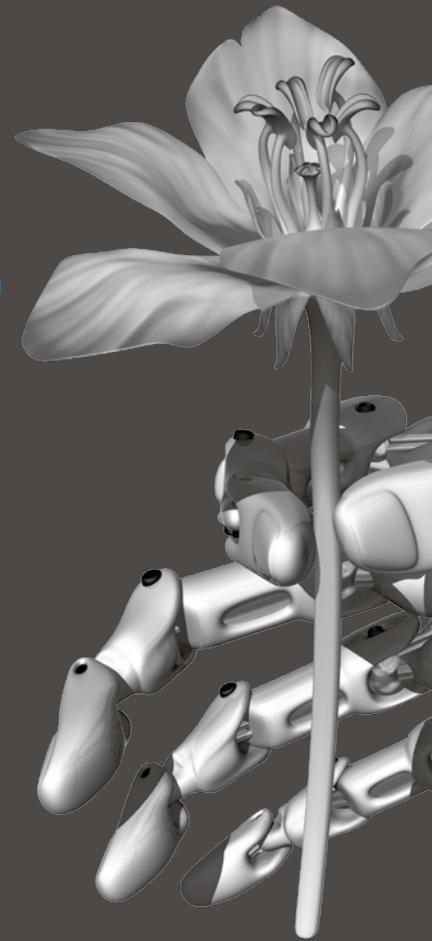


▶ Wolfgang Wahlster: „Künstliche Intelligenz ist besser als natürliche Dummheit“

Roboter, die im Haushalt helfen und alte Menschen pflegen, Systeme mit künstlicher Intelligenz (KI), die sich selbst mit Hilfe von Selbstlern-Algorithmus laufend optimieren – diese Entwicklungen wecken Befürchtungen, Maschinen könnten die Herrschaft über die Menschen übernehmen. Professor Wolfgang Wahlster, Vorsitzender der Geschäftsführung und Wissenschaftlicher Direktor am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH in Saarbrücken, erläutert im Gespräch mit der COMPUTERWOCHE, warum er diese Befürchtungen hinsichtlich der Entwicklung humanoider, selbstlernender Systeme nicht teilt. Er spricht aber sehr wohl auch über die Risiken, die mit der fortschreitenden Entwicklung „intelligenter“ Systeme verbunden sind.



Von Jan-Bernd Meyer,
leitender Redakteur

Mehr zum Thema
künstliche Intelligenz
finden Sie online auf der
Website der COMPUTER-
WOCHE unter:

www.cowo.de/a/2552792

www.cowo.de/a/2557874

www.cowo.de/a/3063495

www.cowo.de/a/3094924

CW: Prominente Unternehmer und Wissenschaftler warnen vor den Gefahren durch künstliche Intelligenz. Stellen solche Systeme eine Gefahr für die Menschen dar?

WAHLSTER: Ich beschäftige mich seit über 30 Jahren mit dem Thema KI. Wenn man so lange an diesem Thema arbeitet, bekommt man trotz aller Forschungserfolge immer mehr Respekt vor der menschlichen Intelligenz. Trotz aller wissenschaftlichen Durchbrüche der KI, an denen auch deutsche Wissenschaftler maßgeblich beteiligt waren, wie dem Sieg über einen Schachweltmeister, der Dialogübersetzung auf dem Smartphone oder dem selbstfahrenden Auto, ist es Science-Fiction, zu behaupten, dass die künstliche Intelligenz sich verselbständigen könne und die Menschheit bedrohe.

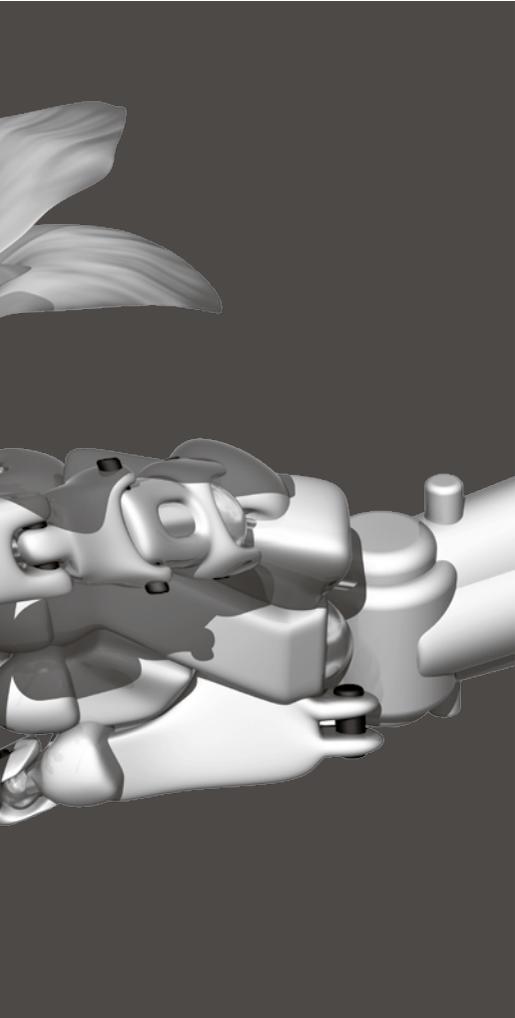
CW: Warum kann das nicht passieren?

