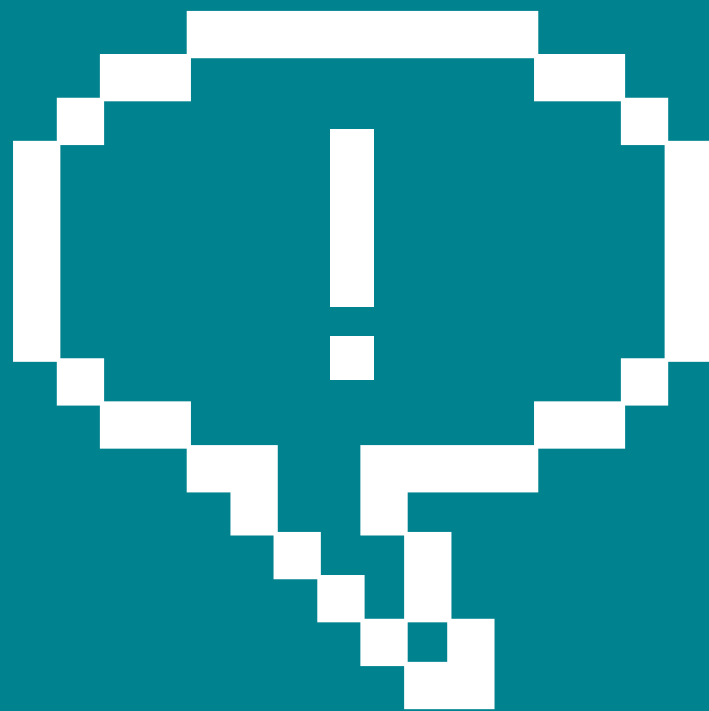


Risikokommunikation zur Künstlichen Intelligenz

RI 



KOMPENDIUM

Mit Kommunikationstipps von Forschenden für Forschende

Eine Initiative des Bundesministeriums
für Bildung und Forschung

Wissenschaftsjahr 2019

**KÜNSTLICHE
INTELLIGENZ**

WISSENSCHAFTSJAHR 2019 – KÜNSTLICHE INTELLIGENZ

Künstliche Intelligenz ist das Thema des Wissenschaftsjahres 2019. Systeme und Anwendungen, die auf Künstlicher Intelligenz basieren, sind schon heute vielfach Bestandteil unseres Lebens: Industrieroboter, die schwere oder eintönige Arbeiten übernehmen oder smarte Computer, die in kurzer Zeit riesige Datenmengen verarbeiten können – und damit für Wissenschaft und Forschung unverzichtbar sind. Ganz abgesehen von virtuellen Assistenzsystemen, die zu unseren alltäglichen Begleitern geworden sind. Digitalisierung und Automatisierung werden in Zukunft weiter fortschreiten. Welche Chancen gehen damit einher? Und welchen Herausforderungen müssen wir uns stellen? Welche Auswirkungen hat diese Entwicklung auf unser gesellschaftliches Miteinander? Im Wissenschaftsjahr 2019 waren Bürgerinnen und Bürger aufgerufen, im Dialog mit Wissenschaft und Forschung Antworten auf diese und weitere Fragen zu finden.

Die Wissenschaftsjahre sind eine Initiative des Bundesministeriums für Bildung und Forschung (BMBF) gemeinsam mit Wissenschaft im Dialog (WiD). Sie tragen als zentrales Instrument der Wissenschaftskommunikation Forschung in die Öffentlichkeit und unterstützen den Dialog zwischen Forschung und Gesellschaft.

Weitere Informationen unter
www.wissenschaftsjahr.de

NATIONALES INSTITUT FÜR WISSENSCHAFTSKOMMUNIKATION

Das Nationale Institut für Wissenschaftskommunikation (NaWik) ist eine gemeinnützige GmbH, die 2012 von der Klaus Tschira Stiftung und dem Karlsruher Institut für Technologie (KIT) gegründet wurde. Das Institut vermittelt Forschenden und Studierenden die Grundlagen verständlicher und guter Wissenschaftskommunikation mit Nicht-Spezialisten. Das Lehrangebot des NaWik umfasst aktuell **elf Seminartypen** mit einem einheitlichen didaktischen Konzept – von Schreib- und Präsentationsseminaren bis hin zu Science Slam-Seminaren. Offene Seminare werden am Sitz des NaWik in Karlsruhe durchgeführt; Inhouse-Seminare können bundesweit gebucht werden. Zudem ist das NaWik bundesweit mit **Workshop-Reihen** und Vorträgen aktiv. Gemeinsam mit Wissenschaft im Dialog und dem KIT betreibt das NaWik die Plattform Wissenschaftskommunikation.de mit einer umfangreichen **Datenbank** zu verschiedenen Formaten in der Wissenschaftskommunikation.

Im Rahmen des Wissenschaftsjahres 2019 hat das NaWik das Projekt RIKI – Risikokommunikation zur Künstlichen Intelligenz realisiert.

INHALT

VORWORT	S. 4
PROJEKTDESCHEIBUNG	S. 6
KI-DEFINITIONEN	S. 8
RISIKOKOMMUNIKATION	S. 12
KI UND UMGANG MIT DATEN	S. 20
KI UND ARBEIT	S. 26
AUTONOMES FAHREN	S. 31
KI IN DER MEDIZIN	S. 39
KI UND ROBOTIK	S. 46
IMPRESSUM	S. 51

VORWORT

In der medialen Berichterstattung und damit auch in der öffentlichen Wahrnehmung genießt das Thema Künstliche Intelligenz (KI) einen hohen Stellenwert. Allerdings erscheint KI oft schwer greifbar und wird gerne mystifiziert oder glorifiziert. Bereits 1946 wurde der Computer ENIAC (Electronic Numerical Integrator and Computer) beispielsweise als „mathematischer Frankenstein“, „elektronisches Gehirn“ oder als Zauberer dargestellt.¹

Einige Jahrzehnte später hat sich der Nebel etwas gelichtet, KI hat zunehmend ein „Gesicht“ bekommen. Denn heute steuern KI-Algorithmen beispielsweise unseren Nachrichtenstrom in den Sozialen Medien. Sie können Bilder, Drehbücher oder Artikel generieren oder nach Profilen von Menschen fahnden. Künstliche Intelligenz beeinflusst unsere Entscheidungen, beispielsweise wie wir einkaufen, was wir lesen, hören und sehen. Das erzeugt bei den Menschen Sorgen vor Manipulierbarkeit und Kontrollverlust. Uns beschäftigen Themen wie Deep Learning, autonomes Fahren, automatische Gesichtserkennung, Profiling oder autonome Kampfdrohnen. Bei der Berichterstattung darüber stehen auch kritische Aspekte zunehmend im Fokus.^{2,3}

Der Journalist Patrick Beuth schreibt beispielsweise zum Thema Deep Learning auf Zeit Online über DeepMind-Mitbegründer Demis Hassabis: „Den Code will er allerdings unter Verschluss halten. AlphaGo Zero soll eine Blackbox bleiben, wie es viele andere Systeme mit künstlicher Intelligenz auch sind. Sogenannte Deep-Learning-Netzwerke, die den derzeitigen KI-Boom ausgelöst haben, sind auf einer bestimmten Ebene selbst für ihre Schöpfer undurchschaubar.“⁴

Aktuelle KI-Entwicklungen verändern unseren Alltag, und werden dies zunehmend tun. Deshalb sind Ängste und Bedenken ernst zu nehmen, und es ist wichtig, dass KI-Forschung nicht als undurchschaubar oder gar als im Elfenbeinturm verschlossen erscheint.

Wissenschaftlerinnen und Wissenschaftler sollten sich verstärkt in Dialogen mit Bürgerinnen und Bürgern engagieren und ihre Sicht der Dinge in die Politik und in die gesellschaftliche Diskussion einbringen. Sie sind die Expertinnen und Experten in ihrem Gebiet. Folglich können sie am besten Einblick in weitere Entwicklungen geben und zudem auf noch offene juristische oder gesellschaftspolitische Fragen hinweisen.

Das Projekt **„RIKI – Risikokommunikation zur Künstlichen Intelligenz“** des Nationalen Instituts für Wissenschaftskommunikation (NaWik), durchgeführt im Rahmen des Wissenschaftsjahres 2019 des Bundesforschungsministeriums, hatte zum Ziel, KI-Forschende für kritische Themen in ihrem Forschungsfeld zu sensibilisieren und ihnen Kommunikationskompetenzen zu vermitteln.

Gemeinsam mit KI-Nachwuchswissenschaftlern und -wissenschaftlerinnen haben wir in Workshops an sechs KI-Standorten diskutiert und erarbeitet,

- was der Stand der Wissenschaft in einzelnen Themenschwerpunkten ist,
- wie, im Vergleich dazu, die mediale Berichterstattung über das jeweilige Thema stattfindet,
- welche Zukunftsszenarien aus Sicht der Forschenden realistisch erscheinen,
- wo die Forschenden potenziell gesellschaftlich relevante kritische Themen sehen und
- wie sie selbst zu einer fundierten Debatte beitragen können.

LITERATUR

1. Schwartz, Oscar: 'The discourse is unhinged': how the media gets AI alarmingly wrong. The Guardian (25/07/2018) <https://www.theguardian.com/technology/2018/jul/25/ai-artificial-intelligence-social-media-bots-wrong>.

2. Lemm, Karsten: Gerät KI außer Kontrolle? Wenn ja, sind wir selber schuld. WIRED (03/09/2018) <https://www.gq-magazin.de/auto-tech-nik/article/geraet-ki-ausser-kontrolle-wenn-ja-sind-wir-selber-schuld>.

3. Furger, Michael & Metzler, Marco: Was müssen unsere Kinder lernen, damit sie gegen intelligente Maschinen bestehen können? Neue Zürcher Zeitung am Sonntag (24/06/2018) <https://nzzas.nzz.ch/hintergrund/wie-verhindere-ich-dass-mein-hausroboter-meine-katze-kocht-ld.1397478?reduced=true>.

4. Beuth, Patrick: Die Automaten brauchen Aufsicht. Zeit Online (25/10/2017) <https://www.zeit.de/digital/internet/2017-10/kuenstliche-intelligenz-deepmind-back-box-regulierung>.

→ Alle hier genannten Webseiten geprüft am: 24.01.2020

Als Themenschwerpunkte haben sich in Abhängigkeit der teilnehmenden KI-Forschenden herauskristallisiert:

- Umgang mit Daten
- Arbeit
- Autonomes Fahren
- Medizin
- Robotik

Dabei zeigten sich wiederholt Grenzen im Entscheidungsrahmen für die Forschenden. Häufig sehen diese die dringende Notwendigkeit gesellschaftspolitischer Debatten, um weit über Fachkreise hinaus zu gemeinsamen Entscheidungen und Regeln im Umgang mit verschiedenen Aspekten des Einsatzes von Künstlicher Intelligenz zu gelangen.

In diesem Kompendium stellen wir je Themenschwerpunkt

- zentrale Aussagen der medialen Berichterstattung der vergangenen Jahre zusammen,
- geben wir einen verschriftlichten Einblick in die Diskussionen der Forschenden während der RIKI-Workshops und den öffentlichen Fishbowl-Diskussionen,
- stellen wir mögliche Antworten der Forschenden auf kritische Fragen zusammen.

Das Kompendium vermittelt die Erkenntnisse aus dem einjährigen Kommunikationsprojekt im Rahmen des Wissenschaftsjahres 2019 an Wissenschaftlerinnen und Wissenschaftler, die im Bereich Künstliche Intelligenz forschen. Es sind insbesondere die kritischen Fragestellungen, bei denen die gesamtgesellschaftliche Auseinandersetzung wichtig ist – und dabei spielen die Forschenden mit ihrer fachlichen Kompetenz eine zentrale Rolle. Deswegen soll dieses Kompendium insbesondere den Forschenden entsprechende Handreichungen geben.

– Beatrice Lugger

DAS RIKI-PROJEKT

Risikokommunikation zur Künstlichen Intelligenz (KI) – dafür steht das Kürzel RIKI. Dieses Projekt des Nationalen Instituts für Wissenschaftskommunikation (NaWik) fand im Rahmen des Wissenschaftsjahres 2019 statt und wurde vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Ziel war es, junge Forschende aus dem Bereich KI zu ermutigen und zu unterstützen, sich in die öffentliche Debatte rund um ihr Forschungsthema einzubringen.

Dazu lud das NaWik **Nachwuchswissenschaftlerinnen und -wissenschaftler** zu jeweils **eintägigen Workshops an sechs ausgewiesenen KI-Hot-Spots** in Deutschland ein (siehe Karte).

In den Workshops haben die Teilnehmerinnen und Teilnehmer zunächst die Grundlagen guter Wissenschaftskommunikation sowie der Risikokommunikation kennengelernt. Im Team wurden dann mögliche kritische Aspekte im Forschungsfeld des jeweiligen Standorts identifiziert. Abschließend erarbeiteten die Forschenden Antworten auf einzelne kritische Fragen in Form einer Kernbotschaft.

Zusätzlich diskutierten Spitzen-Forschende sowie Nachwuchswissenschaftlerinnen und -wissenschaftler an den zwei Standorten mit interessierten Bürgerinnen und Bürgern im sogenannten Fishbowl-Format. Die **Debatte in Karlsruhe** behandelte das Thema **autonomes Fahren**. In **Berlin** stand das Thema **KI in der Medizin** im Mittelpunkt.

Dieses Kompendium fasst die Erkenntnisse aus den sechs Workshops sowie den beiden Diskussionen zusammen.

Neben diesem Kompendium, das einen vertiefenden schriftlichen Einblick gibt, haben einzelne Forschende im Anschluss an die Workshops jeweils gezielt auf einzelne mögliche kritische Fragen Antworten in Form von **Videobotschaften** gegeben. Diese, sowie die **Mitschnitte der Fishbowl-Diskussionen**, sind in einem eigenen YouTube-Kanal und auf der Projekt-Website – www.nawik.de/riki/videos/ – abrufbar.

RIKI - TEAM

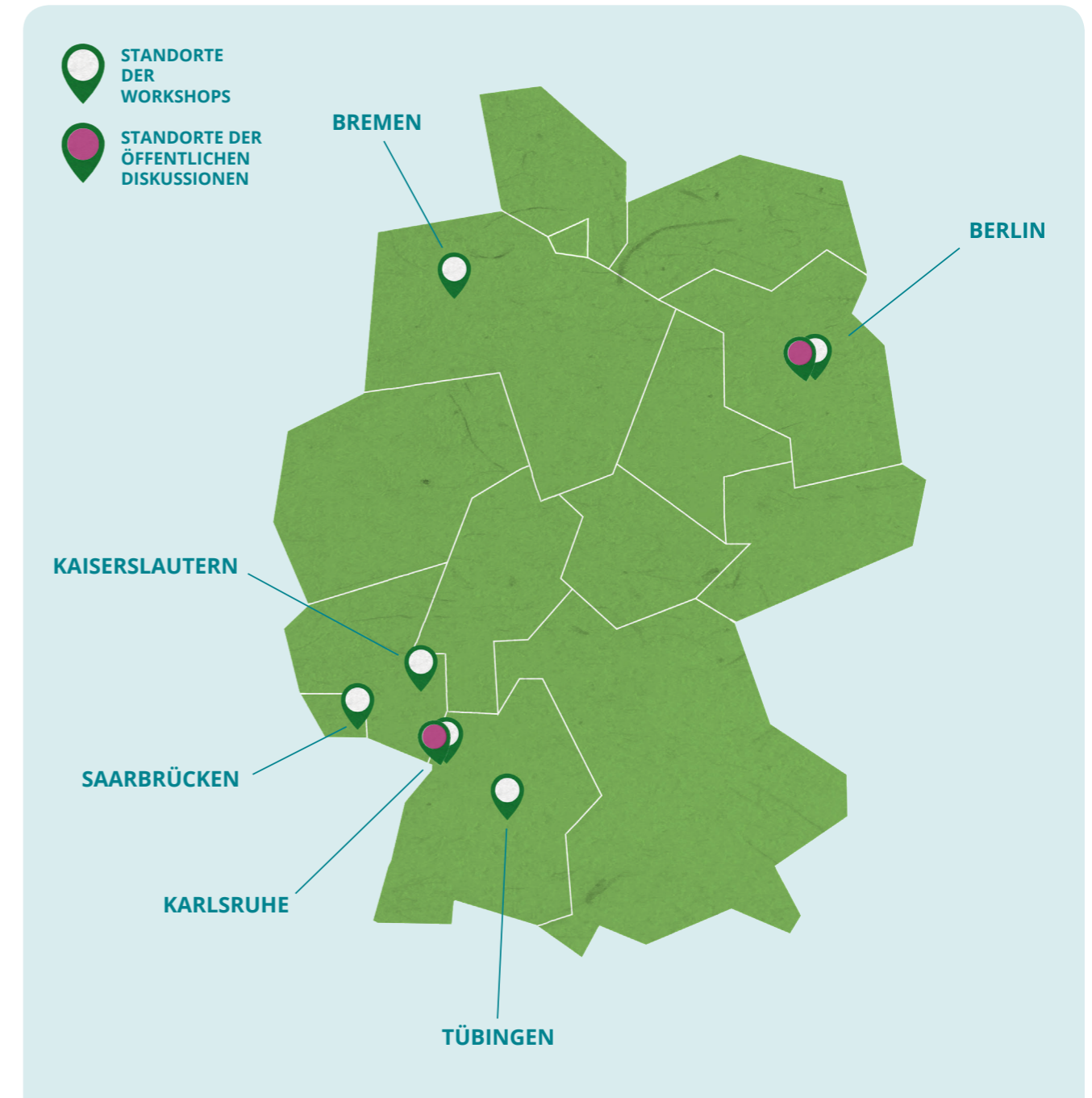
Virginia Albert hat das Projekt Risikokommunikation zur Künstlichen Intelligenz (RIKI) koordiniert. Die Biomedizinerin ist Projektkoordinatorin am NaWik.

Nicola Kuhrt brachte als Dozentin dieses Projekts in zwei Workshops ihre Expertise als Wissenschaftsjournalistin ein. Zudem moderierte sie die öffentliche Diskussion „Künstliche Intelligenz in der Medizin – Reden Sie mit!“ in Berlin. Sie ist freie Medizinjournalistin und Co-Gründerin von MedWatch.de.

Beatrice Lugger ist Geschäftsführerin und Direktorin des NaWik und leitete das RIKI-Projekt. Sie ist Wissenschaftsjournalistin, Social-Media-Expertin und Chemikerin.

Klaus Wingen ist Risiko- und Krisenkommunikationsexperte am NaWik und leitete alle sechs Workshops im Rahmen des RIKI-Projekts. Dabei brachte er insbesondere seine Expertise aus dem Journalismus sowie aus der Öffentlichkeitsarbeit ein.

Eva Wolfangel wirkte als Dozentin bei vier der sechs Workshops dieses Projekts mit. Zudem moderierte sie die Diskussion „Reden Sie mit: Autonomes Fahren und Künstliche Intelligenz“ in Karlsruhe. Die Wissenschaftsjournalistin legt bei der Wahl ihrer Themen einen besonderen Schwerpunkt auf Künstliche Intelligenz.



STANDORTE	BETEILIGTE FORSCHUNGSINSTITUTE	DATUM
KARLSRUHE	Karlsruher Institut für Technologie (KIT) und FZI Forschungszentrum Informatik	30.04.2019
		06.07.2019
SAARBRÜCKEN	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)	14.05.2019
KAISERSLAUTERN	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)	25.06.2019
TÜBINGEN	Tübingen AI Center – Competence Center for Machine Learning	16.07.2019
BREMEN	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)	19.08.2019
BERLIN	Fraunhofer Gesellschaft in Kooperation mit dem Fraunhofer Institut für Digitale Medizin (MEVIS)	25.10.2019
		26.10.2019

KI – DEFINITIONEN

KÜNSTLICHE INTELLIGENZ

Was genau ist eigentlich mit Künstlicher Intelligenz (KI) gemeint? Dies war im Rahmen der Workshop-Reihe zur Risikokommunikation Künstlicher Intelligenz (RIKI) wiederholt ein zentraler Diskussionspunkt. Die Forschenden betonten mehrfach, dass sie bei Gesprächen zur KI fast immer zunächst Missverständnisse in der Einschätzung vom gegenwärtigen Stand der KI-Forschung klären müssen – ob im Familien- und Freundeskreis, ob im Austausch mit Medien oder auch innerhalb der Wissenschaften.

Zu mächtig seien die Bilder aus der Science-Fiction von starken, selbstlernenden Künstlichen Intelligenzen, die über eine dem Menschen ähnliche Intelligenz verfügen. Auch beklagten die Forschenden, dass die Mehrzahl der Berichte in den Medien mit Bildern von Robotern versehen seien. Dabei sei die Robotik doch nur ein Teilgebiet der KI, sozusagen KI verpackt in eine Maschine, die in Aktion mit der physikalischen Welt tritt.

Als zudem problematisch stuften die Forschenden Äußerungen prominenter Physiker und Zukunftsvisionäre zur sogenannten Singularität ein. Gemeint ist damit ein Zeitpunkt, an dem Künstliche Intelligenz menschliche Intelligenz übertrifft – mit unvorhersehbaren Folgen. Gern zitiert wird dazu der Astrophysiker Stephen Hawking mit einer Äußerung aus dem Jahr 2017: „Der Erfolg bei der Schaffung einer effektiven KI könnte das größte Ereignis in der Geschichte unserer Zivilisation sein. Oder das Schlimmste. Wir wissen es einfach nicht.“¹ Und Elon Musk merkte etwa zur gleichen Zeit an: „Künstliche Intelligenz

stellt ein grundlegendes Risiko für die Existenz der menschlichen Zivilisation dar, auf eine Weise wie es Autounfälle, Flugzeugabstürze, schadhafte Drogen oder schlechtes Essen nie waren.“² Der Nachrichtenwert solcher Verlautbarungen ist zweifellos hoch, die Einstufung als „problematisch“ durch KI-Expertinnen und -Experten gleichwohl berechtigt.

DEFINITIONEN VON KÜNSTLICHER INTELLIGENZ

Eine einzige, allgemeingültige oder von allen Akteuren konsistent genutzte Definition von KI gibt es nicht.³ Dies zeigt auch die Zusammenstellung unterschiedlicher Definitionen in der Grafik auf S.9.

Allgemein gilt, dass KI ein Teilgebiet der Informatik ist, das sich damit beschäftigt, Maschinen mit Fähigkeiten auszustatten, die intelligentem (menschlichen) Verhalten ähneln – Lernen, Planen oder Problemlösen. Dies kann einerseits mit Hilfe klar vorgegebener und programmierter Regeln erreicht werden oder durch maschinelles Lernen.

Lernende KI-Systeme verbessern die vorab trainierten Modelle und erweitern so ihre Wissensbasis sowie ihre Fähigkeiten. Sämtliche heute technisch umsetzbaren KI-Systeme ermöglichen eine Problemlösung in beschränkten Kontexten (z. B. Sprach- oder Bilderkennung) und zählen damit zur sogenannten schwachen KI (s.u.).⁴

Heute wird KI mehrheitlich zur Lösung konkreter Probleme eingesetzt. Aktuelle KI-Systeme definieren ihre eigenen Ziele nicht. Allerdings können einige KI-Systeme die Freiheit haben, zu entscheiden, welchen Weg sie einschlagen, um das vorgegebene Ziel zu erreichen.⁵ Eine KI ist dabei praktisch immer ein Spezialist, ein System, das für eine bestimmte Fragestellung entwickelt worden ist.⁶

Künstliche Intelligenz ist...

