auf Alter und Geschlecht der Person. Durch KI lassen sich leicht auch Stimmen künstlich generieren, die täuschend echt klingen. Das fordert uns zunehmend heraus, den Wahrheitsgehalt einer Audio-nachricht richtig einzuschätzen. Wie können sich unter solchen erschwerten Voraussetzungen, Menschen vor kriminellen oder sexuellen Absichten schützen?

Algorithmen beeinflussen unsere Wahrnehmung von Inhalten im Netz und damit potenziell auch die Grundlage unserer Meinungsbildung. Mithilfe generativer KI-Tools können in kürzester Zeit massenhaft gefälschte Nachrichtenartikel und Social-Media-Beiträge erstellt werden. Diese können extrem manipulativ und irreführend sein. Social Bots, die in sozialen Netzwerken automatisiert Reichweite generieren, können dazu genutzt werden diese Inhalte zu verbreiten. Die Gefahr ist, solche Inhalte in den öffentlichen Wahrnehmungsradar zu schieben und sichtbar zu machen. So kann ggf. die öffentliche Stimmungslage beeinflusst und gesellschaftliche Kippmomente verstärkt werden. Etwa indem Popularität und Unterstützung für ein Thema oder eine Person vorgetäuscht oder bestimmte Menschengruppen gezielt angegriffen werden.

Deepfakes haben auch die Qualität von Desinformationen verändert und erschweren es, Inhalte auf ihre Glaubwürdigkeit zu prüfen. Infolge dieser Entwicklung schwinden das Vertrauen und die Orientierung, welchen Informationen man noch glauben kann und welchen nicht. Da generative KI zunehmend in der Berichterstattung eingesetzt wird, stellt sich auch hier die Frage der Überprüfbarkeit. Woran erkennen wir, ob ein Artikel von einem Menschen oder einer KI geschrieben wurde? Um die Urteilsfähigkeit der Menschen zu unterstützen, ist vor allem Transparenz erforderlich, etwa durch verpflichtende Kennzeichnung von KI-generierten Inhalten.

Diskriminierung hängt von Qualität der Trainingsdaten ab

Problematisch ist, wenn KI-Systeme aufgrund ihrer Trainingsdaten Vorurteile und Rassismus verbreiten. Damit stellt sich die Frage, inwieweit wir dies als

Nutzende oder Betroffene von KI-basierten Entscheidungen erkennen können, um bei Ungleichbehandlung zu widersprechen. Aber auch, wie sich eine derartige Diskriminierung grundsätzlich verhindern lässt (Gapski/Oberle & Staufer 2017). Denn der Schutz vor Diskriminierung ist ein grundlegendes Menschenrecht.

Daher gilt es, bei der Entwicklung von Algorithmen Diversität und Repräsentativität im Trainingsdaten-Design zu berücksichtigen. Enthalten Trainingsdaten eine verzerrte Wiedergabe der Realität, entsteht ein verzerrtes Modell (Bias). Rassistische, homophobe oder sexistische Inhalte in den Daten werden im KI-System abgebildet. Bei der Verarbeitung neuer Daten würde das Modell auf Basis dieser erlernten Diskriminierungen Entscheidungen treffen.

Um Menschen vor Benachteiligungen zu schützen, sollten die Trainingsdaten überprüfbar sein und sich divers und repräsentativ zusammensetzen. Eine regelmäßige Qualitäts- und Fairnesskontrolle ist nach der Informatikerin Katharina Zweig (2023) bei nachprüfbareren Entscheidungen gut implementierbar. Damit ließen sich Probleme bei der Entscheidungsqualität schneller erkennen.

Wir brauchen Transparenz für besseren Datenschutz

KI-Tools verarbeiten große Datenmengen. Darauf basierend generieren sie neue Inhalte oder leiten Entscheidungen ab, die unser Leben maßgeblich beeinflussen können. Meistens ist dabei nicht klar, welche Daten genau als relevanter Input für Trainingsdaten genutzt wurden und zu welchem Zweck. Ist die verwendete Datenbasis intransparent, gefährdet dies auch unser Grundrecht auf den Schutz personenbezogener Daten.