... eine Informatik-Anwendung, deren Ziel es ist, intelligentes Verhalten zu zeigen. Dazu sind in unterschiedlichen Anteilen bestimmte Kernfähigkeiten notwendig: Wahrnehmen, Verstehen, Handeln und Lernen.[...] (Bitkom & DFKI, Definition 2)⁸

... ein Teilgebiet der Informatik, das sich damit beschäftigt, Maschinen mit Fähigkeiten auszustatten, die intelligentem (menschlichem) Verhalten ähneln. Dies kann mit vorprogrammierten Regeln oder durch maschinelles Lernen erreicht werden. Starke bzw. generelle KI bezeichnet Maschinen, die generalisierende Intelligenz- und Transferleistungen erbringen können und somit nicht nur auf sehr begrenzte, vordefinierte Aufgabenfelder beschränkt sind. (Fraunhofer)⁹

... the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. (Encyclopedia Britannica)¹⁰

... ein Überbegriff für viele einzelne Teildisziplinen wie Big Data, Statistik oder Machine Learning. (RIKI-Workshop DFKI Saarbrücken)

... die Digitalisierung menschlicher Wissensfähigkeiten. (RIKI-Workshop Saarbrücken)

... ein Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens und dem maschinellen Lernen befasst. (Wikipedia)¹¹

... die Eigenschaft eines IT-Systems, »menschensähnliche«, intelligente Verhaltensweisen zu zeigen. (Bitkom & DFKI, Definition 1)⁸

... der Überbegriff für Anwendungen, bei denen Maschinen menschenähnliche Intelligenzleistungen erbringen. Darunter fallen das maschinelle Lernen oder Machine Learning, das Verarbeiten natürlicher Sprache (NLP – Natural Language Processing) und Deep Learning. Die Grundidee besteht darin, durch Maschinen eine Annäherung an wichtige Funktionen des menschlichen Gehirns zu schaffen – Lernen, Urteilen und Problemlösen. (SAP)¹²

STARKE UND SCHWACHE KÜNSTLICHE INTELLIGENZ

Sämtliche heute technisch umsetzbaren KI-Systeme zählen zur sogenannten **schwachen Künstlichen Intelligenz**. Schwache KI fokussiert jeweils auf die Lösung von bestimmten Anwendungsproblemen. Dafür werden bestimmte Grundlagen vorgegeben und aufbauend kann das KI-System weitere Aspekte durch Simulation oder Analyse großer Datenmengen selbst optimieren. Schwache KI ermöglichen gemäß der Plattform Lernende Systeme bislang Problemlösungen nur in beschränkten Kontexten.⁴ Typische Anwendungen schwacher KI sind unter anderem Musteranalysen wie Sprach- oder Bilderkennung; die Modellierung und das Erfassen von Wissen; die autonome Steuerung von Robotik-Systemen.

Starke Künstliche Intelligenz meint Maschinen, die eine generalisierende Intelligenz hätten und auch Transferleistungen erbringen könnten. Ein starkes KI-System sollte somit nicht auf vordefinierte und begrenzte Aufgabenfelder beschränkt sein, sondern die meisten Aktivitäten ausführen können, die auch Menschen ausführen können.⁵

Von einer solchen starken oder „generellen“ Künstlichen Intelligenz, die vergleichbare intellektuelle Fertigkeiten wie der Mensch habe oder ihn darin sogar übertreffen könnte, sind wir nach Ansicht der Forschenden weit entfernt.

MASCHINELLES LERNEN

Maschinelles Lernen ist eine Schlüsseltechnologie der KI.⁴ Das System kann dabei automatisch lernen, ohne dass es auf von Menschen vorgegebene und definierte Regelwerke zurückgreift oder einfach etwas auswendig lernt. Die KI-Maschine entwickelt selbst auf der Basis von Beispieldaten eigene Regeln und Modelle, die sie dann auf neue Daten oder Situationen anwendet.

So kann die KI auch lernen, wie ein Problem gelöst werden kann, das nicht exakt definiert oder dessen Lösungsweg nicht exakt beschreibbar ist (Beispiele sind Sprache, Sprachverstehen...).⁵ Ein Teilbereich des maschinellen Lernens setzt neuronale Netze für das sogenannte Deep Learning ein, das auf der Basis großer Datensätze (Big Data) komplexe Muster erkennen kann.

Meist unterscheidet man beim maschinellen Lernen drei Ansätze,

- Überwachtes Lernen (Supervised Learning),
- Unüberwachtes Lernen (Unsupervised Learning) und
- Verstärkendes Lernen (Reinforcement Learning).

Überwachtes Lernen – Supervised Learning

Überwachtes Lernen ist beispielsweise in der Bilderkennung im Einsatz. Es sind keine Regeln vorgegeben. Aber das System bekommt viele Daten (etwa von Katzen versus Hunden – oder von Muttermalen versus Hautkrebs), bei denen das Ergebnis bereits zugeordnet ist. Im Trainingsprozess bekommt das System außerdem von Menschen Feedback, ob es richtig zugeordnet hat. So lernt das System im Laufe der Zeit selbst Katzen von Hunden oder Hautkrebs von Muttermalen zu unterscheiden. Es bildet bekannte Gesetzmäßigkeiten nach oder ab. Die Ergebnisse des Lernprozesses lassen sich mit den bekannten, richtigen Ergebnissen vergleichen, also „überwachen“.⁷ Wichtig für die Qualität der Entscheidungen der Systeme ist die Menge der Trainingsdaten.

Unüberwachtes Lernen – Unsupervised Learning

Beim unüberwachten Lernen liegen keinerlei vorgegebene Datenpaare oder zu erreichende Zielwerte vor.⁷ Stattdessen geht es darum, in einem großen, unstrukturierten Datensatz mögliche Korrelationen, Muster und Strukturen sowie alles, was von diesen Clustern abweicht, zu erkennen – Ausreißer, Anomalien, seltene Objekte – und daraus Regeln abzuleiten. Ein Beispiel ist

die Segmentierung von Kundendaten nach Zielgruppen, die man auf ähnliche Weise adressieren möchte.⁴

Verstärkendes Lernen – Reinforcement Learning

In gewisser Weise ist das eigenständige sich verstärkende Lernen eine Unterform des überwachten Lernens. Der Unterschied ist, dass hier keine korrekten Input-Output-Paare vorgegeben sind und das Training eher situativ und dynamisch nach dem Prinzip von „trial & error“ erfolgt.⁷ Das System bekommt aber bei jeder Entscheidung, die es trifft ein Signal, ob es eine gute oder eine schlechte Entscheidung war. Das Ziel des Systems ist es, im Laufe der Zeit die positiven Rückmeldungen zu maximieren.⁵ Dieser Ansatz wird zum Beispiel bei Spielen wie Schach oder Go (AlphaGo) verwendet.

Deep Learning

Deep Learning ist ein häufig gesondert betrachtetes Teilgebiet des maschinellen Lernens. Diese Systeme sind inspiriert von Funktionsweisen menschlicher Nervensysteme. Sie haben ein Netzwerk von kleinen Verarbeitungseinheiten (in Analogie zu menschlichen Neuronen) mit unterschiedlich gewichteten Verbindungen. Diese sogenannten neuronalen Netzwerke verarbeiten Informationen in verschiedenen Schichten von Knotenpunkten, die häufig hierarchisch geordnet sind. Typischerweise haben sie eine Eingabe- und eine Ausgabeschicht und eine variierende Zahl dazwischenliegender Schichten.

Wird in der ersten Schicht ein Muster erkannt, so erkennt die zweite Schicht beispielsweise ein Muster des Musters, usw. Durch die vielen verschiedenen Schichten der neuronalen Netze und der Auswertung enorm großer Datenvolumen ermöglicht dieses Lernverfahren die Lösung komplexer Probleme. Deep Learning Methoden finden Anwendung in der Bilderkennung, beim

autonomen Fahren, der Umwandlung von Fotos in Kunstwerke, der Spracherkennung u.v.m.

GRENZBEREICHE KÜNSTLICHER INTELLIGENZ

Datenverzerrungen

Zentral beim Training von KI-Systemen ist meist der Einsatz großer Datenmengen (Big Data), um weiter zu lernen und gut zu funktionieren. Allerdings kommt es bei bestimmten Datensätzen immer wieder vor, dass die von Menschen erstellten Daten selbst Verzerrungen aufweisen. Wenn nun einer KI ein verzerrter Datensatz zur Verfügung gestellt wird, wird das KI-System diese Verzerrung fortschreiben. Dies kann zu unfairen Entscheidungen, falschen Entscheidungen oder etwa der Bevorzugung einer Gruppe gegenüber einer anderen führen.

Fehlbewertungen

Die Techniken des maschinellen Lernens erreichen immer nur eine gewisse Genauigkeit – beispielsweise eine relativ hohe Wahrscheinlichkeit bei der Auswertung medizinischer Bilddaten. Diese kann unter Umständen sehr hoch liegen. Dennoch ist mit Fehlern immer zu rechnen und es bleibt nach wie vor in der Verantwortung des Menschen, die Bewertungen der KI richtig einzuordnen.

Black Box

Je freier ein maschinelles Lernverfahren von Vorgaben durch Menschen ist, desto weniger sind dessen Entscheidungsfindungen noch nachvollziehbar. Von einer Black Box spricht man, wenn es nicht mehr möglich ist, Entscheidungen nachzuvollziehen.

LITERATUR

1. Wimmer, Barbara: KI könnte schlimmstes Ereignis der Menschheit werden. Futurezone (07/11/2017) <https://futurezone.at/science/stephen-hawking-ki-koennte-schlimmstes-ereignis-der-menschheit-werden/296.805.846>.
2. Wenner, Franz-Georg & Saurenz, Daniel: Tech-Riesen klotzen am Zukunftsmarkt. n-tv (22/07/2017) <https://www.n-tv.de/wirtschaft/Tech-Riesen-klotzen-am-Zukunftsmarkt-article19949113.html>.
3. Die Bundesregierung: Strategie Künstliche Intelligenz der Bundesregierung. (11/2018) https://www.bmbf.de/files/Nationale_KI-Strategie.pdf.
4. Beck, Susanne et al.: Künstliche Intelligenz und Diskriminierung – Whitepaper aus der Plattform Lernende Systeme, München 2019. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG3_Whitepaper_250619.pdf.
5. AI HLEG – High-level expert group on artificial Intelligence: A definition of AI: main capabilities and scientific disciplines. (European Commission, 08/04/2019). <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.
6. Albert, Andreas; Böhm, Markus et al.: Künstliche Intelligenz - endlich verständlich. Spiegel Online (04/01/2017) <http://www.spiegel.de/netzwelt/gadgets/kuenstliche-intelligenz-turing-test-chatbots-neuronale-netzwerke-a-1126718.html>
7. Wagoner, Andreas: Maschinelles Lernen und KI im Marketing: Lernmethoden und ihre Einsatzmöglichkeiten im Marketing. Nerdwärts. de (28/06/2019) <https://nerdwaerts.de/2019/06/maschinelles-lernen-und-ki-im-marketing-lernmethoden-und-ihre-einsatzmoeglichkeiten-im-marketing/>.
8. Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) & Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (Bitkom): Künstliche Intelligenz. Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung. (2017) https://www.dfki.de/fileadmin/user_upload/import/9744_171012-KI-Gipfpapier-online.pdf
9. Fraunhofer Gesellschaft: Trends für die Künstliche Intelligenz. (2017) https://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer_Trendbrosch%C3%BCre_KI_small.pdf
10. Copeland, B.J.: Artificial Intelligence. Encyclopedia Britannica <https://www.britannica.com/technology/artificial-intelligence>
11. Künstliche Intelligenz. Wikipedia https://de.wikipedia.org/wiki/K%C3%BCnstliche_Intelligenz
12. Schick, Uwe: Was ist künstliche Intelligenz? SAP News Center (20/03/2018) <https://news.sap.com/germany/2018/03/was-ist-kuenstliche-intelligenz/>

→ Alle hier genannten Webseiten geprüft am: 24.01.2020

RISIKOKOMMUNIKATION

WARUM RISIKOKOMMUNIKATION ZUR KÜNSTLICHEN INTELLIGENZ?

Viele Menschen machen sich berechtigte oder auch unberechtigte Sorgen darüber, wohin immer ausgereifere Technologien letztlich führen können. Der Begriff „**Sorgen**“ lässt sich dabei auch umdeuten in wahrgenommene „**Risiken**“ – Risiken, Gefahren, Bedrohungen, die eine bestimmte Entwicklung für den Alltag in unserer Gesellschaft womöglich birgt. Diese Sorgen wachsen zum einen umso mehr, je weniger wir verstehen, wie die jeweiligen Technologien eigentlich funktionieren. Und zum anderen werden diese Sorgen umso größer, je schlechter wir sie auf Basis unseres bisherigen Erfahrungswissens einordnen können. Künstliche Intelligenz und ihre sehr unterschiedlichen Anwendungen sind ein prominentes Beispiel für ein Thema, dessen Einschätzung von Unsicherheit geprägt ist.

WIE GEHEN WIR MIT RISIKEN UM?

Risikokommunikation ist Teil unseres Alltags, und sie wird als solche oft gar nicht bewusst wahrgenommen. Das reicht vom Verkehrsschild „Vorfahrt beachten“ bis hin zum Beipackzettel eines Medikaments gegen Migräne. Was wir unbedingt vermeiden wollen, sind übrigens nicht Risiken als solche, sondern Verluste. Etwa den Blechschaden am Auto, aber eben auch die Ankunft genau eine Minute nach Geschäftsschluss. **Risiken und Nutzen** gegeneinander abzuwägen, bevor wir eine Entscheidung treffen: In der Kommunikation über Künstliche Intelligenz stellt dies einen ganz elementaren Aspekt dar.

WIE GELANGEN WIR ZU EINER EINSCHÄTZUNG VON RISIKEN?

Risiken zu berechnen – das ist das Kerngeschäft von Versicherungs-Mathematikern und -Mathematikerinnen. Dabei spielen zwei Faktoren eine entscheidende Rolle: Erstens die **mutmaßliche Schadenshöhe** eines negativen Ereignisses, und zweitens die **Wahrscheinlichkeit**, mit der dieses negative Ereignis mutmaßlich eintritt.

Für potenziell betroffene Bürgerinnen und Bürger zählt aber das **gefühlte Risiko** – und da kommen noch zwei Faktoren hinzu, die eng miteinander zusammenhängen. Zum einen die subjektive Einschätzung – oder etwas wissenschaftlicher ausgedrückt: das individuelle mentale Konstrukt. Und zum anderen die Thematisierung in den Medien.

In den Medien begegnet uns das Thema KI fast täglich, die **mediale Präsenz** ist zweifellos sehr hoch. Folglich auch deren Einfluss auf das individuell gefühlte Risiko. Die Zeiten einer sehr holzschnittartigen Thematisierung einer übermächtigen KI als wahlweise existenzbedrohend oder unentbehrlich scheinen auch in den Medien einer etwas realistischeren Sicht der Dinge gewichen zu sein.

Doch was ist mit der **subjektiven Einschätzung**? Sie ist eben genau dies: subjektiv. Und damit höchst individuell. Ein plastisches Beispiel ist ein Sprungturm im Schwimmbad. Dort herunter zu hüpfen macht Spaß. Aber viele Menschen haben vor dem Sprung ins Wasser auch Angst, denn ungünstig auf dem Wasser aufzukommen, kann unangenehme Folgen haben. Plastischer: Ein Bauchklatscher statt eines eleganten Kopfsprungs tut verdammt weh, und dies macht den Sprung zu einem Risiko. Dies gilt für manche schon für das Ein-Meter-Brett, für andere erst für das Drei-Meter-Brett. Während die einen es vorsichtshalber beim normalen Sprung belassen, finden andere

einen Rückwärts-Salto noch prickelnder als einen Kopfsprung.

So verhält es sich auch mit der Risikowahrnehmung beim Thema Künstliche Intelligenz. Es gibt keine allgemeingültige Grenze, ab wann wir mit der KI verbundene Risiken als solche wahrnehmen.

WELCHE GEFÜHLE BEGLEITEN UNS DABEI?

Wenn wir über Risikokommunikation reden, dann reden wir immer auch über Besorgnis, Angst, Wut, Enttäuschung, über Unverständnis und Ungeduld. Risikokommunikation bedarf folglich der Empathie, also des **Einfühlungsvermögens** in die Gefühlslage der anderen. Dies betrifft übrigens in besonderer Weise die Wissenschaftskommunikation. Denn gerade bei Spitzentechnologien erscheinen beide Seiten ein und derselben Medaille besonders gegensätzlich: Auf der einen Seite Faszination und Bewunderung, auf der anderen Seite Erschrecken und Befürchtungen.

In Anlehnung an die großen Philosophen des alten Griechenlands reicht es bei der Kommunikation nicht, nur auf logische, offenkundig „vernünftige“ Argumente zu setzen – die Griechen nannten dies „**Logos**“. Was ergänzend hinzutreten muss, ist zum einen ein gewisser Sympathiefaktor, verbunden mit der bereits erwähnten Empathie – also der Faktor „**Pathos**“. Als besonders wirksam gilt Kommunikation aber erst dann, wenn auch der Aspekt des „**Ethos**“ berücksichtigt wird. Er beschreibt die Haltung, die Glaubwürdigkeit, also Dinge, die den aufrichtigen Charakter des Kommunizierenden hervorheben. Dazu gehört zum Beispiel, auf gemeinsame Werte hinzuweisen, auf die eigene Sachkunde und Erfahrung mit der Materie, und auf das Bestreben, damit zum Wohl der Gemeinschaft beitragen zu wollen und nicht egoistische Motive zu bedienen.

Wichtig: Betrachten Sie im Gespräch Ihr Gegenüber nicht als Gegner oder gar Feind. Sie dürfen Wertschätzung erwarten dafür, dass Sie sich bereitwillig in die öffentliche Diskussion einbringen. Und umgekehrt ist es an Ihnen, dem Gegenüber mit **Wertschätzung** zu begegnen. Das ist nicht immer leicht – aber es trägt erfahrungsgemäß zur Versachlichung bei, in aufgeregten Momenten im Gegenüber weniger die Person, sondern mehr deren Rolle zu sehen.

WIE STEHT ES UM DIE AKZEPTANZ VON KI?

Ganz allgemein steigt die Technik-Skepsis mit zunehmendem Alter an, wobei Frauen der Technik noch reservierter begegnen als Männer. Die Akzeptanz von Technik steigt mit dem Bildungsgrad, also je höher der erreichte Bildungsabschluss, desto größer die Bereitschaft, sie als Teil des Alltags zu begreifen. In der Tat liegt es ja nahe, dass sich gut Gebildete einer Generation, die mit der Digitalisierung groß geworden ist, bereitwilliger mit neuen Technologien arrangieren.¹

Ein anderer Trend: Je wirkmächtiger die IT, desto eher fühlen sich die Menschen durch neue Entwicklungen abgehängt. Viele Menschen, die sich mit der „normalen“ Computertechnologie noch gut arrangieren können, sehen sich überfordert, noch aufwändigeren Technologien wie der KI zu folgen.² Dazu passt auch ganz gut, dass sich eine große Mehrheit beispielsweise wünscht, dass KI als Urheberin von künstlich generierten Texten erkennbar bleibt.³

WIE UNTERSCHIEDET SICH RISIKO-VON KRISENKOMMUNIKATION?

Gerne werden die Begriffe **Risiko** und **Krise** in einen Topf geworfen. Dass dies nicht zielführend ist, zeigt ein Blick auf die zugrunde liegende Kommunikationsstruktur. Bei der Risikokommunika-

tion ist noch kein negatives Ereignis eingetreten. Deshalb ist es sinnvoll, diese Kommunikation stärker partizipativ anzulegen – etwa mögliche Betroffene unterschiedlicher Risikoszenarien zu Dialogen und zum Austausch einzuladen. Anders ist die Situation im Falle einer eingetretenen Krise. Dann wollen mögliche Betroffene vor allem schnell informiert werden. Hier steht also im ersten Schritt die Information durch den Verursacher im Zentrum der Kommunikation. Dialogische Elemente folgen meist erst später.

LASSEN SICH RISIKEN VORHERSEHEN?

Bei der rechtzeitigen Beurteilung möglicher Risiken hilft ungemein, so etwas wie ein Themen-Radar einzusetzen. Daher hat ein **Monitoring** aktueller Entwicklungen, Berichterstattung und Diskussionen gerade im Bereich der Risiko- und der Krisenkommunikation besondere Bedeutung. Dazu gehört beispielsweise auch, ein Thema vor dem jeweiligen gesellschaftlichen Hintergrund einzuordnen und die eigene Sicht der Dinge in einem Positionspapier zu skizzieren.

Social Media sind hervorragende Indikatoren im Sinne einer Beobachtung der öffentlichen Diskussion. Auch KI-Themen finden sich in diesen Netzwerken zuhauf. Es ist also sinnvoll, wenn schon nicht aktiv, dann wenigstens passiv, regelmäßig ein Stimmungsbild zu bestimmten Themen dort abzurufen. Denn wo sich Nachrichten besonders schnell verbreiten, zeigen sich auch besonders schnell Trends – bis hin zum berüchtigten Shitstorm. Social Media wie etwa Twitter lassen sich auch sehr gut passiv nutzen. In Abhängigkeit davon, wem man folgt, können sie als persönliche Nachrichtenticker auch für Spezialthemen dienen.

WELCHE ROLLE SPIELEN JURISTISCHE GESICHTSPUNKTE?

Bei der Frage, welche Inhalte zur Sprache gebracht werden, gibt es auch rechtliche Gesichtspunkte zu beachten, die Einfluss auf die Art und auf das Ausmaß der Kommunikation haben. Zum einen ist dies ein berechtigter Informationsanspruch der Öffentlichkeit, verbunden mit der Auskunftspflicht öffentlicher Einrichtungen. Zum anderen sind dies aber auch Interna, geschützte Daten, möglicherweise Geschäftsgeheimnisse und vertrauliche Kooperationsvereinbarungen, die nichts in der öffentlichen Diskussion über KI verloren haben.

RISIKEN DER KI KOMMUNIZIEREN – WER?

Wer sollte eigentlich Risiken kommunizieren? Wer wirkt besonders **authentisch**? Doch in aller Regel diejenigen, die sich konkret mit der jeweiligen Sache befassen! Also diejenigen, die in ihrem Alltag daran arbeiten, die aus dieser Perspektive einschätzen können, wie weitreichend die Konsequenzen für unsere Zukunft möglicherweise sind.

Es sind daher die **Wissenschaftlerinnen und Wissenschaftler**, die im Vermittlungsprozess von Informationen und Botschaften als besonders **glaubwürdig wahrgenommen** werden. Aus diesem Grund ist es so wichtig, dass die Forschenden selbst ihre Sicht der Dinge öffentlich machen. Dass sie es sind, die darüber reden, was sie der Technologie abgewinnen können. Dass sie zudem darüber reden, wie sie einzelne Entwicklungen in ihrem Bereich einschätzen. Und, ja, dass sie auch darüber reden, was ihnen selbst Sorgen bereitet. Den KI-Forschenden kommt eine besonders wichtige Rolle zu in dem Prozess, Chancen und Risiken der KI zu vermitteln, damit die Adressaten Chancen und Risiken dann gegeneinander abwägen können.

In Bezug auf KI bedeutet das, dass die KI-Forschenden gefordert sind, frühzeitig auf potenzielle Risi-

ken einer bestimmten KI-Anwendung hinzuweisen und damit die notwendigen Dialoge in der Gesellschaft zu initiieren.

Experten und Expertinnen sind in ihrem Spezialgebiet zu Hause. Sie verfügen über vielfältiges Fach- und Detailwissen, während sich Laien nur oberflächlich auskennen und sich auch nur für bestimmte Aspekte des jeweiligen Fachgebiets interessieren. Fachleute müssen also einschätzen, welche Inhalte und Details für die Zielgruppe konkret von Interesse sein dürften und sich in der Kommunikation entsprechend beschränken.

Auch bei der Risikokommunikation zur KI gilt es genau zu überlegen, was das eigentliche **Thema** ist. Es gilt Schritt für Schritt die möglichen Ebenen durchzugehen, bis das zentrale Thema definiert ist. Zum Beispiel: Informatik – Bildverarbeitung – automatische Objekterkennung.

RISIKEN DER KI KOMMUNIZIEREN – MIT WEM?

Die Adressaten sind keine homogene Gruppe, deswegen heißt es für die Kommunizierenden, in **Zielgruppen** zu denken und sich auf die jeweilige Zielgruppe einzustellen. Das können die Hörerinnen und Hörer eines Interviews mit dem Deutschlandradio sein, die Leserinnen und Leser eines Nachrichtenmagazins oder auch die Schülerinnen und Schüler einer achten Klasse des örtlichen Gymnasiums. Die denkbaren Zielgruppen sind also bunt, manchmal gar schillernd bunt.

Risikokommunikation bietet die Chance, Zielgruppen sorgfältig zu gliedern. Eine denkbare Aufteilung in vier große Gruppen folgt dabei einem Schema, das sich in der Krisenkommunikation bewährt hat. Danach kann es sinnvoll sein, bei der Einschätzung der Adressaten zwischen deren Interesse am Thema (niedrig / hoch) und deren

Einfluss auf andere (niedrig / hoch) zu unterscheiden. Selbstredend verdient die Gruppe der einflussreichen Interessierten (z.B. Aufsichtsgremien einer KI-Forschungseinrichtung) ein besonderes Augenmerk, während es bei den uninteressierten Einflusslosen eher darum geht, sie nicht ganz aus dem Auge zu verlieren.

Solch ein Schema hilft spätestens dann, wenn sich unterschiedliche Zielgruppen zu Wort melden – und dies auch noch unterschiedlich laut. Denn in diesem Moment schärft solch ein Schema den Blick dafür, dass nicht die Lautesten automatisch die Wichtigsten sind.

Das Verständnis von Fachbegriffen und technischen Zusammenhängen ist je nach Zielgruppe sehr unterschiedlich. Und auch die persönliche Aufnahmefähigkeit und die Bereitschaft, sich mit dem Thema auseinanderzusetzen, sind unterschiedlich stark ausgeprägt. Der einen Person ist vielleicht schon die Unterscheidung zwischen starker und schwacher Künstlicher Intelligenz zu abstrakt, die andere möchte dagegen die Unterschiede zwischen Machine Learning und Deep Learning definiert wissen.

Stellen Sie sich also auf ihr jeweiliges Gegenüber ein. Ideal ist, wenn Sie gemeinsame Themen finden, eine Schnittmenge an Wissen zwischen der Zielgruppe und sich, und wenn die Zielgruppe Ihre Botschaft versteht.

Überlegungen zur Zielgruppe setzen voraus, sich über die **Ziele** der eigenen Kommunikationsbemühungen im Klaren zu sein. Möchten (oder sollen) Sie vornehmlich ein komplexes Thema wie die KI aus Expertensicht beschreiben, erklären, erläutern? Oder ist Ihre Expertise auch gefragt, um eine Bewertung, eine Einschätzung mit auf den Weg zu geben? Welche Dinge wollen Sie Ihrem Publikum in Form prägnanter Kernbotschaften unbedingt mit auf den Weg geben? Wel-

Welche Tipps zur Risikokommunikation halten Sie für wichtig?

Emotionen ernst nehmen / sich in das Gegenüber hineinversetzen	
Den möglichen Nutzen ausdrücklich zur Sprache bringen	
Kurze, klare Sätze, ohne ein ständiges Wenn und Aber	
Appell: Risiken in Beziehung setzen zu erwartetem Nutzen	
Sprache nicht mit Fachbegriffen überfrachten	
Es geht darum, die große, ungeschlossene Mitte zu erreichen	
Nicht ungefragt eine Interpretationshilfe zur Einschätzung mitliefern	
Innenleben der eigenen Organisation berücksichtigen	
Quantifizierbare Angaben (wie Prozentwerte) veranschaulichen	
Trotz Risikokommunikation Vorbereitung auf Krisen-Kommunikation	

16

che Überlegungen sollten Sie anstoßen? Gibt es Zusammenhänge, bei denen Sie Überzeugungsarbeit leisten möchten? Stellen Sie sich diese Fragen rechtzeitig!

WIE WICHTIG IST EINE KLARE SPRACHE?

Kommunikation funktioniert nur, wenn der Empfänger die Botschaft des Senders auch versteht.

In den seltensten Fällen tauschen sich beide auf der gern zitierten „Augenhöhe“ aus, meist ist mit der Sender-Rolle ein Wissensvorsprung gegenüber dem Empfänger verbunden.

Eine **verständliche und prägnante Sprache** der Sender – hier der Forschenden – ist deshalb von elementarer Bedeutung. Verständliche Sprache

vermeidet Fremdwörter, führt notwendige Fachbegriffe ein und beschreibt veranschaulichend mittels geeigneter Beispiele. Wichtig ist gerade in der Risikokommunikation das Veranschaulichen von Größenordnungen und Prozentangaben – Beispiel: Wenn für die Kölnerinnen und Kölner in einem Fall X das Risiko, dass sie davon betroffen werden, bei 0,001 Prozent liegt, dann bedeutet dies zehn Bewohner und Bewohnerinnen der Großstadt Köln.

Verständlichkeit und Klarheit sind zudem um so wichtiger, als in Zeiten eines „information overload“ häufig die skurrilen und abseitigen Behauptungen hervorstechen. Wollen sich die seriösen Stimmen Gehör verschaffen, braucht es eingängige, prägnante Aussagen, die sich nicht hinter Fachbegriffen verstecken und damit auch noch das Klischee des abgehobenen, unverständlichen Wissenschaftlertypus verstärken.

Wiederholen Sie zentrale Aussagen, die Ihnen wichtig sind. Und ein Argument, das Sie für zentral halten, sollten Sie auch als Erstes Ihrer Argumente präsentieren (und nicht auf eine Gelegenheit warten, es im Diskussionsverlauf „hervorzuholen“). **Das Wichtigste zuerst!**

Dabei bedeutet „klare, prägnante Sprache“ ausdrücklich nicht, den Leserinnen und Lesern, Hörerinnen und Hörern, Besucherinnen und Besuchern Vorgaben zu machen, wie sie ein Risiko einzuschätzen haben. Die Abwägung zwischen konkretem Risiko und konkretem Nutzen treffen mündige Menschen schon gerne selbst. „Klare, prägnante Sprache“ bedeutet zudem nicht, vage Prognosen und Mutmaßungen abzugeben. Erweisen sich diese als falsch, beschädigen solche Aussagen das vorhandene Vertrauen sehr nachhaltig. Manchmal ist es ehrlicher und richtiger zu sagen: „Wir wissen es nicht.“

TIPPS ZUR RISIKOKOMMUNIKATION VON KI – DIE BESTEN DREI!

In unseren Workshops haben wir die Teilnehmerinnen und Teilnehmer jeweils gebeten, aus einer Liste von zehn Tipps⁴ die aus ihrer Sicht wichtigsten **drei Anregungen** zur Risikokommunikation im Zusammenhang mit KI auszudeuten (siehe Grafik S.16). Daraus haben wir dann eine Gesamtwertung erstellt. Das Rennen machten dabei die folgenden drei Hinweise.

- 1) Emotionen ernst nehmen, sich in das Gegenüber hineinversetzen und die Dinge auch mal aus dessen Perspektive sehen.
- 2) Nicht immer nur über Risiken reden, sondern auch den möglichen Nutzen ausdrücklich zur Sprache bringen.
- 3) Sich in kurzen, klaren, gut verständlichen Sätzen ausdrücken, und zwar ohne ständiges Wenn und Aber.

Dies sind in der Tat wichtige Anregungen, und wer sie in der Debatte über KI beherzigt, macht schon mal ziemlich viel richtig.

LITERATUR

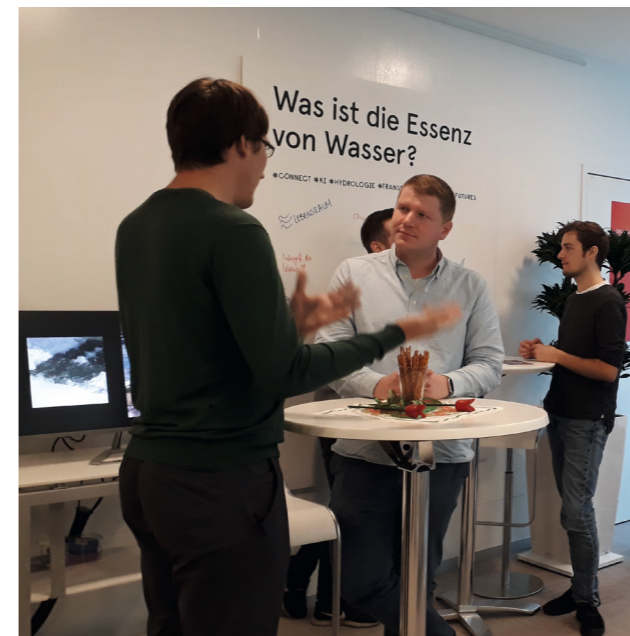
1. acatech & Körber Stiftung: Technikradar 2018. S.6–8 (25/05/2018) <https://www.acatech.de/publikation/technikradar-2018-was-die-deutschen-ueber-technik-denken/>.
2. Kirchner, Stefan: Zeit für ein Update. Was die Menschen in Deutschland über Digitalisierung denken. S.11–12 (Friedrich-Ebert-Stiftung, 07/2019). <http://library.fes.de/pdf-files/fes/15549.pdf>.
3. nextMedia.Hamburg & Statista: Media Innovation Report: So denken die Deutschen über Künstliche Intelligenzen. (13/08/2019) <https://www.nextmedia-hamburg.de/so-denken-die-deutschen-ueber-kuenstliche-intelligenzen/>.
4. Wingen, Klaus: Risiko kommunizieren – zehn Tipps zur Bewältigung einer heiklen Gratwanderung. Wissenschaftskommunikation.de (11/02/2019) <https://www.wissenschaftskommunikation.de/risiko-kommunizieren-zehn-tipps-zur-bewaeltigung-einer-heiklen-gratwanderung-23537/>.

-> Alle hier genannten Webseiten geprüft am: 24.01.2020

17



1



5



6



2



3



7



4



8

1. Das Thema Autonomes Fahren stand im Mittelpunkt der öffentlichen Diskussion in Karlsruhe im Juli 2019. Bürgerinnen und Bürger tauschten sich im sogenannten Fishbowl-Format mit Expertinnen und Experten des Karlsruher Instituts für Technologie (KIT) und des FZI Forschungszentrum Informatik aus. V.l. Bürger, Prof. Dr.-Ing. Anne Koziolok, Christian Hubschneider, Eva Wolfangel (Moderation), Prof. Dr. Marius J. Zöllner, Bürgerin.
2. Forschende am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Saarbrücken bei einer Workshop-Übung in Kleingruppen.
3. Videoaufzeichnung am Tübinger AI Center. Das Kamerteam der Filmproduktionsfirma TeamWERK fängt Kurzstatements einzelner Forscherinnen und Forscher zu Risikothemen zur Künstlichen Intelligenz ein.
4. Forschende am DFKI Saarbrücken diskutieren über kritische Fragen und mögliche Antworten im Zusammenhang mit Künstlicher Intelligenz.
5. Im RIKI-Workshop mit Forschenden der Fraunhofer-Gesellschaft im Fraunhofer-Forum Berlin stand das Thema Medizin und KI im Mittelpunkt.
6. Dozent Klaus Wingen präsentiert den Workshopteilnehmerinnen und -teilnehmern am DFKI Saarbrücken zehn Tipps zur Risikokommunikation.
7. Das Thema Künstliche Intelligenz in der Medizin weckte bei den Bürgerinnen und Bürgern in der öffentlichen Diskussion in Berlin im Oktober 2019 hohes Interesse. Als Expertinnen und Experten beteiligten sich Dr. Andrea Schenk vom Fraunhofer MEVIS (1.v.l.), Dr. med. Alexander Meyer des Deutschen Herzzentrums Berlin (2.v.l.) und Prof. Dr. Thomas Berlage vom Fraunhofer FIT (4.v.l.). Die Moderation hatte Nicola Kuhrt (3.v.l.).
8. Angelehnt an eine Partysituation tauschen sich die Forscherinnen und Forscher während des RIKI-Workshops am Tübinger AI Center über kritische KI-Themen aus.

KI UND UMGANG MIT DATEN

1. MEDIALE BERICHTERSTATTUNG

Wie ein roter Faden zieht sich ein Thema seit Jahrzehnten durch die Berichterstattung über Informationstechnologie: Der Umgang mit den Daten, die wir in einem bisher nie dagewesenen Ausmaß erheben, sammeln und auswerten. In Zeiten von Big Data geht es um Fragen der Datenhoheit, Datensicherheit, Überwachung, Diskriminierung bis hin zu Datenmanipulationen und Fake News.

Daten haben Google, Facebook und Amazon zu mächtigen Unternehmen gemacht. „Heute werden Daten zu Gold gesponnen“, schreibt Adrian Lobe auf Zeit Online.¹

Und im Zusammenhang mit dem Skandal um das Unternehmen Cambridge Analytica, das illegal über 85 Millionen Facebook-Profile ausgewertet und politische Kampagnen wie den Brexit und die Wahl von Donald Trump beeinflusst hat, titelt Patrick Wellinski im Deutschlandfunk: „Wenn Daten die Welt manipulieren“.²

Daten sind in immer größerer Zahl vorhanden und werden in immer mehr Lebensbereichen gesammelt:

- Im Internet: Die Datenspuren, die wir zum Beispiel bei Kaufentscheidungen oder Suchanfragen hinterlassen, sind das oben genannte neue Gold.
- Im automatisierten Auto: Dort werden permanent Daten erfasst, die zur Klärung eines möglichen Unfalls herangezogen werden können.³
- Zu Hause: Einer Umfrage der Postbank zufolge

nutzen ein Drittel der in Deutschland Befragten einen Sprachassistenten.⁴

- In der Medizin: Mit der elektronischen Patientenakte sollen bald alle Gesundheitsdaten überall abrufbar werden.⁵
- Im Arbeitsalltag: Dort werden Daten erfasst, um die Automatisierung von Arbeitsprozessen voranzutreiben.

Die Sorge, dass KI-Systeme zur **Überwachung** und Kontrolle der Bürgerinnen und Bürger verwendet werden könnten, ist groß.⁶ In China sind Straßen mit Überwachungskameras gesäumt. Gesichtserkennung wird laufend eingesetzt, um Daten auch für das dortige Social Scoring System zu erfassen. Medien mahnen: „Ist es der Weg zur totalen Überwachung?“⁷

Zudem leisten KI-Systeme immer mehr. Es gibt Gangerkennungssysteme, die uns identifizieren könnten, selbst wenn wir unser Gesicht bedecken würden. Solche werden beispielsweise in der Uiguren-Region Xinjiang bereits eingesetzt.⁸ Und KI kann Stimmungslagen in unseren Stimmen erkennen. „Amazon hat bereits ein Patent darauf angemeldet, Emotionen und sogar Krankheiten aus der Stimme herauszulesen“, berichtet Eva Wolfangel in der Zeit.⁹

Parallel zu diesen Entwicklungen ist nicht immer klar, welche Daten Internetkonzerne oder staatliche Einrichtungen erheben und speichern – eine zentrale Frage der **Datenhoheit**. Große mediale Aufmerksamkeit hatten 2019 beispielsweise Sprachassistenten wie Siri oder Google Assistant, als herauskam, dass Aufnahmen der Geräte von Menschen abgehört wurden, um ihre Funktionen zu optimieren. Die Mitarbeiter und Mitarbeiterinnen von Drittunternehmen sollten erkennen, wann die Assistenten etwas nicht richtig verstanden haben.¹⁰ Dies ist nur ein Beispiel von vielen

dafür, dass für die Nutzerinnen und Nutzer oftmals nicht einsichtig ist, welche Daten von Internetkonzernen gespeichert werden und wer darauf zugreifen kann.^{3,10}

Insbesondere die Frage der Datenhoheit bei der Speicherung und Nutzung medizinischer Daten wird in Medien und Politik häufig thematisiert.

„DER GLÄSERNE MENSCH WIRD DURCH KI NOCH TRANSPARENTER“

Titelzeile Artikel von Peter Buxmann, Süddeutsche Zeitung²¹

Sehr oft geht es in der medialen Berichterstattung auch um mögliche **Diskriminierungen** durch die Analyse von Daten und Entscheidungen auf Basis von Algorithmen.

Gezielt könnte etwa eine Versicherung krankhaft übergewichtigen Menschen Prämienrabatte geben, wenn diese sich per Sprachassistent nur Gemüse nach Hause liefern lassen.¹¹

In manchen Ländern werden schon heute Interessierten für Hotel- oder Flugbuchungen per „Individual Pricing“ unterschiedliche Preise je nach errechneter Zahlungskraft angezeigt – ermittelt anhand von Faktoren wie Wohnort, verwendeter IT oder besuchter Webseiten.¹²

Amazon hatte experimentell in Bewerbungsverfahren ein Einstellungs-Tool auf der Basis von KI eingesetzt. Es konnte Bewerbenden Noten von eins bis fünf geben. Aber im Jahr 2015 erkannte

das Unternehmen, dass sein System die Kandidatinnen und Kandidaten für eine Anstellung in der Softwareentwicklung nicht geschlechtsneutral bewertete, sondern männliche Kandidaten bevorzugte. Das lag an den Trainingsdaten der Musterlebensläufe von mehrheitlich männlichen Bewerbern.¹³

Die Organisation Algorithmwatch hat Projekte aus der Europäischen Union zusammengetragen, in denen auch staatlicherseits Systeme mit automatischer Entscheidungsfindung eingesetzt werden. In der belgischen Region Flandern werde beispielsweise das Klickverhalten von Jobsuchenden ausgewertet, um herauszufinden, ob sie für mangelnde Aktivität bei der Suche bestraft werden müssen, so ein Report auf Heise Online.¹⁴ In vielen Ländern werden demnach außerdem Systeme zum Predictive Policing getestet oder gar eingesetzt.

Eine solche Software mit dem Namen COMPAS wurde auch in den USA eingesetzt, um das Risiko einer erneuten Straffälligkeit in Bewährungsverfahren einzuschätzen. Journalistinnen und Journalisten von ProPublica deckten bereits 2016 auf, dass die Software dabei Entscheidungen mit einem rassistischen Bias traf.^{15,16}

Und nicht zuletzt sind **Datenmanipulationen** – wie etwa veränderte Bilder oder Videos – ein großes KI-Thema. In den Medien werden auch Bedenken aufgegriffen, dass **Deep Fakes** die Glaubwürdigkeit verfügbarer Informationen schwächen und dazu eingesetzt werden können, um die öffentliche Meinung zu manipulieren.¹⁷ So wird in der Berichterstattung kritisch über gefälschte Videos berichtet, die Menschen persönlich angreifen oder falsche Informationen verbreiten.¹⁸

Um Fragen der Datenhoheit zu klären, Datensicherheit in geeignetem Umfang zu gewährleisten und Diskriminierungen oder Datenmanipulationen entgegenzuwirken, ist ein **ethischer Umgang** mit KI nötig. Sowohl auf nationaler¹⁹ als auch auf europäischer²⁰ Ebene haben Expertengruppen ethische Leitlinien entwickelt, um den Interessen von Nutzerinnen und Nutzern im Umgang mit KI gerecht zu werden.

2. AUS SICHT DER FORSCHENDEN

Im Rahmen der RIKI-Workshops wurde der Umgang mit Daten vor allem an den Standorten Saarbrücken, Kaiserslautern, Bremen und Tübingen angesprochen.

In den Workshops betonten die Forschenden, wie wichtig große Datensätze seien, um KI-Systeme zu trainieren. Das enorme Potenzial von Künstlicher Intelligenz liege darin, dass das System anhand von Daten selbst dazulerne und es keine feste Befehlskette aus Programmbefehlen mehr brauche. Aus Sicht der Forschenden ist es wichtig, der Ge-

sellschaft transparent Auskunft darüber zu geben, wofür Daten gesammelt werden. Denn technisch gesehen sei **Überwachung** durchaus möglich.

Ihrer Ansicht nach muss man Kosten und Nutzen individuell und für jede Situation neu abwägen. Als ein Beispiel für solch ein Abwägen von Vor- und Nachteilen wurde der Einsatz von KI im Schulsystem genannt. Bild- und Sensoraufnahmen im Klassenzimmer kombiniert mit KI sind einerseits problematisch, könnten aber andererseits womöglich dazu beitragen, Schülerinnen und Schüler gezielt und individuell zu fördern.

Gerade mit Blick nach China erscheint es den Forschenden notwendig, sich über Regelungen Gedanken zu machen. Sie sehen Gesellschaft und Politik gefordert. Diese müssten definieren, wer welche Daten sammeln und einsehen dürfe. Dabei sind sie selbst Teil der Gesellschaft und können Diskurse anregen.

0110101110101011010110

Um auszuschließen, dass man durch die gesammelten Daten zum gläsernen Menschen wird, sei die **Privatsphäre** ein wichtiges Thema, urteilten die Forschenden. Sie erklärten, dass Nutzende die Möglichkeit erhalten müssen, selbst zu entscheiden, ob und welche Daten anderen zur Verfügung stehen. Nur so könne das Vertrauen von Verbraucherinnen und Verbrauchern für KI gewonnen werden und erhalten bleiben. Neben den rechtlichen Maßnahmen seien auch technische Maßnahmen zu treffen. Dazu nannten die Forschenden drei zentrale Punkte:

Erstens sei bei jeder KI zu überlegen, ob es für die jeweilige Anwendung nicht genügt, wenn die dabei entstehenden Daten **lokal** gespeichert werden oder, ob sie wirklich in eine Cloud hochgeladen werden müssen. Aus Sicht der Forschenden ist es oftmals nicht notwendig, Daten permanent mit der Umgebung auszutauschen. Als Beispiel verwiesen sie auf ein KI-Überwachungssystem bei Pflegebedürftigen. Hier genüge, solange es zu keinen akuten Situationen kommt, ein Update am Tag, das auch als Information an Angehörige gehen kann.

Zweitens forderten die Wissenschaftlerinnen und Wissenschaftler eine sogenannte **Datendiät**. Aus ihrer Sicht ist es sinnvoll, nur Daten zu erfassen, die für die jeweils zu bewältigende Aufgabe tatsächlich notwendig sind. Sehr vereinfacht: Für eine Unterstützung bei Alltagsarbeiten in der Küche benötigt das System keine Daten aus dem Bad. Zudem sprachen sich die Forschenden dafür aus, nicht benötigte Daten regelmäßig und dadurch zeitnah zu löschen.