Aus den Erfahrungen der Datenschutz-Regulierung haben wir gelernt, dass wir uns nicht auf die Selbstregulierung der Konzerne verlassen sollten, wenn es um den Datenschutz geht. Die DSGVO liefert eine gute Grundlage, was die Transparenz über verwendete Daten und Widerspruchsmöglichkeiten angeht. Diese Bestimmungen könnte man auf KI-Systeme anwenden und entsprechend anpassen. Die Nachprüfbarkeit des Entscheidungsprozesses ist dabei Voraussetzung sowie Herausforderung.

Wir brauchen ethische Regeln und Werte für KI

Der Zugang zu generativen GPT-Modellen hat viele strukturelle und moralische Fragen aufgeworfen, die

aktuell nicht ausreichend geklärt sind. Expertinnen und Experten mahnten daher, die Gesellschaft sei auf diese disruptiven Entwicklungen nicht vorbereitet. Vorgeschlagen wurde ein KI-Moratorium, um sich zu positionieren und um gemeinsam Regeln und Grenzen im Umgang mit KI-Systemen zu formulieren (Wagner 2023a). Im medizinischen Bereich gibt es z. B. schon klare Regeln zum Umgang mit KI. Kaum reguliert ist jedoch das Feld der „Plattform-Ökonomie“. Gerade im Bereich der Social-Media-Plattformen braucht es aber schnellstmöglich verbindliche Regelungen und transparente Richtlinien, wie etwa eine verpflichtende Kennzeichnung für KI-generierte Inhalte (Wagner 2023b).

Der AI Act der EU möchte einen verantwortungsvollen Umgang mit KI sicherstellen und verbindliche Regeln und Ethik-Standards schaffen. Dabei wird ein risikobasierter Ansatz verfolgt, indem bewertet wird, ob KI-Anwendungen ein unannehmbares, hohes, geringes oder minimales Risiko darstellen. Verletzen KI-Systeme Werte der Union, z. B. Grundrechte, fallen sie in die risikoreichste Kategorie und sollen verboten werden. Auch Techniken, die Personen unterschwellig manipulieren und damit physischen oder psychischen Schaden verursachen können, sind demnach verboten (Wagner 2023b). Diese Regelung bietet eine erste Grundlage. Entscheidend werden jedoch die konkrete, nationale Ausgestaltung und Anwendbarkeit sein.

Wichtig ist zudem der Aspekt der Teilhabe. Viele KI-Systeme sind leicht zugänglich. Meist genügt eine einfache Registrierung. Konzerne wie Microsoft und Google integrieren KI-Technologien bereits in ihre Produkte und Plattformen. Kinder und Jugendliche kommen also zwangsläufig mit KI-Inhalten in Kontakt. KI-Systeme sollten daher den Schutz von Minderjährigen und anderen schutzbedürftigen Gruppen, z. B. Menschen mit Behinderung, stets berücksichtigen.

Ausblick

KI hat das Potenzial, unser Leben in vielerlei Hinsicht zu verbessern. Jedoch muss sichergestellt werden, dass alle davon profitieren und nicht nur die Akteurinnen und Akteure, die die Technologie entwickeln, kontrollieren oder finanzieren (Stichwort: Chancengleichheit). Als Nutzende und Betroffene von KI-Anwendungen haben wir das Recht auf faire Entscheidungen und Transparenz über die Prozesse. Hierfür benötigen wir ein besseres Verständnis darüber, auf welcher Grundlage ein KI-System eine Entscheidung getroffen hat. Erst dann können wir kri-

tisch die Qualität der Outputs hinterfragen und bewerten und bei Fehlern oder Ungerechtigkeit widersprechen. Eine wichtige Voraussetzung dafür ist, dass algorithmische Entscheidungsprozesse nachvollziehbar und damit überprüfbar sind.

Grundsätzlich müssen wir uns bei der Nutzung von KI immer fragen, warum wir die Technologie eigentlich nutzen und kritisch reflektieren, ob die berechnete Entscheidung automatisch die bessere Wahl ist. Zudem muss geklärt werden, wer die Verantwortung für KI-Entscheidungen trägt und in welchen Fällen ein Mensch entscheiden soll. Für diese Abwägung ist entscheidend, besonders die Risiken im Blick zu haben, die man mit der Nutzung von KI in Kauf nimmt.