Drittens sehen es die Forschenden als notwendig an, bindende Regeln für den **Datenzugriff** festzulegen. Es sei wichtig, dass die Nutzenden frei darüber entscheiden können, wem sie die Daten zur Verfügung stellen. Wer darf auf persönliche Daten zugreifen? So sei es nicht wünschenswert,

wenn die Krankenversicherung die individuellen Gesundheitsdaten von Patientinnen und Patienten einsehen könne. Die Forschenden rieten daher, solche Daten anonym zu verarbeiten. Damit bliebe für die Krankenversicherung der gesamte Datenschatz für allgemeine Auswertungszwecke erhalten und zugleich die Datenhoheit bei den Versicherten.

Insbesondere in der KI-Forschung spielen Daten eine zentrale Rolle. Denn je höher die Menge der eingesetzten Daten ist, desto bessere Trainingsergebnisse können KI-Systeme erzielen. Ein Beispiel dafür ist die Mustererkennung in der medizinischen Diagnostik. Hier wünschen sich die Forschenden den Zugriff auf die hochsensiblen Daten. Die Privatsphäre der Patientinnen und Patienten könne dabei ähnlich wie beim Vorschlag für die Krankenversicherungen durch Anonymisierungen und Pseudonymisierungen geschützt werden.

100101010110001110101011010

Auf die Frage, inwieweit und in welchen Bereichen durch den Einsatz von KI **Entscheidungen** getroffen werden können oder sollen, betonten die Forschenden, dass Künstliche Intelligenz dafür in noch keinem Bereich verlässlich und gut genug entwickelt sei. Jede Empfehlung einer KI müsse letztlich noch durch einen Menschen überprüft werden. So hielten es viele Forschende für kritisch, einer KI die Entscheidung über eine Kreditvergabe zu überlassen. Insbesondere dann, wenn der angewendete Algorithmus nicht wirklich nachvollziehbar erscheint.

Dennoch halten es einige Forschende für denkbar, Künstlicher Intelligenz in manchen Bereichen sukzessive die Kontrolle übergeben zu können. Dann müsse die Entscheidung des jeweiligen KI-Systems aber zwingend **nachvollziehbar, fair** und **sicher** sein.

<p>Wie geht man mit Vorurteilen in Datensätzen um?</p> <p>Lilli Smal DFKI, Saarbrücken</p>	<p>Werden Lehrer ihre Schüler mit KI überwachen?</p> <p>Dr. Nicolas Großmann DFKI Kaiserslautern</p>
<p>Führt KI zum Überwachungsstaat?</p> <p>Alexander Meinke Eberhard Karls Universität Tübingen</p>	<p>Wird KI den Menschen automatisch klassifizieren?</p> <p>Sabine Klein DFKI, Saarbrücken</p>

Videostatements von Forschenden sind online abrufbar.²¹

Für die **Nachvollziehbarkeit** müsse das System transparent und die Entscheidung erklärbar sein, so die Forschenden. Sei ein Algorithmus transparent, dann sei auch bekannt, auf Basis welcher Merkmale er zu einer Entscheidung gelange. Könnten sich zudem alle Beteiligten diese Entscheidung plausibel erklären, dann sei das Kriterium der Nachvollziehbarkeit erfüllt.

Mit der Forderung nach Transparenz sehen die Forschenden potenziell Eingrenzungen für ihre Forschung. Denn insbesondere leistungsstarke Algorithmen seien häufig nicht nachvollziehbar, wie eine Black Box. Eine potenzielle Lösung sehen die Forschenden darin, einen zweiten Algorithmus nachzuschalten, der den Entscheidungsmechanismus des eigentlichen Algorithmus transparent machen könnte.

Was die **Fairness** anbelangt, so stellten die Forschenden heraus, dass sowohl die eingesetzten Datensätze als auch die jeweiligen KI-Methoden die Fairness eines KI-Systems beeinträchtigen können.

So würden KI-Systeme häufig mit Datensätzen trainiert, die in der Vergangenheit gesammelt wurden. Dementsprechend könnten die Datensätze veraltete Ansichten und Werte auf das KI-System übertragen. Auch warnten sie davor, KI-Systeme beispielsweise mit Datensätzen zu trainieren, die nur einen begrenzten Personenkreis repräsentieren.

Doch selbst wenn die Trainingsdaten auf potenzielle Fehler oder Daten, die Vorurteile befördern könnten, überprüft würden, sei nicht gesagt, dass das KI-System am Ende faire Entscheidungspfade aufstellt. Denn viele Merkmale sind miteinander verknüpft und geben dem KI-System so auch indirekt Hinweise, beispielsweise zum Geschlecht. Schließe man das Geschlecht als Entscheidungsmerkmal aus, die Berufserfahrung aber nicht, könnte ein System diejenigen benachteiligen, die

einen lückenhaften Lebenslauf haben – etwa Frauen, die auch heute noch längere und mehr Erziehungspausen im Lebenslauf haben als Männer.

3. KOMMUNIKATIONSTIPPS VON FORSCHENDEN FÜR FORSCHENDE

- Kann KI zur Überwachung eingesetzt werden? Führt KI zum Überwachungsstaat? Solche und ähnliche Fragen werden häufig gestellt. Antworten Sie darauf sehr transparent: Es ist technisch durchaus möglich, dass Unternehmen oder auch der Staat – siehe China – Mustererkennungssysteme zur Überwachung einsetzen.
- Sensibilisieren Sie Laien auch proaktiv für unerwünschte Entwicklungen und regen Sie Diskussionen über mögliche Regeln für Unternehmen an.
- Bemühen Sie sich um eine ausgewogene Darstellung bezüglich einer breiten Datenerfassung. Kosten und Nutzen sind individuell gegeneinander abzuwägen. Veranschaulichen Sie dies mit einem Beispiel aus der Lebenswelt Ihres Gegenübers.
- Heben Sie hervor, dass es sich bei den Ergebnissen von KI-Systemen stets um Wahrscheinlichkeiten handelt. Das Ergebnis einer KI sollte deshalb nicht als ein eindeutig analytisches Ergebnis eingestuft werden.
- Sensibilisieren Sie Ihr Gegenüber dafür, dass auch KI-Systeme und deren Ergebnisse fehlerhaft sein können. Erstens können Fehler und Mängel in Datensätzen die Ergebnisse eines KI-Systems beeinflussen. Hier ist es wichtig sicherzustellen, dass Verzerrungen nicht die Fairness des Algorithmus und dessen Entscheidungen beeinflussen. Zweitens können Algorithmen nicht ausreichend robust gegenüber Schwankungen in den eingegebenen Daten sein. Kleine Schwankungen sollten das Ergebnis kaum verändern. So sollte ein Algorithmus der Huskys erkennt, einen Husky

auch dann erkennen, wenn kein Schnee im Bild zu sehen ist. Oder ein Stoppschild müsse von einem automatisierten Fahrsystem auch erkannt werden, wenn auf ihm ein Aufkleber Teile der Oberfläche bedeckt.

- Will man einer KI die Entscheidungshoheit überlassen, muss die Entscheidung nachvollziehbar, fair und sicher sein. Erklären Sie konkret, was Sie dazu beitragen, damit diese Kriterien erfüllt werden.
- Nutzen Sie Beispiele aus der Lebenswelt Ihres Gegenübers, um hervorzuheben, warum die Entscheidungen einer KI nachvollziehbar sein müssen. Ein mögliches Beispiel: Bei der Auswahl von Bewerbungsunterlagen muss transparent sein, auf Basis welcher Merkmale die KI Bewerberinnen und Bewerber aussucht. Zudem sollte bei jeder einzelnen Bewerbung erklärbar sein, warum der Algorithmus eine konkrete Bewerbung für gut oder nicht gut befindet.
- Zeigen Sie beim Thema Privatsphäre, wie wichtig Ihnen diese ist. Nennen Sie mögliche Maßnahmen, die ergriffen werden können, um persönliche Daten zu schützen. Wer Daten lokal speichert, minimiert beispielsweise Sicherheitsrisiken, die mit dem Datentransfer verbunden sind.
- Ein weiterer Aspekt im Zusammenhang mit der Privatsphäre ist die Frage, wer auf welche Daten zugreifen darf. Regen Sie einen gesellschaftspolitischen Diskurs darüber an. Weisen Sie darauf hin, dass rechtliche Rahmenbedingungen idealerweise auf internationaler Ebene festgelegt werden müssen.

LITERATUR

1. Lobe, Adrian: Rage Against the Machine. Zeit Online (21/01/2019) <https://www.zeit.de/kultur/2019-01/kuenstliche-intelligenz-widerstand-angriffe-usa/komplettansicht?print>.
2. Wellinski, Patrick: Wenn Daten die Welt manipulieren. Deutschlandfunk Kultur (23/07/2019) https://www.deutschlandfunkkultur.de/netflix-doku-ueber-cambridge-analytica-wenn-daten-die-welt.1013.de.html?dram:article_id=454606.
3. Müller, Christina: Was das Auto über seinen Fahrer weiß. Süddeutsche Zeitung (24/04/2019) <https://www.sueddeutsche.de/auto/>

[daten-auto-blackbox-unfall-1.4418337](https://www.sueddeutsche.de/auto/daten-auto-blackbox-unfall-1.4418337).

4. Postbank: Digitalstudie 2019. Untersuchung zur Internet- und Mobilnutzung der Bevölkerung.. (06/2019) https://www.postbank.de/postbank/pr_presseinformation_2019_06_12_studie_ein_drittel_der_deutschen_spricht_mit_alex_siri_und_co.html.
5. Hegemann, Lisa: Wie sicher sind meine medizinischen Daten? Zeit Online (28/12/2018) <https://www.zeit.de/digital/datenschutz/2018-12/elektronische-patientenakte-medizinische-daten-sicherheit-zweifel-datenschutz>.
6. Hegemann, Lisa: Eine Frage der Ethik. Zeit Online (08/04/2019) <https://www.zeit.de/digital/internet/2019-04/kuenstliche-intelligenz-eu-kommission-richtlinien-moral-kodex-maschinen-ethik>.
7. Dorloff, Axel: Auf dem Weg zur totalen Überwachung. tagesschau.de (24/03/2019) <https://www.tagesschau.de/ausland/ueberwachung-china-101.html>.
8. Dorloff, Axel: Künstliche Intelligenz als Staatsziel. Deutschlandfunk (11/02/2019) https://www.deutschlandfunk.de/china-kuenstliche-intelligenz-als-staatsziel.724.de.html?dram:article_id=440743.
9. Wolfangel, Eva: Die Seele auf der Zunge. Die Zeit Nr.07/2019 (07/02/2019) <https://www.zeit.de/2019/07/stimme-biometrie-messbarkeit-emotionen-persoenlichkeit>.
10. Reichelt, Patrick & Hegemann, Lisa: Alexa, hörst du grad zu? Zeit Online <https://www.zeit.de/digital/datenschutz/2019-09/smart-speaker-amazon-echo-google-home-daten>.
11. Spiekermann, Sarah: Der Mensch als Fehler. Süddeutsche Zeitung (23/03/2019) <https://www.sueddeutsche.de/kultur/kuenstliche-intelligenz-ethik-menschenbild-philosophie-1.4378898>.
12. Beck, Susanne et al.: Künstliche Intelligenz und Diskriminierung – Whitepaper aus der Plattform Lernende Systeme, München 2019. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG3_Whitepaper_250619.pdf.
13. Dastin, Jeffrey: Amazon scraps secret AI recruiting tool that showed bias against women. Reuters (10/10/2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.
14. Holland, Martin: Künstliche Intelligenz: Überall in Europa entscheiden schon Algorithmen. Heise Online (29/01/2019) <https://www.heise.de/newsticker/meldung/Kuenstliche-Intelligenz-Ueberall-in-Europa-entscheiden-schon-Algorithmen-4290676.html>.
15. Angwin, Julia; Larson, Jeff et al.: Machine Bias. ProPublica (23/05/2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
16. Herzog, Lisa: Glückwunsch, Sie haben die KI überzeugt! Zeit Online (13/06/2019) <https://www.zeit.de/arbeit/2019-05/kuenstliche-intelligenz-arbeitsplatz-fairness-algorithmen-diskriminierung>.
17. Hurtz, Simon: Echt falsch. Süddeutsche Zeitung (06/12/2019) <https://www.sueddeutsche.de/politik/kuenstliche-intelligenz-echt-falsch-1.4711761>.
18. Tschopp, Marisa: Kein Vertrauen in KI? Das steckt hinter der Angst vor Künstlicher Intelligenz. Maschinenmarkt (22/10/2019) <https://www.maschinenmarkt.vogel.de/kein-vertrauen-in-ki-das-steckt-hinter-der-angst-vor-kuenstlicher-intelligenz-a-876226/>.
19. Datenethikkommission der Bundesregierung: Gutachten der Datenethikkommission. (23/10/2019) https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6.
20. High-Level Expert Group on Artificial Intelligence set up by the European Commission: The Ethics Guidelines for Trustworthy Artificial Intelligence. <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>.
21. Buxmann, Peter: Der gläserne Mensch wird durch KI noch transparenter. (03/12/2018) <https://www.sueddeutsche.de/digital/kuenstliche-intelligenz-algorithmen-ethik-1.4235816>

→ Alle hier genannten Webseiten geprüft am 24.01.2020

KI UND ARBEIT

1. MEDIALE BERICHTERSTATTUNG

„Jobverlust durch KI“¹, „Roboter werden immer mehr Menschen bei der Arbeit ersetzen“² – das sind Schlagzeilen, die wir nicht gerne lesen. Die Frage, inwieweit KI Arbeitsplätze vernichtet, ist gesellschaftlich von zentraler Bedeutung und wird daher häufig in den Medien angesprochen. Ebenso wie die weiterführende Frage, in welchem Ausmaß KI möglicherweise neue Arbeitsplätze schafft. Diese beiden Perspektiven werden auch in der medialen Berichterstattung häufig eingenommen.

Mit einem optimistischen Blick, bietet das Zusammenspiel zwischen Mensch und KI vielseitige Chancen, um die Arbeit aufzuwerten und damit die Arbeitsqualität zu verbessern.³ Mittels KI-Unterstützung wird Arbeit effizienter.⁴ Das bedeutet unter Umständen auch mehr Freizeit für die Menschen, wenn KI Teile der Arbeit erledigt. Um gesamtgesellschaftlich das Steueraufkommen zu sichern, könnte eine Maschinensteuer erhoben werden. Allgemein könnten KI-Maschinen Arbeiterinnen und Arbeiter produktiver und wertvoller machen, was die Nachfrage nach ihnen und damit auch die Löhne steigen lasse.⁵ „Maschinen schaffen mehr Jobs als sie vernichten“, titelte Alexander Hagelüken gar in der Süddeutschen Zeitung.⁶

Mit einem pessimistischen Blick ist umgekehrt von massiven Verwerfungen auf dem Arbeitsmarkt durch weitreichende Verdrängung von Arbeitskräften aufgrund des Einsatzes von KI auszugehen.³ Laut dem Eurobarometer 2017 der Europäischen Kommission sind 74 Prozent der in

Deutschland Befragten überzeugt, dass durch den Einsatz von Robotern und KI mehr Jobs wegfallen als neue entstehen und 72 Prozent glauben, KI werde Menschen den Arbeitsplatz wegnehmen.⁷ Überraschend pessimistisch ist laut dem Magazin t3n eine Studie der Oxford-Wissenschaftler Carl Benedikt Frey und Michael A. Osborne aus dem Jahr 2013 ausgefallen, die sogar „Bürotätigkeiten, Dienstleistungen und vermeintlich kreative Tätigkeiten“ als gefährdet einstufte.^{8,9}

Egal, welche Zukunftsperspektive man einnimmt, Tatsache ist, dass bereits heute viele Firmen KI-Anwendungen nutzen, um ihre Prozesse zu optimieren. Deren Einsatz ist dabei nicht nur auf Logistik, Produktion oder Marketing beschränkt. Vielmehr halten sie Einzug in alle primären und unterstützenden Aktivitäten von Unternehmen entlang der Wertschöpfungskette.³ Aus Sicht der Wirtschaft sind KI-gesteuerte Prozesse präzise, ausdauernd und finanziell gut kalkulierbar. KI streikt nicht, um Lohnforderungen durchzusetzen. KI bekommt keinen Bandscheibenvorfall, macht keinen Urlaub, braucht keine freien Tage. Sie ist immer verfügbar.

Ein Beispiel ist der Wandel im Supermarkt. Walmart setzt in 50 Supermärkten mehr als 2.000 Roboter ein, um die Regale nach ausverkauften Artikeln, falschen Warenauszeichnungen und anderen Dingen zu durchsuchen, was normalerweise Aufgabe des Verkaufspersonals wäre.¹⁰ Die Metro-Gruppe untersucht maschinelles Sehen an der Kasse: Die Warenkörbe der Kunden werden von Kameras aufgezeichnet und entsprechende Rechnungen ausgestellt. Nach Schätzungen können die unbemannten Kassen etwa 50 Kunden pro Stunde – und damit etwa doppelt so viel wie menschliches Kassenpersonal – bedienen.¹⁰

01010010010101110101011010

„KÜNSTLICHE INTELLIGENZ: SIND UNSERE ARBEITSPLÄTZE NUN BEDROHT ODER NICHT?“

Titelzeile Artikel von Nick Gretzinger auf Ingenieur.de¹¹

010100101010100000011111010

Der Wandel findet also bereits statt. Entscheidend ist, wen er in besonderem Maße trifft. Tätigkeiten, die Kreativität und vorausschauendes Denken erfordern, werden weniger betroffen sein, als Tätigkeiten, die viele repetitive Arbeitsschritte enthalten. Maschinen können besser Wiederholungstätigkeiten abspulen und riesige Datenmengen analysieren, stoßen jedoch schnell an ihre Grenzen, sobald Kreativität und Interaktionen mit Menschen gefragt sind – denn hier sind Empathie und Emotionen am wichtigsten.⁴

Es ist daher sehr wahrscheinlich, dass zumindest für einige Jahre oder Jahrzehnte die Arbeitslosigkeit – nicht nur in Deutschland – enorm steigen wird, weil viele Berufe durch Maschinen und Computer ersetzt werden. Insbesondere die Politik muss in den nächsten Jahren gute Voraussetzungen schaffen, damit die Menschen, die ihren Job tatsächlich verlieren werden, wieder in den Arbeitsmarkt integriert werden können.¹¹ Carl Benedikt Frey von der Oxford Martin School der Oxford University schreibt in einem Gastbeitrag in der Süddeutschen Zeitung: „Um starke Gegenbewegungen zur Automatisierung zu vermeiden, müssen Regierungen eine Politik verfolgen, die das Produktivitätswachstum ankurbelt, aber gleichzeitig die Arbeitnehmer dabei unterstützt, sich auf die Automatisierungswelle einzustellen.“¹²

0101011100101011101010111010

KI verändert nicht nur Arbeitsplätze, sondern auch Bewerbungsverfahren. Bereits vor oder während eines Bewerbungsgesprächs analysieren zunehmend auch Algorithmen potenzielle Arbeitskräfte. Anhand von Mimik, Wortwahl und Tonfall bewertet beispielsweise der Algorithmus von HireVue die Berufseignung und kategorisiert die Bewerbungen.¹² Auch hier liegen Vor- und Nachteile eng beieinander. Die Hoffnung ist, dass eine KI neutrale analytische Entscheidungen anhand mehrerer Faktoren treffen kann.

Jedoch bevorzugte eine von Amazon eingesetzte KI männliche Bewerber.¹³ Diese KI wirkt sich für Frauen im Bewerbungsverfahren diskriminierend aus. Es gilt wiederum, dass eine KI nur so gut ist, wie die Daten, mit denen sie trainiert wurde (mehr dazu siehe Kapitel Umgang mit Daten S.20).

2. AUS SICHT DER FORSCHENDEN

Im Rahmen der RIKI-Workshops wurde die Frage, wie Künstliche Intelligenz den Arbeitsmarkt verändern wird, an allen sechs Standorten von den Forschenden diskutiert. Auch dies belegt die enorme Bedeutung des Themas.

Die Wissenschaftlerinnen und Wissenschaftler sind sich darin einig, dass Künstliche Intelligenz den Arbeitsmarkt stark verändern wird. Sie sprechen von der nächsten technologischen Revolution. Die Forschenden rechnen damit, dass die zu erwartenden Umwälzungen des Arbeitsmarktes nicht ohne Reibungspunkte vonstatten gehen werden. Daher appellieren sie an die Politik, frühzeitig darüber nachzudenken, wie die Veränderungen abgefedert werden könnten.

Bedroht KI Arbeitsplätze?

Konstantin-Klemens Lurz
Eberhard Karls
Universität Tübingen



Kostet KI Arbeitsplätze?

Guillermo
Reyes Fuentes
DFKI, Saarbrücken



Wie verändert KI Arbeitsplätze? – Ein Beispiel

Tanja Simeonovski
DFKI Saarbrücken



Videostatements von Forschenden sind online abrufbar.¹⁴

Als mögliche Lösungsansätze nannten die Forschenden die Einführung eines **Grundeinkommens** oder einer **KI-Steuer**. Wichtig ist ihnen, dass die Gewinne, die eine KI erwirtschaftet, insbesondere denjenigen Arbeitnehmerinnen und Arbeitnehmern zugutekommen, die ihren Job durch KI verlieren.

0101011110101010011110

Bei den Diskussionen in den Workshops wurde sowohl über Nachteile als auch über Chancen gesprochen, die der Einsatz von KI bringen kann.

Der offenkundigste Nachteil, den die Forschenden im Einsatz von KI im Arbeitsalltag sehen, ist der **Wegfall von Arbeitsplätzen**. Sie schätzen, dass repetitive Arbeiten und Tätigkeiten, die wenig Schulung erfordern, am stärksten von den Ver-

änderungen betroffen sein könnten – Fließbandarbeiten in der Industrie etwa. Aber auch hoch qualifizierte Arbeitskräfte, wie Investmentbanker oder Teile der Filmindustrie, könnten durch die fortschreitende Entwicklung von KI ersetzt werden.

Der Einsatz von KI berge jedoch auch Chancen, betonen die Forschenden. Die Entwicklungen hätten das Potenzial, den **Menschen zu entlasten** und ihn zu **unterstützen**. Pflegekräften könnte so zum Beispiel die aufwendige Dokumentation erleichtert werden, folglich hätten sie mehr Zeit für die Patientinnen und Patienten.

Viele Teilnehmerinnen und Teilnehmer zeigten sich überzeugt, dass durch KI die **Wochenarbeitszeit reduziert** und die **Sicherheit und Qualität** vieler Arbeitsprozesse **erhöht** werden könnte.

Einzelne Forschende verweisen darauf, dass KI auch neue Nischen erschließen wird. Aufgaben, die für den Menschen bislang nicht lukrativ genug oder zu komplex gewesen seien, könnten von KI erledigt werden. Ein Beispiel: Elektronische Simultanübersetzer für seltene Sprachen.

01010111101010101010111010

Gerne orientieren sich die Forschenden bei ihren Überlegungen an dem Ideal, dass jede Arbeitnehmerin und jeder Arbeitnehmer die Möglichkeit hat, einen sinnstiftenden Job auszuüben. Damit verbinden sie die Erwartung, dass für wegfallende Arbeitsplätze äquivalente neue Jobs entstehen. Und zwar im Umfeld der KI-Forschung und -Betreuung, denn KI-Systeme müssten entwickelt, trainiert und überprüft werden. Auch müsse es Menschen geben, die die Datensätze, mit denen ein Algorithmus trainiert wird, vorbereiten.

Die Mehrheit der Forschenden ist sich aber dessen bewusst, dass damit **Umschulungen und Weiter-**

bildungen vieler Arbeitnehmerinnen und Arbeitnehmer notwendig werden. So könnten viele neu entstehende Arbeitsplätze mit qualifizierten Kräften besetzt werden.

010100100101011101010111010

Die Teilnehmerinnen und Teilnehmer der Workshops sahen sehr klar den Bedarf an einer gesellschaftlichen Debatte darüber, in welchem Ausmaß bestimmte Aufgaben eines Jobs von KI übernommen werden könnten oder sollten und wo dies insbesondere wünschenswert sei oder nicht. Als hilfreiches Werkzeug erachteten sie dabei **Technikfolgenabschätzungen** auch der KI-Forschungsprojekte, die potenziell Auswirkungen auf die Zukunft der Arbeit haben. Die Mehrheit der Forschenden möchte, dass derartige Technikfolgenabschätzungen bereits parallel mit der Entwicklung von KI-Arbeitssystemen erfolgen. Gleichzeitig möchten sie aber die Unabhängigkeit ihrer Forschung wahren.

Ihr Appell: Wer forscht, muss frühzeitig darüber nachdenken, welche Veränderungen seine Entwicklung hervorrufen dürfte. Als Beispiel wurde ein Programmierer oder eine Programmiererin genannt, der oder die einen Algorithmus entwickelt, der das Scannen an der Kasse überflüssig macht. Er oder sie müsse sich der Tragweite bewusst sein, dass die Entwicklung Kassiererinnen und Kassierer ersetzen kann.

Die Forschenden äußerten vielfach den Wunsch durch andere Disziplinen wie Gesellschafts-, Sozial- und Politikwissenschaften unterstützt zu werden. Diese seien gefragt, die gesellschaftlichen Auswirkungen zu bewerten und einzuordnen.

110101101110101010111010010

Jeder technologische Fortschritt führe zu einer gewissen **Abhängigkeit**, urteilten die Forschenden.

Risiken und Nutzen seien auch hier gegeneinander abzuwägen. Dahinter steckt die Frage, wie weit wir uns mehr oder weniger komplett auf die KI verlassen. So sehen es viele Forschende kritisch, einer KI die Entscheidung über eine Kreditvergabe zu überlassen, insbesondere dann, wenn der angewendete Algorithmus nicht wirklich nachvollziehbar erscheint.

Die Wissenschaftlerinnen und Wissenschaftler der RIKI-Workshops rechnen damit, dass der Mensch durch den Einsatz von KI einige bisherige **Fähigkeiten verlieren** wird. So könne man verlernen, einen Straßenatlas zu lesen, wenn man ausschließlich mit einem Navigationssystem fahre. In den meisten Fällen sei dies jedoch kein Problem, zumal sich die Menschen zugleich ja auch neue Kenntnisse und Fähigkeiten aneigneten.

3. KOMMUNIKATIONSTIPPS VON FORSCHENDEN FÜR FORSCHENDE

- Wer Sorgen äußert, erwartet Verständnis. Zeigen Sie dieses Verständnis. Begegnen Sie Menschen, die ihren Arbeitsplatz durch die KI bedroht sehen, nicht nur sachlich, sondern auch empathisch – also mit Einfühlungsvermögen. Klar ist: Künstliche Intelligenz wird den Arbeitsmarkt verändern. Es ist jedoch nicht eindeutig absehbar, wie viele Arbeitsplätze und welche Berufssparten von der Einführung von KI betroffen sein werden.
- Heben Sie im Gespräch hervor, dass es vielfach nicht möglich sein wird, einen Arbeitsprozess vollständig zu automatisieren – etwa die Versorgung einer Krankenstation mit Essen. Gerade dort geht es aber auch gar nicht um einen Ersatz, sondern um eine Unterstützung im Arbeitsalltag – denn die Essensausgabe ist nicht nur eine Versorgung mit Nahrungsmitteln, sondern auch willkommener Anlass für eine soziale Interaktion.

- Sprechen Sie Chancen an, die durch den Einsatz von KI entstehen: Eine Vielzahl von Arbeitnehmerinnen und Arbeitnehmern könnte durch KI-Anwendungen entlastet und unterstützt werden, bis hin zu einer Reduzierung der Arbeitszeit oder einer intensiveren Konzentration auf die eigentliche Aufgabe. Ein Beispiel, das Sie dafür anführen können, sind die Dokumentationstätigkeiten von Pflegekräften – würden diese KI-gestützt weniger Zeit kosten, hätten die Pflegekräfte mehr Zeit für ihre eigentliche Aufgabe.
- Achten Sie im Gespräch mit besorgten Arbeitnehmerinnen und Arbeitnehmern darauf, bestimmte Tätigkeiten nicht pauschal herabzuwürdigen. Das Kassenpersonal kann seiner Arbeit mit großer Freude nachgehen und ist um so mehr darüber erschüttert, wenn der Job von einer KI übernommen wird.
- Je erfolgreicher die KI, desto größer die Wahrscheinlichkeit, dass manche Beschäftigte tatsächlich ihren Arbeitsplatz verlieren. Die Sorge, den Arbeitsplatz zu verlieren, geht Hand in Hand mit existenziellen Ängsten – das Einkommen könnte entfallen. Skizzieren Sie mögliche Lösungsansätze wie das Grundeinkommen oder eine KI-Steuer. Machen Sie jedoch auch klar, dass hier politische Entscheidungen gefordert sind.
- Machen Sie sich bereits zu Beginn Ihres Projekts Gedanken über mögliche gesellschaftliche Auswirkungen Ihrer Forschung. Wird Ihre Entwicklung Arbeitskräfte ersetzen? Setzen Sie sich gegebenenfalls mit Forschenden anderer Disziplinen in Verbindung, um eine Technikfolgenabschätzung zu erarbeiten. Sprechen Sie offen darüber, welche Auswirkungen Ihre Forschung auf die Gesellschaft haben könnte.
- Heben Sie hervor, wo ein Dialog nötig ist: Welche Anwendungen sollen an eine KI übergeben werden und welche nicht. Ein

Beispiel: Soll ein Algorithmus darüber entscheiden, ob und welche Weiterbildungen den Arbeitssuchenden finanziert werden?

LITERATUR

1. Fluhr, David: Jobverlust durch KI. Autonomes Fahren & Co (26/01/2019) <https://www.autonomes-fahren.de/jobverlust-durch-ki/>.
2. Frey, Carl Benedikt: In der Technologiefalle. Süddeutsche Zeitung (21/08/2019) <https://www.sueddeutsche.de/kultur/serie-kuenstliche-intelligenz-in-der-technologiefalle-1.4570763?reduced=true>.
3. Apt, Wenke & Priesack, Kai: Künstliche Intelligenz, KI und Arbeit – Chance und Risiko zugleich. Hrsg.: Wittpahl, Volker (Springer Berlin Heidelberg, 2019). ISBN 978-3-662-58042-4 (e-Book)
4. Genesys Telecommunications Laboratories GmbH: Künstliche Intelligenz am Arbeitsplatz: Kollege oder Konkurrent? Marketing Börse (03/09/2019) <https://www.marketing-boerse.de/news/details/1936-kuenstliche-intelligenz-am-arbeitsplatz-kollege-oder-konkurrent/159885>.
5. Joffe, Josef: Robby, fass! Die Zeit Nr.25/2019 (13/06/2019) <https://www.zeit.de/2019/25/automatisierung-arbeitsplatz-roboter-jobverlust-ausbildung-arbeitskraefte>.
6. Hagelüken, Alexander: Maschinen schaffen mehr Jobs als sie vernichten. Süddeutsche Zeitung (09/01/2019) <https://www.sueddeutsche.de/wirtschaft/automatisierung-jobs-roboter-1.4278998>.
7. Europäische Kommission: Special Eurobarometer 460 'Attitudes towards the impact of digitisation and automation on daily life'. S.152-158 (03/2017) ISBN: 978-92-79-68471-5, <https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/yearFrom/1974/yearTo/2017/surveyKy/2160>.
8. t3n Redaktion: Wir nannten es Arbeit! Was Künstliche Intelligenz auch für deinen Arbeitsplatz bedeutet. t3n Nr.51 (15/03/2019) <https://t3n.de/magazin/nannten-arbeit-kuenstliche-intelligenz-arbeitswelt-244325/>.
9. Frey, Carl Benedikt & Osborne, Michael A.: The future of employment: How susceptible are jobs to computerisation? (17/09/2013) https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf.
10. Suich Bass, Alexandra: GrAlt expectations. Non-tech businesses are beginning to use artificial intelligence at scale. The Economist Nr.9085 Vol. 426, S. 3-12 (28/03/2018). <https://www.economist.com/special-report/2018/03/28/non-tech-businesses-are-beginning-to-use-artificial-intelligence-at-scale>
11. Gretzinger, Nick: Künstliche Intelligenz: Sind unsere Arbeitsplätze nun bedroht oder nicht? Ingenieur.de <https://www.ingenieur.de/karriere/arbeitsleben/alltag/kuenstliche-intelligenz-sind-unsere-arbeitsplaetze-nun-bedroht-oder-nicht/>.
12. Harwell, Drew: A face-scanning algorithm increasingly decides whether you deserve the job. The Washington Post (06/11/2019) <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>.
13. Algorithmus brachte sich Sexismus bei. Deutschlandfunk Nova <https://www.deutschlandfunknova.de/nachrichten/kuenstliche-intelligenz-amazon-verabschiedet-sich-von-sexistischem-algorithmus>.
14. Videostatements des RIKI-Projekts. <https://www.youtube.com/playlist?list=PLYFa5hehwobjh1EzdHD3Tjy5B98UCZWJ>.

→ Alle hier genannten Webseiten geprüft am 24.01.2020

AUTONOMES FAHREN

1. MEDIALE BERICHTERSTATTUNG

Autonomes Fahren: Fahren ohne Fahrer oder FahrerIn – das haben viele vor Augen, wenn es um die Mobilität von morgen geht, und nicht selten stammen dabei besonders visionäre Bilder aus der Science-Fiction-Welt. Nun hat sich auch im realen Leben in den vergangenen Jahrzehnten einiges getan: Der Tempomat, die Abstandsmessung oder die Einparkhilfe sind erste Schritte, mit denen Teilbereiche automatisiert wurden. Wirklich autonom fahrende Fahrzeuge – also ohne jeglichen Eingriff der Insassen – werden den Markt allerdings auf absehbare Zeit nicht erobern. Die mediale Berichterstattung bemüht sich um realistische Einschätzungen: „Das vollkommen autonome Fahren wird vorerst nicht kommen“¹, „Roboterautos brauchen noch viele Jahre“². Begründet wird dies mit „Hürden“³ auf dem Weg hin zum autonomen Fahren. Es werden jedoch immer mehr Teststrecken frei gegeben, und Fach- wie Publikumsmedien berichten über wichtige Erfolge auf dem Weg zum autonomen Mobil.⁴⁻⁶

01010101001010111010101111

Ein Grund für die anhaltend großen Erwartungen mag sein, dass die Begriffe „**autonom**“ und „**automatisiert**“ oft synonym verwendet werden. So schreibt zum Beispiel der Spiegel: „Volkswagen hat den Testbetrieb für autonomes Fahren in Hamburg gestartet. Ab sofort fahren die automatisierten Fahrzeuge des Konzerns völlig eigenständig mitten durch den Stadtverkehr.“⁴ Selten werden bei der Begriffswahl die allgemeingültigen und vielfach beschriebenen Definitionen zum autonomen Fahren der International Society of Automotive Engineers berücksichtigt.^{3,7-9} Dabei differenziert

dieses Modell anhand eines Sechs-Stufen-Systems (Stufe 0 bis Stufe 5) sehr deutlich die beiden Begriffe und ist darüber hinaus hilfreich für die weitere Einschätzung (siehe Grafik S.32).

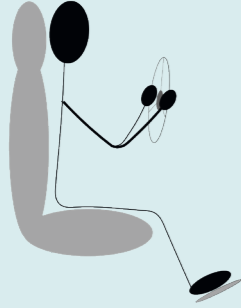
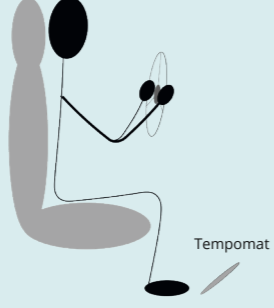
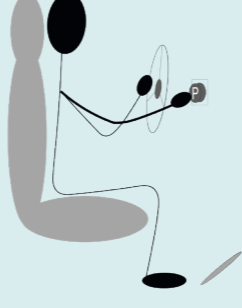
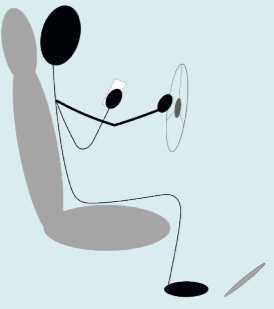
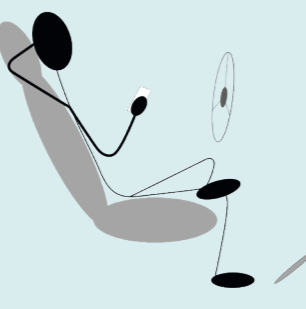

Umfragen in Deutschland zeigen, dass die Befragten auch ohne Kenntnis der beiden Begriffe in ihrer Bewertung genau am Übergang vom automatisierten zum autonomen Fahren kritischer werden (acatech & Körber Stiftung 2019¹⁰, Ernst & Young 2019¹¹). So lehnt es fast die Hälfte ab, die Kontrolle komplett an ein autonomes Fahrzeug abzugeben (Stufe 5).^{10,11} Fahrten mit vollautomatisierten Fahrzeugen der Stufe 4, bei denen noch die Möglichkeit besteht, in Gefahrensituationen einzugreifen, lehnt dagegen nicht einmal jede dritte befragte Person ab.¹¹ Auf einzelne konkrete Situationen hin befragt, können sich 54 Prozent vorstellen, die Kontrolle im Stau abzugeben. 65 Prozent würden dem Fahrzeug auch beim Ein- und Ausparken die Kontrolle überlassen – Bitkom Research 2018.¹²

Immerhin erwarten von den Befragten der Bitkom-Erhebung rund 60 Prozent mehr Sicherheit für Verkehrsteilnehmerinnen und -teilnehmer durch selbstfahrende Autos. Umgekehrt befürchten 68 Prozent technische Probleme, 63 Prozent sorgen sich vor potenziellen Hackerangriffen und 52 Prozent sehen die Gefahr eines Datenmissbrauchs.¹²

„IHR AUTO WURDE GEHACKT! ZAHLEN SIE!! ODER IHRE BREMSEN VERSAGEN!!!“

Titelzeile Artikel von Anja Reiter, Die Zeit 12/2019 ¹⁴

Automatisierungsstufen

STUFE 0 keine Unterstützung	STUFE 1 assistiert	STUFE 2 teilautomatisiert
<p>Alle Verantwortung liegt bei der Fahrerin oder dem Fahrer. Es gibt keine eingreifenden, sondern lediglich warnende Symsteme.</p> 	<p>Einzelne Assistenzsysteme unterstützen den Fahrer oder die Fahrerin bei bestimmten Fahraufgaben. Bsp.: ESP, Tempomat, Spurhalteassistent</p> 	<p>Ein oder mehrere Assistenzsysteme übernehmen Lenk- und Beschleunigungsvorgänge. Die Fahrerin oder der Fahrer muss jederzeit die Steuerung des Fahrzeugs übernehmen können. Bsp.: automatische Einparkhilfe, Stauassistent, Überholassistent</p> 
STUFE 3 hochautomatisiert	STUFE 4 vollautomatisiert	STUFE 5 autonom
<p>Das System übernimmt die Ausführung aller Aspekte der dynamischen Fahraufgaben. Der Fahrer oder die Fahrerin darf sich in speziellen Anwendungsfällen für eine begrenzte Zeit vom Fahrgeschehen abwenden, muss jedoch jederzeit wieder eingreifen können. Bsp.: Autobahnfahrt (ohne Baustelle)</p> 	<p>In bestimmten Anwendungsfällen kann das Fahrzeug alle Fahraufgaben selbständig übernehmen. Der jeweilige Anwendungsfall beinhaltet den Straßentyp, den Geschwindigkeitsbereich und die Umfeldbedingungen. Bsp.: Autobahnfahrt inkl. Baustellen und Auf- und Abfahrt</p> 	<p>Das Fahrzeug kann vollumfänglich auf allen Straßentypen, in allen Geschwindigkeitsbereichen und unter allen Umfeldbedingungen die Fahraufgabe vollständig allein durchführen.</p> 

32

Mit zunehmendem Grad der Automatisierung steigt die Menge der **Daten**, die erfasst und mit Netzwerken ausgetauscht werden. An diesen Schnittstellen können Hacker potenziell in das System eindringen, Daten entwenden, diese missbrauchen oder manipulieren. Die Schnittstellen zwischen Auto und anderen Netzwerken

bedürfen daher besonderer Schutzmechanismen. Diese Sorge vor Hackerangriffen wird in den Medien vielfach aufgegriffen. Dabei gibt es ein breites Spektrum von sachlichen bis hin zu effektheischenden Darstellungen.¹³⁻¹⁵

Darüber hinaus beschäftigen sich die Medien mit dem Thema Datenhoheit. „Wem gehören die Daten meines Autos?“¹⁶ Wem soll also der Zugriff auf gespeicherte Daten innerhalb des Fahrzeugs oder im Datentransfer erlaubt werden – Behörden, Marketing-Unternehmen, Autowerkstätten, Versicherungsunternehmen, den Halterinnen und Haltern oder Fahrenden des automatisierten oder autonomen Fahrzeugs?

In diesem Zusammenhang beschreiben viele Medien kritisch, dass für eine Umsetzung des autonomen Fahrens noch rechtliche Rahmenbedingungen im Bereich Datenschutz und Haftbarkeit fehlen.¹⁶⁻¹⁹

010100100101011101010111010

Die ethischen Debatten um autonomes Fahren fokussieren mehrheitlich auf theoretisch mögliche **Dilemma-Situationen**.^{20,21} Soll das autonome Auto die Insassen oder andere Verkehrsteilnehmer retten oder dafür jemand anderen überfahren? Wer wird geopfert? Kind oder Rentner? Fußgängerin oder Radfahrer? Solche Entscheidungsszenarien sind vielfach dargestellt auf der Plattform „Moral Machine“ des Massachusetts Institute of Technology.^{22,23} Wie in einem Online-Spiel können sich Menschen weltweit zu diesen Szenarien äußern, wie sie entscheiden würden, bzw. wie ein selbstfahrendes Auto entscheiden sollte.

Bis Ende 2018 haben auf der Plattform Daten zu den Entscheidungen von über 40 Millionen Personen aus über 200 Ländern vorgelegen. Die Forscher konnten anhand dieser Stimmabgaben Antworten nach Personengruppen und Ländern vergleichen.²⁴ In bestimmten Dilemma-Situationen kristallisieren sich universale Entscheidungsmuster heraus. So erachteten die meisten Teilnehmer und Teilnehmerinnen das Überleben eines Menschen im Falle eines Unfalls als wichtiger als das Überleben eines Tieres. Ebenso lässt sich be-

obachten, dass manche Personengruppen besonders häufig geschont werden sollen: Babys, Kinder und schwangere Frauen. Ordnet man die Ergebnisse bestimmten Kulturkreisen zu, sind jedoch deutliche Unterschiede festzustellen. Die Festlegung europa- oder gar weltweiter Regelungen dürfte sich demnach nicht nur durch individuelle, sondern auch durch kulturelle Unterschiede als schwierig bis unmöglich erweisen.²⁴ Es gibt außerdem grundlegende Zweifel an der Sinnhaftigkeit derartiger Entscheidungsabfragen. Die Verkehrsszenarien – so ein Vorwurf – wirkten beispielsweise zu konstruiert.²⁵

Die deutsche Ethik-Kommission „Automatisiertes und vernetztes Fahren“ bleibt bei diesem Thema in ihrem 2017 vorgelegten Bericht eher vage²⁶ und fordert, dass die Technik darauf ausgelegt sein müsse, derartige kritische Situationen per se zu vermeiden.

33

2. AUS SICHT DER FORSCHENDEN

Im Rahmen der RIKI-Workshops wurde autonomes und automatisiertes Fahren vor allem an den Standorten in Karlsruhe und Kaiserslautern thematisiert. Zudem fand im Rahmen des RIKI-Projekts im Juli 2019 eine öffentliche Diskussion zum Thema autonomes Fahren und Künstliche Intelligenz mit Expertinnen und Experten des Karlsruher Instituts für Technologie und des FZI Forschungszentrum Informatik in Karlsruhe statt. Bei dieser Diskussion im Fishbowl-Format konnten Bürgerinnen und Bürger ihre Fragen und Meinungen einbringen.

Sowohl in den Workshops als auch in der Diskussion betonten die Wissenschaftlerinnen und Wissenschaftler, dass die **Unterscheidung zwischen den Begriffen autonom und automatisiert** sehr wichtig sei. Sie wünschten sich mehrheitlich, dass in der medialen Berichterstattung mehr darauf geachtet werden sollte, diese beiden Begriffe zu

Wie gut funktioniert automatisiertes Fahren heute?

Dr.-Ing. Oliver Wasenmüller
DFKI Kaiserslautern

Autonomes Fahren – die Schuldfrage

Christian Hubschneider
FZI Forschungszentrum Informatik, Karlsruhe

Reden Sie mit:
Autonomes Fahren und Künstliche Intelligenz
Diskussion mit Bürgern in Karlsruhe

Videostatements von Forschenden und ein Mitschnitt der öffentlichen Diskussion in Karlsruhe im Juli 2019 sind online abrufbar.²⁷

34 trennen (siehe Abbildung S.32). Gleichzeitig sehen die Forschenden sich selbst mit in der Verantwortung, wiederholt auf die Unterschiede zwischen autonom und automatisiert hinzuweisen und bei der eigenen Kommunikation auf die Verwendung des jeweils zutreffenden Begriffs zu achten.

Dadurch wäre auch eine der von Bürgerinnen und Bürgern wie auch Medien am häufigsten gestellten Fragen besser zu beantworten: Wann wird das autonome Fahren kommen? Seitens der Forscherinnen und Forscher ist das Unbehagen groß, auf diese Frage zu antworten. Zumal häufig versucht werde, sie zu **konkreten Jahresangaben** zu drängen. Das Wissen um die verschiedenen Stufen der Automatisierung könnte aber aus Sicht der Forschenden entscheidend dazu beitragen, dass

Bürgerinnen und Bürger die Zeithorizonte selbst besser einschätzen können: Welche Stufen der Automatisierung bis hin zum autonomen Fahren könnten in welchen ungefähren Zeithorizonten realistisch auf die Straße kommen?

010100111110101000110101010

Einzelne der Forschenden im RIKI-Projekt gaben in diesem Zusammenhang sogar an, dass sie glauben, das **autonome Fahren (Stufe 5)** könne **niemals flächendeckend** realisiert werden.

Erstens sei der Straßenverkehr in gewissen Situationen dafür zu **komplex**. Als Beispiel wurde hier etwa die **Verkehrssituation** rund um den Arc de Triomphe in Paris genannt, wo zwölf Zufahrtsstraßen in eine nicht definierte Anzahl an Spuren münden.