Vor diesem Hintergrund brauchen wir eine zukunftsgerichtete, gesellschaftliche Debatte darüber, wann wir welche Form von automatisierten Entscheidungssystemen nutzen wollen und wie die Nutzung abgesichert werden kann (Zweig 2023).

Literatur

- Anyoha, Rockwell (2017): The History of Artificial Intelligence. Can Machines Think? Harvard. Verfügbar unter: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>.
- Brammer, Robert (2023): Zum 100. Geburtstag von Joseph Weizenbaum: Die Macht der Computer. Verfügbar unter: <https://www.deutschlandfunkkultur.de/weizenbaum-100.html>.
- Deutscher Ethikrat (2023): Stellungnahme: Mensch und Maschine – Herausforderungen durch KI. Verfügbar unter: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
- Dia, Mohammed (o. A.): Eine kurze Geschichte der KI. KI von der Antike bis heute, eine Reise durch Höhen und Tiefen. Verfügbar unter: <https://www.thats-ai.org/de-CH/units/eine-kurze-geschichte-der-ki>.
- European Union (2023): Artificial intelligence act. Verfügbar unter: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf).
- Gapski, Harald/Oberle, Monika und Staufer, Walter (2017.): Medienkompetenz. Herausforderung für Politik, politische Bildung und Medienbildung. Bonn: Bundeszentrale für politische Bildung.
- Geuter, Jürgen (2018): Nein, Ethik kann man nicht programmieren. Verfügbar unter: <https://www.zeit.de/digital/int.ernet/2018-11/digitalisierung-mythen-kuenstliche-intelligenz-ethik-juergen-geuter/komplettansicht>.
- Ludwig, Werner (2023): Künstliche Intelligenz: „Maschinen sind nicht intelligent“. Verfügbar unter: <https://www.stuttgarter-zeitung.de/inhalt.kuenstliche-intelligenz-maschinen-sind-nicht-intelligent.84a24ac5-8e29-4a5b-a7ca-55c0b6b36c21.html>.
- o. A. (2023): KI-Software ChatGPT knackt Rekord beim Benutzerwachstum. Verfügbar unter: <https://www.deutschlandfunkkultur.de/ki-software-chatgpt-knackt-rekord-beim-benutzerwachstum-100.html>.
- Rusch, Lina (2023): Der „Ketzer der Informatik“: Was ChatGPT und Internet mit Joseph Weizenbaum zu tun haben. Verfügbar unter: <https://www.tagesspiegel.de/wissen/der-ketzer-der-informatik-was-chatgpt-und-internet-mit-joseph-weizenbaum-zu-tun-haben-9137928.html>.

- Spiekermann, Sarah (2021): Zum Unterschied zwischen künstlicher und menschlicher Intelligenz und den ethischen Implikationen der Verwechslung. In Mainzer, Klaus (Hrsg.): Philosophisches Handbuch Künstliche Intelligenz. Wiesbaden: Springer VS, S. 1-20.
- Wagner, Lukas (2023a): Forderung nach Forschungsstopp: Wie gefährlich kann KI für die Welt werden? Verfügbar unter: <https://www.zdf.de/nachrichten/digitales/gedahren-kuenstliche-intelligenz-entwicklung-stopp-gpt-100.html>
- Wagner, Lukas (2023b): Ethischer Umgang mit der Technik: Wie KI in Zukunft reguliert werden soll. Verfügbar unter: <https://www.zdf.de/nachrichten/politik/ki-regeln-gesetz-ai-act-eu-ethik-experten-100.html>.
- Zweig, Katharina (2019): Ein Algorithmus hat kein Taktgefühl: Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können. München: Heyne Verlag.
- Zweig, Katharina (2023): Die KI war's!: Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz. München: Heyne Verlag.

Zu den Personen



Foto: © Susanne Lencinas

Dominique Facciorusso (Kommunikationswissenschaftlerin, M.A.) ist Referentin für Medienkompetenz bei der EU-Initiative klicksafe.



Foto: © privat

Deborah Woldemichael (Dipl. Pädagogin) ist Leiterin der EU-Initiative klicksafe und Leiterin des Bereichs Medienkompetenz bei der Medienanstalt RLP.