Zweitens sei das Zusammenspiel aller Verkehrsteilnehmenden zu berücksichtigen – neben dem Verkehrsverhalten anderer Autos müsste dies nämlich auch für Rad- und Fußverkehr sowie weitere Verkehrsteilnehmende präzise eingeschätzt werden. In diesem **multikomplexen Interaktionssystem** mit anderen Verkehrsteilnehmern und -teilnehmerinnen sehen viele Forschende ein großes Problem bei der Umsetzung des autonomen oder automatisierten Fahrens der Stufen 4 und 5. Realistischer erscheint es vielen, dass auch langfristig nur bestimmte Verkehrssituationen von automatisierten Fahrzeugen übernommen werden können, aber eben nicht pauschal alle.

Drittens gebe es im Alltag **laufend unvorhersehbare Veränderungen im Straßenverkehr**. Selbstfahrende Autos müssten etwa auch belebte Verkehrsschilder richtig interpretieren, bei widrigen Witterungsbedingungen die Umgebung richtig erfassen oder unübersichtliche Baustellen meistern. Werden bei Baustellen beispielsweise

zusätzliche, vorübergehend geltende Fahrbahnmarkierungen auf die Straße aufgebracht, müssen diese von einem vollautomatisierten oder autonomen Fahrzeug richtig interpretiert werden. Es kann also sein, dass das Fahrzeug eine durchgezogene Linie überfahren muss, um die Ausweichstrecke auf dem Standstreifen zu nutzen. Dies widerspricht außerhalb dieser Situation aber den allgemeinen Verkehrsregeln.

Ein weiteres Problem sehen Forschende bei einem automatisierten oder autonomen Fahrzeug darin, dass es sich an die **Verkehrs- und Abstandsregeln** hält, sobald das System die Kontrolle übernimmt. Insbesondere Abstandsregelungen nach vorne und hinten und auch zur Seite hin, etwa beim Überholen von Fahrradfahrern und Fahrradfahrerinnen, müssen eingehalten werden. Deshalb sei es für ein vollautomatisiertes oder autonomes Fahrzeug eine Herausforderung, sich in den laufenden Verkehr zu integrieren.

Die Gruppe der Optimisten unter den Forschenden im Hinblick auf autonome Fahrzeuge sieht für diese Punkte zwei Lösungswege. Zum einen müsse mit vielen Daten trainiert werden, um auch unterschiedlich komplexe Situationen korrekt einschätzen zu können. Zum anderen sei es wichtig, ab einem gewissen Punkt den gesamten Verkehr auf autonome Fahrzeuge umzustellen, die im Datenaustausch stehen.

000101001000001101010111010

In den Workshops wie auch in der Debatte betonten die Wissenschaftlerinnen und Wissenschaftler, dass ihnen die **Sicherheit der Datensysteme**, die zum Einsatz kommen, sehr wichtig sei. Dies sehen sie als eine maßgebliche Herausforderung bei der Entwicklung. Denn schließlich würden mit der wachsenden Anzahl an automatisierten und autonomen Fahrzeugen immer größere Datenmengen

erhoben und auch mit Netzwerken ausgetauscht. Dabei sehen viele Forschende insbesondere an den Schnittstellen im **Datentransfer** ein großes Sicherheitsrisiko. Die Schnittstellen bieten potenzielle Angriffsflächen für Hacks, Systemmanipulationen oder Datenmissbrauch. Dringen Hacker in ein System ein, könnten sie etwa ein Fahrzeug unter ihre Kontrolle bringen und in den Verkehr eingreifen. So könnten sie verschiedene Fahrzeuge ineinander fahren lassen und Unfälle provozieren.

Für die Automobilhersteller sei es daher eine zentrale Herausforderung, potenziellen Hackern immer einen Schritt voraus zu sein. Dies sei auch ein nicht zu unterschätzender Kostenfaktor.

0101001011101001101010111010

Auch das Thema **Datenhoheit** und die potenzielle Möglichkeit, Daten zur **Überwachung** zu nutzen, stuften viele der Forschenden als sehr bedeutsam ein.

Denn um das automatisierte Fahren sicher zu gestalten, werden Daten des Fahrzeugs und dessen Umgebung erfasst, festgehalten und gegebenenfalls mit anderen Fahrzeugen ausgetauscht. Wenn die Daten in die Hände Dritter gelangten, könnten sie auch für andere Zwecke gebraucht oder sogar missbraucht werden. Ein Beispiel: Wie hochauflösend sollen Sensoren sein, die Fuß- und Fahrradverkehr erkennen, um im Zweifel das Notbremsystem oder Ausweichmanöver zu starten? Einige Forschende weisen darauf hin, dass durch die hohe Auflösung der Sensoren, auch einzelne Personen identifiziert werden könnten. Mithilfe aller erhobenen Daten könnten Personen überwacht werden.

Auch Daten aus dem Fahrzeuginnenraum könnten entwendet und missbraucht werden. Dabei handelt es sich um Aufnahmen des Innenraums,

die beispielsweise künftig erfassen könnten, ob der Fahrer oder die Fahrerin gelangweilt oder aggressiv ist. Dazu werden Fahrer oder Fahrerin während der Fahrt gefilmt. Behörden könnten – wie teilweise bereits in China²⁸ oder Australien²⁹ im Einsatz – mithilfe dieser Daten überprüfen, ob der Fahrer oder die Fahrerin mit dem Handy oder anderweitig abgelenkt war und gegebenenfalls entsprechende Strafen verhängen.

Doch wann fängt Überwachung an? Ein Teil der Forschenden empfindet es bereits als Überwachung, wenn überhaupt Daten über die Fahrer oder Fahrerinnen erhoben werden. Ein anderer Teil befürwortet eine Erhebung von Daten der Fahrzeuglenker, wenn die Außensensoren des Automobils Gefahrensituationen erfasst haben. Vor allem aber müsse geklärt werden, wer solche Daten wann auslesen und nutzen dürfe.

Häufig werden die Forschenden auf **moralische Dilemmata** angesprochen (siehe S.33). Zu wessen Gunsten soll das System im Falle eines unvermeidbaren Unfalls entscheiden? Die KI-Forscher und -Forscherinnen können dazu keine Entscheidungen treffen. Zudem müsse ein autonomes Auto nach Ansicht der Forschenden so konzipiert sein, dass eine solche Situation nicht auftreten könne – dann bräuchte man beispielsweise auch nicht über den Vorschlag zu diskutieren, eine solche Entscheidung einem Zufallsgenerator zu überlassen. Eigentlich müssten autonome Fahrzeuge dann so defensiv fahren, dass Unfälle nur durch das Verschulden anderer hervorgerufen werden könnten.

Letztlich sehen die Forschenden hier den einzig richtigen Weg in einer gesellschaftlichen Debatte mit dem Ziel allgemeingültiger Vorgaben.

Unabhängig von diesen ethischen Entscheidungen wiesen die Forschenden häufig darauf hin, dass zudem **rechtliche Rahmenbedingungen** geklärt werden müssen. So sehen sie bei automatisierten Fahrzeugen Lücken in der Rechtslage. Aus ihrer Sicht ist strafrechtlich nicht eindeutig geklärt, wer die Verantwortung trägt, wenn das System zum Zeitpunkt des Unfalls die Kontrolle über das Fahrzeug hatte. Sowohl beim autonomen Fahren wie auch beim automatisierten Fahren ab Stufe 3 müsse geklärt werden, inwieweit Fahrzeughalter oder-halterin, das herstellende Unternehmen oder eine Zertifizierungsstelle in die Verantwortung genommen werden könnten. Ein interessanter Punkt insbesondere vor dem Hintergrund, dass es den „Fahrer“ oder die „Fahrerin“ ja in diesen Stufen im klassischen Sinne gar nicht mehr gibt. Auch die Straßenverkehrsordnung müsse weiter angepasst werden, damit autonome Fahrzeuge ohne Sondergenehmigung am Verkehr teilnehmen können.

3. KOMMUNIKATIONSTIPPS VON FORSCHENDEN FÜR FORSCHENDE

- Unterscheiden Sie in Ihrer Kommunikation bewusst zwischen „autonom“ und „automatisiert“. Das Sechs-Stufen-System kann dabei helfen, den Unterschied für Laien verständlicher zu machen^{30,31} (siehe S.32).
- Lassen Sie sich nicht zu Antworten drängen. Ein Beispiel: Wann wird das autonome Fahren kommen? Hier antworten Sie am besten mit einem groben zeitlichen Rahmen und begründen, warum eine konkrete Festlegung schwierig oder nicht möglich ist. Geben Sie zusätzlich eine realistische Einschätzung, wie weit die Forschung ist und was noch zu tun bleibt. Verweisen Sie dabei auf das Sechs-Stufen-System.
- Automatisiertes und autonomes Fahren soll die Sicherheit von Verkehrsteilnehmerinnen und -teilnehmern verbessern. Weisen Sie je-

doch auch auf mögliche Probleme im alltäglichen Straßenverkehr hin: Da sich das System an Verkehrs- und Abstandsregeln hält, kann es eine Herausforderung sein, sich in den laufenden Verkehr zu integrieren.

- Veranschaulichen Sie die Komplexität der Verkehrssysteme anhand von Beispielen, etwa dem Verkehrsgewirr in einer indischen Großstadt mit Autos, Mopeds, Rad- und Fußverkehr. Eine Veranschaulichung aus unserem Alltag hierzulande kann eine Baustelle mit verengten Fahrstreifen sein, die bei widrigem Wetter passiert werden muss.
- Nehmen Sie mögliche Bedenken Ihres Gegenübers ernst. Dazu zählt vor allem der Aspekt Sicherheit. Erläutern Sie, was Sie als Forschende tun können, um die Sicherheit von Systemen, Daten und Personen zu gewährleisten.
- Heben Sie im Gespräch hervor, dass an der Sicherheit der Systeme sowohl von Seiten der Forschung als auch der Industrie stetig gearbeitet wird. Die Schnittstellen im Datentransfer von automatisierten Fahrzeugen ab Stufe 3 bergen ein großes Sicherheitsrisiko, da sie Angriffsflächen für Datenmanipulationen bieten.
- Machen Sie auch deutlich, dass die erhobenen Daten an Dritte gelangen und zu anderen Zwecken verwendet oder auch missbraucht werden könnten. Aufnahmen des Fahrers oder der Fahrerin könnten an Behörden weitergeleitet werden. Die Wahrnehmung, ab wann eine Überwachung stattfindet, ist subjektiv. Vermeiden Sie deshalb Ansichten vorzugeben.
- Scheuen Sie sich nicht davor, darauf hinzuweisen, wo Ihre Expertise endet und andere Akteure in der Verantwortung stehen. So müssen beispielsweise Gesellschaft und Politik entscheiden, wie ein autonomes Fahrzeug in einer Dilemma-Situation entscheiden soll – oder ob dies gar einem Zufallsgenerator überlassen werden sollte.

- Machen Sie auf ungeklärte Fragen aufmerksam. Wobei Sie als Forscherin oder Forscher natürlich Teil der Gesellschaft sind. Wie viele und welche Daten werden erhoben? Wem gehören die Daten? Wer kann sie einsehen? Wie lange und wo werden sie gespeichert? Wer haftet bei einem Unfall mit einem automatisierten oder autonomen Fahrzeug? Diese Fragen erfordern Diskurse in Gesellschaft und Politik über rechtliche Rahmenbedingungen.

LITERATUR

1. Delhaes, Daniel & Murphy, Martin: Das vollkommen autonome Fahren wird vorerst nicht kommen. Handelsblatt Online (17/07/2019) <https://www.handelsblatt.com/politik/deutschland/hohe-kosten-das-vollkommen-autonome-fahren-wird-vorerst-nicht-kommen/24597246.html>.
2. Germis, Carsten: „Roboterautos brauchen noch viele Jahre“. Frankfurter Allgemeine Zeitung (06/09/2019) <https://www.faz.net/aktuell/wirtschaft/zukunft-der-mobilitaet-realitaet-von-roboter-autos-16368536.html>.
3. Tyborski, Roman: Das sind die fünf Hürden auf dem Weg zum autonomen Fahren. Handelsblatt Online (28/08/2019) <https://www.handelsblatt.com/technik/digitale-revolution/digitale-revolution-das-sind-die-fuenf-huerden-auf-dem-weg-zum-autonomen-fahren/24946006.html>.
4. Frahm, Christian: VW fährt in Hamburg jetzt autonom. Spiegel Online (03/04/2019) <https://www.spiegel.de/auto/aktuell/hamburg-volkswagen-testest-autonomes-fahren-in-deutscher-grossstadt-a-1251825.html>.
5. Briegleb, Volker: Teststrecke für automatisiertes Fahren mitten in Berlin. Heise Online (19/09/2019) <https://www.heise.de/newsticker/meldung/Teststrecke-fuer-automatisiertes-Fahren-mitten-in-Berlin-4534423.html>.
6. Nefzger, Emil: Karlsruhe wird Testgebiet für Roboter-Autos. Spiegel Online (03/05/2018) <https://www.spiegel.de/auto/aktuell/karlsruhe-testfeld-fuer-autonomes-fahren-eroeffnet-a-1206008.html>.
7. Autonomes Fahren. Wikipedia [https://de.wikipedia.org/wiki/Autonomes_Fahren#Autonomiestufen_\(Level\)](https://de.wikipedia.org/wiki/Autonomes_Fahren#Autonomiestufen_(Level)).
8. Grundhoff, Stefan: Bei schlechtem Wetter unzuverlässig: Autonomes Fahren weiter weg als gedacht. Focus Online (01/04/2019) https://www.focus.de/auto/news/technikbericht-autonomes-fahren-autonomes-fahren-noch-weiter-weg-als-gedacht_id_10520057.html.
9. Hommen, Mario: Roboterautos ungewiss. Zeit Online (22/08/2018) <https://www.zeit.de/mobilitaet/2018-08/autonomes-fahren-robotik-kuenstliche-intelligenz-auto-mobilitaet>.
10. acatech & Körber-Stiftung: TechnikRadar 2019. (28/05/2019) <https://www.acatech.de/wp-content/uploads/2019/05/Langfassung-Technik-radar-2019-final.pdf>.
11. Ernst & Young GmbH: Elektromobilität und autonomes Fahren in Deutschland. (09/2019). https://assets.ey.com/content/dam/ey-sites/ey-com/de_de/news/2019/09/ey-automobilbarometer-2019.pdf?download
12. Bühler, Dr. Joachim & Rohleder, Dr. Bernhard: Autonomes Fahren und vernetzte Mobilität. (bitkom und VdTÜV, 18/04/2018). <https://www.bitkom.org/sites/default/files/file/import/Bitkom-Charts-Autonomes-Fahren-und-vernetzte-Mobilitat-18-04-2018-final.pdf>.

13. Maier-Borst, Haluka: Hacken ins Auto der Zukunft. Deutschlandfunk (08/07/2015) https://www.deutschlandfunk.de/autonomes-fahren-hacken-ins-auto-der-zukunft.676.de.html?dram:article_id=324879.

14. Reiter, Anja: 'Ihr Auto wurde gehackt! Zahlen Sie!! Oder Ihre Bremsen versagen!!!' Die neue Bedrohung? Die Zeit Nr.12/2019 (14/03/2019) <https://www.zeit.de/2019/12/selbstfahrende-autos-sicherheit-strassenverkehr-gefahren>.

15. Lewicki, Martin: Hacker machen Autos zu ferngesteuerten Geschossen. Die Welt (14/09/2014) <https://www.welt.de/motor/article132197392/Hacker-machen-Autos-zu-ferngesteuerten-Geschossen.html>.

16. Doll, Nikolaus: Wem gehören in Zukunft die Daten meines Autos? Die Welt (22/01/2019) <https://www.welt.de/wirtschaft/article187462368/Autonomes-Fahren-Wer-hat-die-Hoheit-ueber-die-Daten.html>.

17. Krempf, Stefan: EU-Parlament: Haftungs- und Ethikprobleme beim autonomen Fahren rasch lösen. Heise Online (15/01/2019) <https://www.heise.de/newsticker/meldung/EU-Parlament-Haftungs-und-Ethikprobleme-beim-autonomen-Fahren-rasch-loesen-4276700.html>.

18. Reis, Dennis: Autonomes Fahren und die rechtlichen Fragen. Automobil Industrie (06/06/2019) <https://www.automobil-industrie.vogel.de/autonomes-fahren-und-die-rechtlichen-fragen-a-836396/>.

19. Doll, Nikolaus: Selbst der unschuldige Fahrer soll für das autonome Auto haften. Die Welt (21/11/2018) <https://www.welt.de/wirtschaft/article184207482/Autonomes-Fahren-Der-Fahrer-soll-immer-haften-selbst-wenn-er-unschuldig-ist.html>.

20. Hucko, Margret: Todlangweilig? Sicher! Spiegel Online (28/04/2016) <https://www.spiegel.de/auto/aktuell/autonomes-fahren-horrorszenarien-sind-unwahrscheinlich-a-1084850.html>.

21. dpa: Autonome Autos: Wen überfahren? Heise Online (25/10/2018) <https://www.heise.de/newsticker/meldung/Autonome-Autos-Wen-ueberfahren-4203485.html>.

22. Köppe, Julia: Einer muss sterben - nur wer? Spiegel Online (24/10/2018) <https://www.spiegel.de/wissenschaft/technik/unfaelle-mit-selbstfahrenden-autos-wer-soll-leben-wer-soll-sterben-a-1234901.html>.

23. Rahwan, Iyad; Bonnefon, Jean-Francois et al.: Moral Machine. <http://moralmachine.mit.edu/>.

24. Awad, Edmond; Dsouza, Sohan et al.: The Moral Machine experiment. Nature Nr.7729 Band 563, S. 59–64 (11/2018).DOI: 10.1038/s41586-018-0637-6, <https://doi.org/10.1038/s41586-018-0637-6>

25. Köhler, Konstantin: Wenn die Maschine entscheidet, wer sterben muss. Deutschlandfunk Nova (25/10/2018) <https://www.deutschlandfunknova.de/beitrag/ethik-fuer-roboter-autonomes-fahren-wer-soll-sterben>.

26. Ethik-Kommission Automatisiertes und vernetztes Fahren: Bericht Juni 2017. (Bundesministerium für Verkehr und digitale Infrastruktur, 06/2017). <https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf>.

27. Videostatements und Diskussionsmitschnitt des RIKI-Projekts. <https://www.youtube.com/playlist?list=PLYFa5shehwobLklrueyaLB-QCygxSw0-0Bs>.

28. Dorloff, Axel: Künstliche Intelligenz als Staatsziel. Deutschlandfunk (11/02/2019) https://www.deutschlandfunk.de/china-kuenstliche-intelligenz-als-staatsziel.724.de.html?dram:article_id=440743.

29. Frahm, Christian: Australien setzt künstliche Intelligenz gegen Handynutzung am Steuer ein. Spiegel Online (26.09.2019) <https://www.spiegel.de/auto/aktuell/australien-ueberwacht-handynutzung-am-steuer-mit-neuer-technologie-a-1288705.html>

30. Autonomes Fahren: Die 5 Stufen zum selbstfahrenden Auto. ADAC (07/11/2018) <https://www.adac.de/rund-ums-fahrzeug/autonomes-fahren/grundlagen/autonomes-fahren-5-stufen/>.

31. Verband der Automobilindustrie: Automatisierung von Fahrerassistenzsystemen zum automatisierten Fahren. (03/09/2015) <https://www.vda.de/de/services/Publikationen/automatisierung.html>.

→ Alle hier genannten Webseiten geprüft am 24.01.2020

KI IN DER MEDIZIN

1. MEDIALE BERICHTERSTATTUNG

„Doktor KI“, „Doktor Data“ oder „Doktor Algorithmus“ – die Bezeichnungen in der medialen Berichterstattung sind vielfältig. Doch die Einschätzung ist einheitlich: Künstliche Intelligenz wird die Medizin in vielen Gebieten verändern.^{1–4} Dabei geht es primär nicht darum, den Arzt oder die Ärztin zu ersetzen, sondern diese zu entlasten und zu unterstützen.^{1,2,4–6}

Auch wenn der Einzug der KI in den medizinischen Alltag teilweise als Wettstreit zwischen KI und Arzt oder Ärztin inszeniert wird, bewerten viele Medien eine Ergänzung der medizinischen Diagnostik, Therapieempfehlungen oder Überwachung des Klinikalltags mittels KI als positiv.^{4,7}

Die **Erwartungen** an KI in der Medizin reichen von einer präziseren Diagnostik bis zu individuelleren Präventions- und Therapieempfehlungen, verbunden mit mehr Zeit für den persönlichen Austausch zwischen Behandelnden und Behandelten.^{1,8}

- KI kann mit relativ hoher Wahrscheinlichkeit Hautkrebs erkennen¹, bei der Behandlung von Schlaganfällen helfen² oder Anzeichen von Depressionen in Stimmaufzeichnungen⁹, Videos¹⁰ oder Instagram-Bildern¹¹ erkennen.
- Manche KI-Expertinnen und -Experten hoffen, KI-Algorithmen könnten medizinische Leistungen kostengünstiger machen.⁴
- Die Entwicklung von Arzneimitteln wird laut Medienberichten von KI profitieren, ebenso wie das Feld der personalisierten Medizin mit für Patientinnen und Patienten maßgeschneiderten Therapieoptionen.^{12,13}
- Neben Ärztinnen und Ärzten könnten auch

Pflegekräfte durch KI-Anwendungen unterstützt werden.

- App-Systeme erlauben ein fortlaufendes Gesundheitstracking im Alltag – sofern erwünscht.^{14,15}

110100010100001111010101011

„DIAGNOSE AUS DER BLACKBOX“

Titelzeile Artikel von Astrid Viciano, Süddeutsche Zeitung¹⁶

01001010100100001010111010

Bedenken gegenüber KI-Einsätzen in der Medizin werden in der Medienberichterstattung vor allem im Zusammenhang mit Datensicherheit und Datenhoheit genannt, sowie in Bezug auf die Verlässlichkeit und die Aussagekraft der algorithmischen Systeme. Außerdem setze datengetriebene Forschung zu sehr auf Korrelationen und vernachlässige Kausalitäten.¹⁷

010100100101011101010111010

Schon heute sind intelligente Programme in der Medizin insbesondere in der **Analyse von Bild-daten** und entsprechender **Mustererkennung** leistungsstark. So hilft KI, seltene Erkrankungen zu erkennen oder radiologische Daten auszuwerten.¹⁶ Bereits 2017 berichtete ein Team der Stanford University in der Fachzeitschrift Nature beispielsweise, dass ihre KI Hautkrebs vergleichbar gut zu Fachärztinnen und Fachärzten diagnostizieren könne.¹⁸ Eine Studie von Forschenden rund um Holger Hänßle von der Universität Heidelberg legte im Mai 2018 nach. Demnach konnte eine mit mehr als 100.000 Aufnahmen trainierte KI schwarzen Hautkrebs besser erkennen als dies 58 Hautärztinnen und Hausärzten im Schnitt gelang.¹⁹

Um die Aussagekraft solcher Mustererkennungen zu stärken, muss die jeweilige KI mit möglichst **verlässlichen Daten** trainiert werden. Nicht nur die Menge der Daten ist also entscheidend, die Daten sollten vor allem eine **hohe Qualität** aufweisen.¹¹ Wichtig sind „standardisierte, saubere Daten mit gesicherten medizinischen Befunden“.²⁰

Mediziner und Medizinerinnen, die KI einsetzen, brauchen zudem Einblick in die **Architektur des Systems**. Wie sind die Algorithmen definiert? Das Trainieren von Mustererkennungen ist nicht immer eindeutig nachvollziehbar. Algorithmen können beispielsweise auf Basis falsch vorgegebener Annahmen falsche Ergebnisse produzieren. Für eine Verlässlichkeit muss aber jederzeit nachvollziehbar sein, nach welchen Kriterien die jeweiligen Algorithmen Entscheidungen treffen.⁶ Schlussendlich müssen die Ärzte und Ärztinnen die Empfehlungen der KI immer kritisch hinterfragen.

40

010001010011101010111010

Der Einsatz von Künstlicher Intelligenz in der **Vorsorge ist ein wachsender Markt**. Gesundheits-Apps, Fitnessarmbänder oder Smartwatches sind im Alltag vieler Menschen angekommen. Die KI-Systeme der sogenannten tragbaren Gadgets und **Fitness-Tracker** könnten über die Erfassung des Fitness-Level hinaus auch Krankheitsmuster wie etwa chronisch hohen Blutdruck frühzeitig erkennen und damit Ärztinnen und Ärzten zuarbeiten.^{14,15,21}

Ebenso können Handy-Apps zusätzlich diverse Daten zur Früherkennung sammeln. Für den Beispielsfall Parkinson könnten dies Informationen sein, wie etwa Sprachmerkmale beim Telefonieren, zurückgelegte Distanzen im Verlauf eines Tages oder emotionale Merkmale, die in Textnachrichten enthalten sind.²¹

11101011001001110101011101

Besonders in der Medizin ist der **Schutz personenbezogener Daten** wichtig. Hier stellt sich zudem die Frage der Datenhoheit. Häufig werden zur Verschlüsselung von Daten, um sie etwa zu Forschungszwecken nutzen zu können, zwei Möglichkeiten betrachtet: **Pseudonymisierung** und **Anonymisierung**.

- Bei pseudonymisierten Daten werden der Name und andere zentrale Merkmale ersetzt, Bezüge zu anderen Datensätzen bleiben jedoch erhalten. Pseudonymisierte Daten können mithilfe eines Schlüssels wieder zugänglich gemacht werden. Hier wird weiterhin von personenbezogenen Daten gesprochen.
- Bei anonymisierten Daten werden die personenbezogenen Daten sehr stark verändert, sodass sie nur sehr schwer wieder einem Individuum zuzuordnen wären. Es gibt keinen Bezug mehr zu anderen Datensätzen. Allerdings kann bei einer aussagekräftigen Datensammlung auch eine Anonymisierung rückgängig gemacht werden.^{22,23}

0101001001010101110101011010

Bislang dürfen Gesundheitsdaten nur zweckgebunden vom Arzt oder von der Ärztin erhoben und verwendet werden.²⁴ So können Ärztinnen und Ärzte aufgrund des Datenschutzes häufig nicht die gesamte Krankenakte auch mit Vermerken anderer Ärzte einsehen. Ironisch wird in den Medien angemerkt, dass Google genauere Informationen habe als der jeweilige Hausarzt oder die Hausärztin.¹¹

Dem Mangel an relevanten Trainingsdaten soll das **Digitale-Versorgung-Gesetz** Abhilfe schaffen, das Anfang November 2019 vom Bundestag verabschiedet wurde: Damit auch die Forschung Ge-

sundheitsdaten besser nutzen kann, sollen künftig Daten von **elektronischen Patientenakten** in eine **zentrale Gesundheitsdatenbank** fließen. In ihr sollen pseudonymisiert die Informationen von 73 Millionen gesetzlich Versicherten erfasst sein – dazu zählen Personendaten wie Alter und Geschlecht, Informationen zum Versicherungsverhältnis, zum Leistungsbezug und zum Gesundheitsstatus. Eine Widerspruchsmöglichkeit für die Versicherten oder Löschfristen sind nicht vorgesehen.²⁵

Damit werden die Gesundheitsdaten von gesetzlich Versicherten über die eigentliche Behandlung hinaus Behörden, Forschungseinrichtungen oder Universitätskliniken zur Verfügung gestellt.²⁶ Dieses Gesetz stieß auf starke Bedenken.²⁷

Umfragen des Bundesverbands Digitale Wirtschaft sowie der Technologie- und Methodenplattform für die vernetzte medizinische Forschung im Jahr 2019 in Deutschland zeigen übrigens, dass vier von fünf Befragten dazu bereit wären, Gesundheitsdaten anonym und unentgeltlich digital für die medizinische Forschung zur Verfügung zu stellen.^{28,29}

Etwas allgemeiner wird dieses Thema häufig unter dem Stichwort „Elektronische Patientenakte“ aufgegriffen. Als problematische Aspekte gelten dabei unter anderem fehlende einheitliche Standards oder rechtliche Regelungen. Offen ist auch die Frage, wie und zwischen welchen Beteiligten eigentlich der Transfer medizinischer Daten erfolgt – letzteres wird gerne illustriert am Beispiel der Zugriffsrechte von Krankenkassen.³⁰

2. AUS SICHT DER FORSCHENDEN

Im Rahmen der RIKI-Workshops wurde der Einsatz und die Forschung an KI-Systemen in der Medizin vor allem am Standort Berlin thematisiert. Zudem fand im Rahmen des RIKI-Projekts im Oktober 2019 eine öffentliche Diskussion zum Thema Künstliche Intelligenz in der Medizin mit Expertinnen und Experten der Fraunhofer-Gesellschaft und des Deutschen Herzzentrums Berlin im Fraunhofer-Forum in Berlin statt. Dabei konnten Bürgerinnen und Bürger ihre Fragen und Meinungen einbringen.

41

Videostatements von Forschenden und ein Mitschnitt der öffentlichen Diskussion in Berlin im Oktober 2019 sind online abrufbar.³¹

Die beteiligten Forschenden sehen **Chancen** in medizinischen KI-Anwendungen sowohl in verbesserten Diagnosen als auch passgenaueren Therapien. In beiden Fällen dient KI ausschließlich als Werkzeug, das Ärztinnen und Ärzte unterstützt. KI-Anwendungen werden nach Meinung der Forschenden die Mediziner und Medizinerinnen **nicht ersetzen** – auch nicht in Fachbereichen wie beispielsweise der Radiologie. Eine Radiologin oder ein Radiologe erkenne schließlich nicht nur Muster, sondern bespreche auch Befunde und Therapiemöglichkeiten mit der Patientin oder dem Patienten und gehe individuell auf das Gegenüber ein. Dementsprechend werde den Fachärzten und Fachärztinnen in der Radiologie nur ein Teil ihrer Arbeit abgenommen.

Viele Forschende hoffen, dass Medizinerinnen und Mediziner künftig mehr Zeit für Arzt-Patienten-Gespräche haben werden, weil KI-Algorithmen zeitaufwendige Aufgaben übernehmen könnten. Das Entwicklungsziel der Forschenden ist, für unterschiedlichste Anwendungsbereiche möglichst gute Entscheidungshilfen für Medizinerinnen und Mediziner bereitzustellen. Unumstritten ist unter allen Beteiligten, dass sich der Berufsalltag von Ärztinnen und Ärzten verändern wird.

0100000101101011101010111010

Wo es die erwähnten Chancen gibt, dort lauern auch **Risiken**. Darüber sind sich die Forschenden sehr wohl bewusst. So geben sie zu bedenken, dass es nie auszuschließen sei, dass die beim Training der Algorithmen eingesetzten **Datensätze unvollständig** und teils **fehlerhaft** waren. Deshalb müssten die anwendenden Ärztinnen und Ärzte die Empfehlung eines Algorithmus letztlich immer kritisch hinterfragen und ein eigenes Urteil treffen. Um fehlerhafte Entscheidungen eines KI-Systems frühzeitig zu erkennen, ist es aus Sicht der Forschenden notwendig, dass KI-Systeme nach-

vollziehbar sind (siehe Kapitel Umgang mit Daten S.23). Es müsse ersichtlich und **nachvollziehbar** sein, welche Informationen für die Entscheidung herangezogen und als wie relevant diese eingestuft wurden. Der Arzt oder die Ärztin müsse sich darauf verlassen können, dass die Entscheidung der KI auf maximal validen Merkmalen beruhe. Gern diskutiert werden in diesem Zusammenhang Fälle, in denen ein Algorithmus seine Ergebnisse auf Basis von Informationen generiert hat, die etwa in der Bildbeschriftung zu finden waren.

11101010101011101010111011

Jenseits möglicher Fehler und Unvollständigkeiten in Datensätzen sehen die Forschenden deutliche Vorteile im Einsatz diagnostischer KI-Werkzeuge. Schließlich würden diese mit einer großen Anzahl an verfügbaren Datensätzen trainiert. Dabei sind die Bilddaten meist mit den Einschätzungen erfahrener Medizinerinnen und Mediziner kombiniert – deren jeweilige Bewertungen und Diagnosestellungen durchaus variieren. Aus der Fülle der berücksichtigten Informationen erreichen KI-Systeme damit einen deutlich größeren „Erfahrungsschatz“ als einzelne Ärztinnen und Ärzte. Die KI-gestützte Diagnostik kann damit eine individuelle Bewertung eines Arztes oder einer Ärztin entweder erhärten oder in Frage stellen.

Die Mehrheit der Forschenden betont, dass Künstliche Intelligenz lediglich **Handlungsempfehlungen** geben solle. Es sind am Ende die Medizinerinnen und Mediziner, die die KI-Diagnose verifizieren müssen und im Austausch mit den jeweiligen Patientinnen und Patienten die darauf aufbauende Therapie planen. Nicht zuletzt sind es die Ärztinnen und Ärzte, die bei einer fehlerhaften Diagnose haftbar gemacht werden können.

000101001010101000110111

Viel diskutiert wurde darüber, wie die **Sicherheit von Gesundheitsdaten** gewährleistet werden kann. Die Forschenden sehen zwei zentrale Aspekte:

Zum einen müssen Patientinnen und Patienten verstehen, was mit ihren Daten geschieht und zu welcher Art von Datennutzung sie jeweils einwilligen. Dabei sehen die Forschenden die Ärzte und Ärztinnen in der Pflicht, bestmöglich zu **informieren** und **aufzuklären**.

Der Schutz der Daten müsse zudem technisch sichergestellt werden. Dies könne erreicht werden, indem einzelne **Diagnosedaten lokal** auf dem jeweiligen Gerät gespeichert werden. Im Idealfall könne das Gerät, das die Werte erfasst hat, diese auch direkt verarbeiten und auswerten. Die Daten müssten dann nicht auf andere Geräte übertragen werden. Dies minimiere das Risiko, dass die Daten in falsche Hände gelangen.

Großen Handlungsbedarf sehen die Forschenden in der Regelung der Gesundheitsdaten, die über Smartphone-Apps oder Fitnessarmbänder geteilt und von privaten Anbietern erhoben werden. Hier sehen die Forschenden die Notwendigkeit internationaler Rahmenbedingungen, denn häufig würden die Daten auf ausländischen Servern gespeichert und ausgewertet.

Zudem forderte die Mehrheit der Forschenden, darüber nachzudenken, welche Daten wirklich sinnvollerweise erfasst und verarbeitet werden müssen. Eine ungerichtete „Sammelwut“ sei nicht zielführend. Entscheidend sei vielmehr der Bezug zu konkreten Fragestellungen. Hier verweisen die KI-Forschenden auf das Konzept der **Datendiät** (siehe Kapitel Umgang mit Daten S.23).

Viele Forschende geben zu bedenken, dass eine Anonymisierung der Daten aus ihrer Sicht nicht

zwangsläufig ausreicht, um die Sicherheit der Gesundheitsdaten zu gewährleisten. Bei seltenen Krankheitsbildern beispielsweise sei es wegen der deutlich geringeren Fallzahl nicht sonderlich schwer, Zusammenhänge herzustellen.

Je mehr Merkmale erfasst werden, desto leichter sei es, Rückschlüsse auf einzelne Personen zu ziehen, so die Quintessenz der Forschenden. Trotz dieses Risikos halten es die Forschenden für wichtig, auch den Nutzen im Blick zu behalten.

0101001001011010101011010

Die Forschenden drängen darauf, einheitliche Standards für die Erfassung von Gesundheitsdaten zu definieren, um die Potenziale von KI-Algorithmen für alle Kliniken und Praxen zugänglich zu machen. Durch fehlende Standards könnten viele Algorithmen häufig nur an der Institution eingesetzt werden, an der sie entwickelt wurden. Momentan sei nicht gewährleistet, dass ein Algorithmus, der an der Klinik A entwickelt wurde, auch an einer anderen Klinik B oder C erfolgreich eingesetzt werden könne.

1001001001011010101011010

Kosten und Nutzen von KI-Systemen im Gesundheitswesen sollten nach mehrheitlicher Meinung der Forschenden kritisch gegeneinander abgewogen werden. Sie betonen, dass die Technik nur eingesetzt werden soll, wenn dadurch das Ergebnis mutmaßlich verbessert werden kann. Als Beispiele dienen aktuelle Studien etwa zur verbesserten Hautkrebsdiagnostik oder zur Verringerung von Blutverlusten bei Operationen mittels des Einsatzes von KI.

Nach den Einschätzungen der Forschenden können sich einzelne Algorithmen durch die Bereitstellung immer neuer Daten laufend weiterentwickeln

und so die Aussagekraft der Diagnostik sukzessive verbessern. Für die **Zulassung** solcher, sich in der Praxis **stetig verbessernder KI-Algorithmen** gibt es aber bisher keine gesetzlichen Regelungen.

3. KOMMUNIKATIONSTIPPS VON FORSCHENDEN FÜR FORSCHENDE

- Wenn Ihr Gegenüber Befürchtungen hegt, dass KI-Systeme Ärztinnen und Ärzte ersetzen könnten, machen Sie klar, was KI-Systeme in der Medizin konkret leisten können. Verdeutlichen Sie, dass es sich dabei stets um einzelne Anwendungsbereiche handelt – in der Diagnostik, im OP-Saal etc.
- Vermitteln Sie nicht den Eindruck eines Wettbewerbs zwischen Arzt und KI. KI-Anwendungen im Medizinbetrieb stellen keine Konkurrenz dar, sondern sind ein zusätzliches, unterstützendes Werkzeug. Der Vorteil von diagnostischen KI-Anwendungen liegt in der Kombination der großen Anzahl an verfügbaren Daten mit Einschätzungen erfahrener Medizinerinnen und Mediziner. Damit verfügt KI über einen besonders großen „Erfahrungsschatz“. Die letztendliche Diagnose – und die Verantwortung dafür – verbleiben bei den jeweils behandelnden Ärzten oder Ärztinnen. Gesamtanamnesen, wie sie Mediziner und Medizinerinnen im Rahmen des persönlichen Arzt-Patienten-Gesprächs erheben, sind von KI-Systemen nicht leistbar.
- Verweisen Sie auf Chancen beim Einsatz von medizinischen KI-Anwendungen. Durch KI können Diagnosen und Therapien für Patienten und Patientinnen verbessert werden. Algorithmen könnten zeitaufwendige Aufgaben übernehmen, sodass mehr Zeit für das persönliche Arzt-Patienten-Gespräch zur Verfügung steht.
- Eine häufig gestellte Frage ist die nach der **Fehlerhäufigkeit** des jeweiligen KI-Systems.

Zu einer ehrlichen Antwort gehört die Tatsache, dass KI-Systeme von den Daten abhängig sind, die ihnen zur Verfügung gestellt werden. Schon darin kann eine mögliche Fehlerquelle liegen. Auch die Programmierung des Systems kann Fehlerquellen bergen. Umso wichtiger ist es, dass die KI-Analysen nur ein Werkzeug für Medizinerinnen und Mediziner sind, die möglicherweise fehlerhafte Interpretationen des Systems erkennen und richtig einordnen können. Die Ergebnisse einer medizinischen KI sollten daher stets nachvollziehbar sein.

- Zeigen Sie beim Thema **Datensicherheit**, wie wichtig auch Ihnen dies ist. Sie sind schließlich selbst nicht nur Forschende, sondern unter Umständen auch Betroffene. Erklären Sie konkret, wie Sie die Gesundheitsdaten von Patientinnen und Patienten, mit denen Sie arbeiten gleichzeitig schützen – auch gemeinsam mit Ärztinnen und Ärzten. Bringen Sie dabei auch zur Sprache, was Anonymisierung und Pseudonymisierung bedeutet und was die Verfahren leisten können – und was nicht.
- Weisen Sie neben möglichen Risiken auch auf Chancen hin. So kann eine Anonymisierung keinen hundertprozentigen Schutz der Daten gewährleisten. Dennoch kann ohne Daten auch kein Algorithmus trainiert werden, der die Diagnose oder Behandlung enorm verbessern würde.
- Erläutern Sie, warum es wichtig ist, für weitere KI-Anwendungen in der Medizin immer mehr Daten elektronisch zu erfassen: Das enorme Potenzial von KI kann umso besser ausgeschöpft werden, je mehr zuverlässige Daten für das Training der jeweiligen Algorithmen zur Verfügung stehen.
- Sensibilisieren Sie Laien für einen sorgsamen Umgang mit persönlichen Daten – insbesondere auch dafür, dass im Hinblick auf die Anhäufung von Daten weniger oft mehr ist. Daten sollten gezielt erhoben und ausgewertet wer-

den, um sinnvolle Zusammenhänge herstellen zu können. Dabei spielt auch die Datenhoheit eine wichtige Rolle.

- Machen Sie Ihre Gesprächspartner neben den Vorteilen auch auf die Problematik der über Fitness-Apps in diversen Netzwerken geteilten Daten aufmerksam.

LITERATUR

1. Kröplin, Tim: Dr. KI hat nun Zeit für Sie. Zeit Online (13/11/2018) <https://www.zeit.de/wissen/gesundheit/2018-11/bildererkennung-kuenstliche-intelligenz-gesundheits-arzt-diagnose-smart-devices>.
2. Ekkenkamp, Axel: Am Krankenbett mit Doktor KI. Der Tagesspiegel (24/07/2019) <https://www.tagesspiegel.de/wissen/kuenstliche-intelligenz-in-der-medizin-am-krankenbett-mit-doktor-ki/24691876.html>.
3. Bartens, Werner: Dr. Data. Süddeutsche Zeitung (18/03/2018) <https://www.sueddeutsche.de/wissen/medizin-dr-data-1.3908072>.
4. Buck, Christian: Dr. Algorithmus. Die Welt (26/02/2019) <https://www.welt.de/wirtschaft/bilanz/article189418949/Dr-Algorithmus-Kuenstliche-Intelligenz-in-der-Medizin.html>.
5. Urbaneck, Margarethe: Die natürliche Angst vor dem Algorithmus. Ärzte Zeitung Online (04/11/2019) <https://www.aerztezeitung.de/Wirtschaft/Die-natuerliche-Angst-vor-dem-Algorithmus-403227.html>.
6. Hütten, Felix: Die große Zukunft. Süddeutsche Zeitung Nr.283, Wissen (7. Dezember) <https://www.sueddeutsche.de/gesundheits/kuenstliche-intelligenz-medizin-1.4712637>.
7. Meier, Christian J.: Der ersehnte Konkurrent des Arztes. Riff Reporter https://www.riffreporter.de/ki-fuer-alle/kuenstliche_intelligenz_medizin/.
8. Menn, Andreas: „Algorithmen sehen mehr als Ärzte“. Wirtschaftswoche Online (22/02/2019) <https://www.wiwo.de/unternehmen/it/ki-in-der-medizin-algorithmen-sehen-mehr-als-aerzte/24025456-all.html>.
9. Tertilt, Mathias: Wie Krankheiten aus uns sprechen. Zeit Online (18/08/2016) <https://www.zeit.de/2016/33/menschliche-stimme-lun-ge-sprache-krankheit-mund-kehlkopf>.
10. Knight, Will: KI erkennt Depression. Heise Online (14/01/2019) <https://www.heise.de/tr/artikel/KI-erkennt-Depression-4271725.html>.
11. Lauterbach, Anastasia: KI in der Medizin bietet große Chancen – was fehlt, sind die Daten. Wirtschaftswoche Online (15/10/2019) <https://www.wiwo.de/erfolg/management/fuehrungswechsel-ki-in-der-medizin-bietet-grosse-chancen-was-fehlt-sind-die-daten/25117568.html>.
12. Höpner, Axel: So verändert Künstliche Intelligenz die Diagnostik. Handelsblatt Online (09/07/2019) <https://www.handelsblatt.com/technik/thespark/medizin-der-zukunft-so-veraendert-kuenstliche-intelligenz-die-diagnostik/24535864.html>.
13. Volkmann, Anne: Was kann Künstliche Intelligenz in der Medizin? Gesundheitsstadt Berlin (23/08/2019) <https://www.gesundheitsstadt-berlin.de/was-kann-kuenstliche-intelligenz-in-der-medizin-13571/>.
14. von Liebe, Sylvaine: Mit Smartwatches und Fitnessarmbändern Krankheiten verhindern? Bayerischer Rundfunk (22/05/2019) <https://www.br.de/nachrichten/wissen/mit-smartwatches-und-fitnessarmbaendern-krankheiten-verhindern,RQXVojS>.
15. Nagel, Till Simon: Medizin-Wearables gehen weiter als bisherige Diagnosemethoden. Die Welt (11/03/2019) <https://www.welt.de/wirtschaft/webwelt/article190092755/Medizin-Wearables-Was-taugen-die-neuen-Gesundheitsgeraete.html>.
16. Viciano, Astrid: Die Blackbox-Diagnosen. Süddeutsche Zeitung Nr.34 (09/12/2019) <https://www.sueddeutsche.de/wissen/big-data-die-blackbox-diagnosen-1.4321839>.
17. Bartens, Werner: Der große Bluff. Süddeutsche Zeitung Nr.283 (07/12/2019) <https://www.sueddeutsche.de/gesundheits/kuenstliche-intelligenz-medizin-1.4712639>.
18. Esteva, Andre; Kuprel, Brett et al.: Dermatologist-level classification of skin cancer with deep neural networks. Nature Nr.542 S. 115–118 (29/06/2017) DOI:10.1038/nature21056. <https://doi.org/10.1038/nature21056>.
19. Hänßle, Holger A. et al.: Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology Nr.29, Issue 8 S. 1836–1842 (08/2018) doi:10.1093/annonc/mdy166. <https://doi.org/10.1093/annonc/mdy166>.
20. Roemheld, Lars: Hype-basierte Medizin: Blockchain und KI. Tagesspiegel Background (22/08/2019) <https://background.tagesspiegel.de/digitalisierung/hype-basierte-medizin-blockchain-und-ki>.
21. Blasius, Helga: Handy-App zur Früherkennung von Parkinson. DAZ online (06/06/2018) <https://www.deutsche-apotheker-zeitung.de/news/artikel/2018/06/06/handy-app-zur-frueherkennung-von-parkinson>.
22. Thesing, Melissa: Unterschied zwischen Anonymisierung, Pseudonymisierung und Verschlüsselung. Datenschutz Notizen (11/01/2017) <https://www.datenschutz-notizen.de/unterschied-zwischen-anonymisierung-pseudonymisierung-und-verschluesselung-2916984/>.
23. Hurtz, Simon: Warum 'anonyme' Daten eine Illusion sind. Süddeutsche Zeitung (30/07/2019) <https://www.sueddeutsche.de/digital/anonyme-daten-studie-1.4542458>.
24. Schonschek, Oliver: Anonymisierung und Pseudonymisierung von Kundendaten. Datenschutz Praxis (12/06/2018) <https://www.datenschutz-praxis.de/fachartikel/anonymisierung-und-pseudonymisierung-von-kundendaten/>.
25. Jens Spahn weist Kritik an Sammlung von Patientendaten zurück. Zeit Online (07/11/2019) <https://www.zeit.de/politik/deutschland/2019-11/digitale-versorgung-gesetz-elektronische-patientenakte-gesundheitsapps-datenschutz>.
26. Digitale-Versorgung-Gesetz Was sich für Patienten ändert. tagesschau.de (07/11/2019) <https://www.tagesschau.de/inland/digitale-versorgung-gesetz-hintergrund-101.html>.
27. Hamich, Christopher: Bundestag beschließt Gesundheitsdatenbank und Apps auf Rezept. Netzpolitik.org (07/11/2019) <https://netzpolitik.org/2019/bundestag-beschliesst-gesundheitsdatenbank-und-apps-auf-rezept/>.
28. BVDW-Studie: Hohe Akzeptanz für künstliche Intelligenz und digitale Anwendungen im Gesundheitswesen. (25/06/2019) <https://www.bvdw.org/der-bvdw/news/detail/artikel/bvdw-studie-hohe-akzeptanz-fuer-kuenstliche-intelligenz-und-digitale-anwendungen-im-gesundheitswesen/>.
29. Befragung 'Datenspende für die medizinische Forschung'. TMF-Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. <http://www.tmf-ev.de/News/articleType/ArticleView/articleId/4456.aspx>.
30. Plattform Lernende Systeme (Hrsg.): Lernende Systeme im Gesundheitswesen – Bericht der Arbeitsgruppe Gesundheit, Medizintechnik Pflege, München 2019. <https://www.plattform-lernende-systeme.de/publikationen-details/lernende-systeme-im-gesundheitswesen.html>.
31. Videostatements und Diskussionsmitschnitt des RIKI-Projekts. <https://www.youtube.com/playlist?list=PLYFa5hehwobloEtRmlS-QSfBtmqsn3rSy9>.

→ Alle hier genannten Webseiten geprüft am 24.01.2020

KI UND ROBOTIK

1. MEDIALE BERICHTERSTATTUNG

Kaum ein Beitrag über Künstliche Intelligenz, der nicht mit einem Roboter bebildert ist¹. Sei es der weiße Roboter ASIMO, der einem Astronauten gleicht. Seien es die Roboter-Robbe Paro im Altenheiminsatz, der digitale Buttler Roboter Hollie, Sexpuppen mit Künstlicher Intelligenz oder intelligente Kuscheltiere im Kinderzimmer. So verwundert es eigentlich nicht, dass in der Wahrnehmung von Laien die Begriffe **KI und Robotik** gerne **verschwimmen**. Bei einer Umfrage in Großbritannien setzten 25 Prozent der Befragten Künstliche Intelligenz mit Robotern gleich.²

46

In dystopischen Berichten ist in den Medien darüber hinaus von „Killerrobotern“³⁻⁶ oder einer „**Robokalyipse**“⁷ die Rede. Die Vorstellung von übermächtigen, intelligenten Robotern wirft Fragen auf wie „Werden wir bald von Robotern beherrscht oder gar versklavt?“⁸

Dabei suggeriert die Verpackung oft mehr an Intelligenz als in den Maschinen bislang steckt. Roboter sind im Grunde Apparaturen, die bestimmte Funktionen ausführen können. Dies können sie zum einen dank der in ihnen steckenden KI, von der sie gesteuert werden. Und zum anderen können sie dies wegen ihrer Sensoren, die die Umgebung erfassen und der Mechanik, die in ihnen steckt. Da wir bisher jedoch nur sogenannte schwache KI-Systeme haben (siehe KI-Definitionen S. 9) sind Roboter noch sehr weit von den teils sogar übermenschlichen Fähigkeiten entfernt, die wir aus der Science Fiction kennen.

So gibt es auch realistisch einordnende Medienberichte, wie etwa von Michael Moorstedt, der in der Süddeutschen Zeitung nüchtern formuliert: „Die Roboter wissen nicht, was sie tun.“⁹ Maschinen wie Pepper, Sophia oder die Hindernisläufer von Boston Dynamics wirkten verblüffend menschlich. Dabei seien sie aber ferngesteuert und täuschten über den eigentlichen Stand der Robotertechnologie hinweg.

110101110101011101011101

In der Fertigung und Logistik werden seit Jahrzehnten Roboter eingesetzt. Sie können jedoch nur wenige, eindeutig definierte Aufgaben ausführen. Ein Beispiel ist etwa ein Schweißroboter in der Automobilindustrie. Wenn Roboter nun zunehmend als fahrerlose Transportsysteme in der Fabrik eingesetzt werden oder in der Pflege beim Umbetten von Patienten, dann steht das Thema Sicherheit des Menschen im Zentrum. Aus dem Roboter wird der sogenannte **Cobot**¹⁰, ein kollaborierender Roboter. Dafür brauche es eine neue Roboter-Generation, die etwa in Nicht-Standard-Situationen trainiert werde, um auf mögliche Gefahrensituationen wie Kollisionen zu reagieren und diese zu vermeiden.¹⁰

Neben dem Thema Sicherheit gibt es noch eine andere Sorge, etwa in Bezug auf Pflegeroboter. Denn einerseits schätzten 40 Prozent der Befragten des von acatech und der Körber Stiftung in Auftrag gegebenen Technikradars 2018 die technische Unterstützung durch Pflegeroboter positiv ein, andererseits lehnten 32 Prozent der Befragten diese Form der Hilfe jedoch grundsätzlich ab.¹¹ Womöglich ist der Grund der Ablehnung, dass über 80 Prozent damit rechnen, dass Pflegebedürftige durch den Einsatz der Pflegeroboter weniger menschliche Zuwendung erhalten.

„KRIEGE DER ZUKUNFT – KILLERROBOTER AUF DEM VORMARSCH“

Titelzeile Artikel von Markus Brauer/dpa, Stuttgarter Nachrichten⁴

00010010111101010111010

Es liegt auf der Hand, dass eine Bilderkennungssoftware nicht nur dazu eingesetzt werden kann, Katzen- von Hundebildern zu unterscheiden, sondern auch, um militärisch bedeutsame Objekte ausfindig zu machen.¹² Fast alle Entwicklungen der Robotik können sowohl zu zivilen als auch militärischen Zwecken genutzt werden. In der Tat wird KI bereits zu diversen militärischen Zwecken eingesetzt – beispielsweise in der Logistik oder in der Cyber- und Raketenabwehr. Die USA, Großbritannien, China und Russland setzen zunehmend auf militärische Systeme, die über einen gewissen Grad an Intelligenz verfügen.⁶

Bedenken gegenüber dem Einsatz von KI zu militärischen Zwecken werden in der medialen Berichterstattung vor allem im Zusammenhang mit autonomen Waffensystemen genannt.^{6,13} „Im Gegensatz zu ferngesteuerten Drohnen, die jetzt schon in Kriegsgebieten im Einsatz sind, arbeiten autonome Waffen ganz ohne menschliche Steuerung“, schreibt Julia Merlot auf Spiegel Online.¹³ „Künftig könnten autonome Waffen damit in deutlich komplexeren Kriegssituationen entscheiden, wen sie in welcher Form angreifen.“

2. AUS SICHT DER FORSCHENDEN

Im Verlauf der RIKI-Workshops kam das Thema Robotik vor allem am Standort in Bremen aber

Wie ähnlich wird KI einem menschlichen Gehirn?

Julian Bitterwolf
Eberhard Karls
Universität Tübingen



Soll KI künftig Entscheidungen treffen?

Alexander Fabisch
DFKI, Bremen



Wird KI eines Tages klüger sein als der Mensch?

Matthias Tangemann
Eberhard Karls
Universität Tübingen



Kann man einem Roboter alles befehlen?

Fabian Paus
Karlsruher Institut für
Technologie (KIT)



Videostatements von Forschenden sind online abrufbar.¹⁴

auch an den Standorten Tübingen, Saarbrücken und Karlsruhe zur Sprache.

In der Tiefsee oder auf dem Mars, im OP-Saal oder in der häuslichen Pflege, als Haushaltshelfer oder Bergungsroboter – die Einsatzmöglichkeiten der Robotik erstrecken sich über viele **verschiedene Bereiche**. Die Forschenden stellten heraus, dass die Robotik den Menschen idealerweise bestmöglich unterstützt. Sie verwiesen beispielhaft auf

47

Roboter, die in schwer zugänglichen oder menschenfeindlichen Umgebungen eingesetzt werden könnten.

Die Robotik profitiere laut den Workshop-Teilnehmenden dabei aktuell enorm von jüngsten Entwicklungen im Forschungsfeld der Künstlichen Intelligenz. Komplexe Bewegungsabläufe, die mit großem Rechenaufwand verbunden sind, könnten nun realisiert werden. Zudem ließen sich heute schon Roboter teils ohne zusätzlichen Programmieraufwand in unterschiedlichen Umgebungen einsetzen.

010100100101011101010111010

Die Mehrheit der Forschenden sieht jedoch ausgerechnet in den vielfältigen Einsatzmöglichkeiten auch Risiken. Unter dem Stichwort „Dual Use“ diskutierten die Teilnehmenden den Einsatz von Robotik und KI auch zu militärischen Zwecken. Auf einen Roboter könne beispielsweise ein Maschinengewehr montiert werden, und schon werde aus dem Helfer eine Waffe. Doch auch andere Technologien wie Blickerfassung oder Objekterkennung bieten sich für eine militärische Verwendung geradezu an.

So zeigten sich viele Forschende davon überzeugt, dass Künstliche Intelligenz sowohl vermehrt im zivilen Bereich wie auch zu militärischen Zwecken verwendet werde. Die Mehrheit der Forschenden erwartet, dass sich auf lange Sicht keine Nation solch einer Entwicklung verschließen kann und fürchtet in diesem Zusammenhang ein neuerliches Wettrüsten verfeindeter Mächte.

Die Teilnehmenden der Workshops sprachen sich allerdings klar gegen eine **militärische Verwendung** von KI und Robotik aus. Die Mehrheit lehnte eine Zusammenarbeit mit Institutionen ab, die KI erklärtermaßen militärisch einsetzen wollen. Um einen Dual Use für militärische Zwecke zumindest

zu erschweren, veröffentlichten einige Forschende bereits ihre Programmcodes mit Lizenzzeilen, die eine militärische Verwendung ausschließen sollen. Sie gaben jedoch zu bedenken, dass auch dies nicht sicher verhindern könne, dass der Code letztlich auch für militärische Zwecke verwendet wird. Die Expertinnen und Experten sehen in dieser heiklen Frage einen Regulierungsbedarf und die Notwendigkeit gesellschaftlicher Diskurse.

0101001110101010101010111010

Um einer KI Verantwortung übertragen zu können, müsse das jeweilige System aus Sicht der Teilnehmerinnen und Teilnehmer der Workshops und Fishbowl-Debatten klar definierten Regeln folgen, die Menschen festgelegt sind und die diese jederzeit wieder ändern könnten. Darüber hinaus müssten die **Entscheidungen** der KI **transparent**, **regelkonform** und **überprüfbar** sein, um zu gewährleisten, dass das KI-System den menschlichen Vorstellungen entsprechend handelt. Zumindest zum jetzigen Zeitpunkt sei es aber viel zu früh, einer KI Entscheidungen zu überlassen.

Ein weiterer, viel diskutierter Punkt zum Thema Verantwortung war: Wer steht bei einem von einem Roboter verursachten Fehler, etwa einer Kollision oder anderem, in der Verantwortung? Die Programmierer? Die Besitzer? Oder gar die KI selbst? Müsse man eventuell Robotern mit Künstlicher Intelligenz je nach Grad der Eigenständigkeit beispielsweise eine Rechtspersönlichkeit zusprechen?

010100100101011101010111010

„Habt ihr nicht Angst, dass der Roboter abends aus eurem Forschungslabor abhaut?“ – diese oder ähnliche Fragen zeigten den Wissenschaftlerinnen und Wissenschaftlern, welche verzerrten und falschen Vorstellungen Laien häufig hätten. Bei bestimmten Aufgaben könnten Roboter zwar in klar

definierten Tätigkeiten Fähigkeiten von Menschen übertreffen. Aber Roboter seien bis heute nicht mehr als hochentwickelte Werkzeuge, wie sie betonten. Es bedürfe noch vieler Entwicklungsarbeiten, ehe Roboter mit komplexeren, menschlichen Fähigkeiten mithalten könnten.

3. KOMMUNIKATIONSTIPPS VON FORSCHENDEN FÜR FORSCHENDE

- Klären Sie zu Beginn, wie Robotik und Künstliche Intelligenz miteinander verbunden sind.
- Machen Sie sich vor einem Gespräch bewusst, welche Voraussetzungen Ihr Gegenüber mitbringt: Sind die Vorstellungen maßgeblich geprägt durch Science-Fiction-Figuren wie Terminator oder R2D2? Oder haben Sie es mit jemandem zu tun, der bereits mit dem Thema vertraut ist?
- Ist KI klüger als der Mensch? Bei solchen Fragen ist es ratsam, den Unterschied zwischen starker und schwacher KI anzusprechen. Diese Unterscheidung sollte zu einer realistischen Einschätzung der Dinge beitragen (siehe S.9).
- Wenn Ihr Gegenüber Sorge hat, dass Roboter die Menschen bald übertrumpfen und im äußersten Fall sogar die Menschheit unterwerfen könnten, ist es höchste Zeit, ein realistisches Bild zu zeichnen: Roboter sind hochentwickelte Werkzeuge, die den Menschen entlasten und unterstützen können. Arbeiten Sie dabei mit Beispielen, etwa mit den von der Pariser Feuerwehr in Notre Dame nach dem Brand in einsturzgefährdeten Bereichen eingesetzten Robotern.
- Menschen neigen dazu, menschliche Eigenschaften auf Roboter zu übertragen.¹⁵⁻¹⁷ Versuchen Sie Ihr Gegenüber dafür zu sensibilisieren. Ein Roboter ist weder traurig, noch verfügt er über ein eigenes Bewusstsein.
- Dual Use: Dass technologischer Fortschritt fast immer nicht nur zu zivilen, sondern auch zu

militärischen Zwecken genutzt wird, ist nicht neu. Dies lässt sich über die gesamte Ära der Technikentwicklung nachzeichnen. Dual Use ist also kein spezifisches Problem von KI. Das entbindet insbesondere Sie als KI-Forschende aber nicht davon, sich darüber Gedanken zu machen. Versuchen Sie für sich eine klare Position zu beziehen, die Sie auch gut vertreten und vermitteln können.

- Die Robotik ist vielseitig einsetzbar – im Guten wie im Bösen. Wenn Ihre Forschung häufig Kritik ausgesetzt ist, nehmen Sie dies keinesfalls persönlich, sondern betrachten Sie dies als notwendigen gesellschaftlichen Prozess. Auch wenn Kritik nicht immer sachlich vorgebracht wird – bleiben Sie gelassen, achten Sie auf eine neutrale Wortwahl, bedienen Sie sich einer klaren und verständlichen Ausdrucksweise, räumen Sie Risiken ein, verweisen Sie aber stets auch auf die Chancen.
- Zögern Sie nicht in einem Gespräch Gegenfragen zu stellen und so Missverständnissen vorzubeugen. Ein Beispiel: Ist KI gefährlich? Versuchen Sie zu präzisieren, ob sich der oder die Fragende dabei auf KI oder etwa einen Roboter bezieht. Indem Sie die Frage präzisieren, haben Sie die Möglichkeit differenziertere Antworten zu geben. Präzisierende Rückfragen fördern zudem einen Austausch auf Augenhöhe.
- Weisen Sie auf ungeklärte Fragen hin und regen Sie einen Meinungs-austausch darüber an: Wollen wir KI Entscheidungen überlassen? Wo sollen Roboter eingesetzt werden? Wer trägt die Verantwortung im Falle von Fehlern mit Folgen?

LITERATUR

1. Bild der Wissenschaft. Spezial.: Künstliche Intelligenz. S. 7, 12, 37–41, 49, 57, 58, 61–69, 76, 90,91 (02/05/2019).

2. Cave, Stephen & Coughlan, Kate: 'Scary Robots': Examining Public Responses to AI. researchgate (Jauar/2019) https://www.researchgate.net/publication/334386635_Scary_Robots_Examining_Public_Responses_to_AI.

3. Hielscher oder Haase: Vollautomatische Killer-Roboter: UN-Staaten können sich nicht einigen. <https://www.deutschlandfunknova.de/beitrag/krieg-un-staaten-sprechen-ueber-killerroboter>.

4. Brauer, Markus/dpa: Killerroboter auf dem Vormarsch. Stuttgarter Nachrichten (20/08/2019) <https://www.stuttgarter-nachrichten.de/inhalt.kriege-der-zukunft-killerroboter-auf-dem-vormarsch.172527fd-48ee-4fe9-8616-dd0be619f145.html>.

5. 120 Staaten beschließen Regeln für Killerroboter. Tagesspiegel (16/11/2019) <https://www.tagesspiegel.de/politik/ruestungsgegner-fordern-verbot-120-staaten-beschliessen-regeln-fuer-killerroboter/25236098.html>.

6. Schwarz, Elke: Die Killerroboter kommen. Süddeutsche Zeitung (14/10/2019) <https://www.sueddeutsche.de/kultur/essay-die-killerroboter-kommen-1.4638223>.

7. Paradies oder Robokalypse? Das Erste (15/04/2019) <https://www.daserste.de/information/reportage-dokumentation/dokus/sendung/paradies-oder-robokalypse-106.html>.

8. Thomsen, Prof. Bernd: Keine Angst vor der KI! Was Menschen Maschinen voraushaben. Handelsblatt (11/01/2019) <https://www.handelsblatt.com/meinung/kolumnen/expertenrat/thomsen/expertenrat-prof-bernd-thomsen-keine-angst-vor-der-ki-was-menschen-maschinen-voraushaben/23856180.html>.

9. Moorstedt, Michael: Die Roboter wissen nicht, was sie tun. Süddeutsche Zeitung (22/10/2018) <https://www.sueddeutsche.de/digital/kuenstliche-intelligenz-die-roboter-wissen-nicht-was-sie-tun-1.4178766>.

10. Welchering, Peter: Der Roboter als Sensibelchen. Deutschlandfunk (16/03/2019) https://www.deutschlandfunk.de/mensch-roboter-interaktion-der-roboter-als-sensibelchen.684.de.html?dram:article_id=443827.

11. acatech & Körber-Stiftung: Technikradar 2018. https://www.koerber-stiftung.de/fileadmin/user_upload/koerber-stiftung/redaktion/technikradar/pdf/2018/Technikradar-2018_Langfassung.pdf.

12. Peitz, Dirk: Google wird einfach ersetzt. Zeit Online (05/06/2018) <https://www.zeit.de/digital/internet/2018-06/maven-militaerprojekt-google-ausstieg-ruestungsexperte-paul-scharre>.

13. Merlot, Julia: Autonome Waffen außer Kontrolle. Spiegel Online (24/02/2019) <https://www.spiegel.de/wissenschaft/technik/autonome-waffen-ausser-kontrolle-a-1253320.html>.

14. Videostatements des RIKI-Projekts. <https://www.youtube.com/playlist?list=PLYFa5hehwobjjh1EzdHD3TjySB98UCZJW>.

15. Prüfer, Tillmann: Roboter. Sie kommen. Zeit Magazin Nr.21/2019 (15/05/2019) <https://www.zeit.de/zeit-magazin/2019/21/roboter-mensch-technologie-robotik-privatleben/seite-2>.

16. Gruber, Angela: Gefühle für Maschinen. Wenn die Roboter kommen, werden wir sie lieben. (11/01/2017) <https://www.spiegel.de/netzwelt/gadgets/anthropomorphismus-koennen-menschen-roboter-lieben-a-1128553.html>.

17. Seng, Leonie: Ganz schön menschlich. Deutschlandfunk (01/11/2012) https://www.deutschlandfunk.de/ganz-schoen-menschlich.676.de.html?dram:article_id=226262.

→ Alle hier genannten Webseiten geprüft am 24.01.2020

IMPRESSUM

HERAUSGEBER

Nationales Institut für
Wissenschaftskommunikation
(NaWik) gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg
www.nawik.de

REDAKTION

Virginia Albert
Beatrice Lugger
Klaus Wingen

FOTOS

Virginia Albert
Beatrice Lugger

VIDEOPRODUKTION

Redaktion: Virginia Albert
Kamera, Ton & Schnitt: teamWERK.
Die Filmproduktion GmbH

GRAFIKDESIGN/
ENTWICKLUNG

Chris Spatschek

GRAFISCHE GESTALTUNG
KOMPENDIUM

Lena Haselmann

STAND

Januar 2020

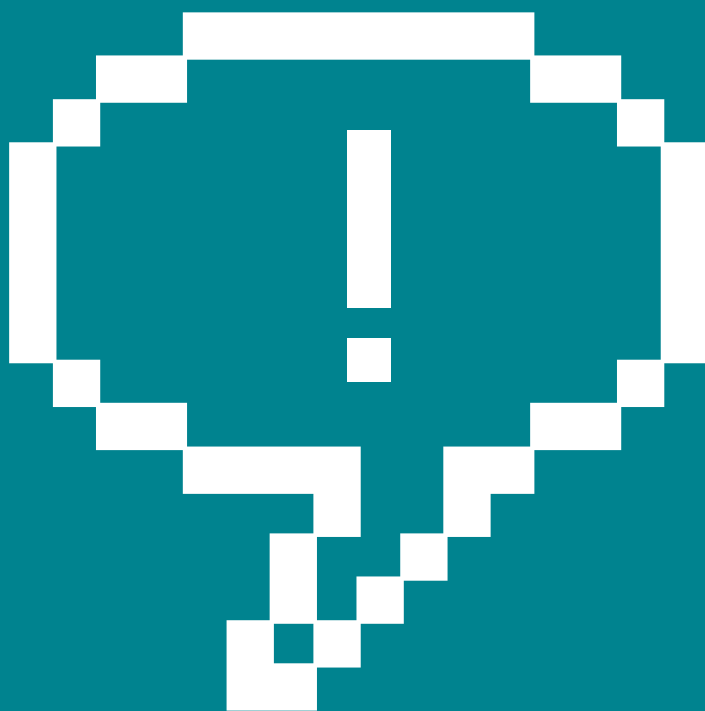
Bei Fragen oder Anmerkungen zu dieser
Publikation kontaktieren Sie uns über
nawik@nawik.de

Folgen sie uns auf Twitter: @NaWik

EMPFOHLENE ZITIERWEISE

Nationales Institut für Wissenschaftskommunikation (NaWik) gGmbH (Hrsg.): Risikokommunikation zur Künstlichen Intelligenz. Kompendium mit Kommunikationstipps von Forschenden für Forschende, Karlsruhe 2020.

ISBN: 978-3-00-064939-4



ISBN 978-3-00-064939-4

Nationales Institut für
Wissenschaftskommunikation



Wissenschaft.
Verständlich.



Risikokommunikation
zur Künstlichen Intelligenz

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Eine Initiative des Bundesministeriums
für Bildung und Forschung

Wissenschaftsjahr 2019

**KÜNSTLICHE
INTELLIGENZ**