WAHLSTER: Maschinelles Lernen unterscheidet sich prinzipiell vom menschlichen Lernen. Es geht in der KI keinesfalls darum, menschliches Verhalten mit all seinen Stärken, aber auch den vielen Schwächen perfekt nachzubilden,

sondern wir wollen dem Menschen mit KI assistieren, wo dies sinnvoll ist. Beim Lernen von Regelmäßigkeiten in sehr großen Datenmengen sind maschinelle Lernverfahren dem Menschen heute bereits überlegen. Andererseits kann aber Common Sense, der auch die soziale und emotionale Intelligenz des Menschen stark fordert, bislang nur äußerst beschränkt von maschinellen Systemen erworben werden. Technologien der KI zur automatischen Perzeption, zum Verstehen, Lernen, Suchen und Schlussfolgern nutzen wir heute beispielsweise auf Smartphones tagtäglich: bei der Wegsuche, im Sprachdialog mit Assistenzsystemen wie Siri und Cortana oder in Empfehlungssystemen auf Web-Portalen. Dem liegen zwar äußerst komplexe Softwaresysteme zugrunde, aber ich kann hier kein prinzipielles Problem einer sich aufbauenden Superintelligenz erkennen.

CW: Aber Firmen forschen daran. Kann dabei so eine Superintelligenz herauskommen?

WAHLSTER: Da Internet-Firmen derzeit stark im Bereich der KI investieren, entsteht schnell auch ein Hype-Zyklus, den wir leider in der



Geschichte der KI schon mehrfach durchlaufen haben. Dabei werden oft durch Persönlichkeiten, die in der internationalen Forschergemeinschaft zur KI selbst wissenschaftlich gar nicht anerkannt sind, in der Öffentlichkeit mit populärwissenschaftlichen Texten überzogene Erwartungen an die Leistungsfähigkeit von bestimmten Technologien erweckt. Dies war in der Vergangenheit schon bei Begriffen wie „Neuronale Netze“ und später „Expertensysteme“ der Fall, und jetzt wird rund um das maschinelle Lernen erneut eine solche Blase erzeugt. Die deutsche Forschung ist immer gut damit gefahren, hier auf dem Boden der Tatsachen zu bleiben, seriöse Grundlagenforschung zu betreiben, sinnvolle Anwendungen voranzutreiben und zusammen mit der Industrie umzusetzen. Dadurch blieben wir in Deutschland von den sogenannten KI-Wintern in den USA verschont, wenn völlig überzogene Erwartungen herb enttäuscht wurden und die Forschungsinvestitionen dann drastisch heruntergefahren wurden. KI kann auch als Abkürzung für „künftige Informatik“ verstanden werden: Die künstliche Intelligenz bildet eine Avantgarde in der Informatik, weil dieses Forschungs-

gebiet seit seiner Gründung immer wieder an die jeweiligen Grenzen der Digitalisierung und Algorithmisierung heranführt.

CW: Aber Skeptiker wie der Technikvisionär Raymond Kurzweil finden viel Gehör.

WAHLSTER: Ja, leider. Aber unter Wissenschaftlern sind seine Thesen sehr umstritten. In den USA haben sich namhafte Informatiker von Kurzweils extremen Vorstellungen einer Superintelligenz bis hin zur Unsterblichkeit klar distanziert. Trotz der großen Fortschritte auf vielen Teilgebieten der KI sind wir selbst von den alltäglichen Intelligenzleistungen eines Menschen noch weit entfernt und erreichen hier nicht einmal Vorschulniveau. Laien finden es oft paradox, dass dagegen KI-Systeme in speziellen Bereichen tatsächlich dem Menschen überlegen sind. So bestätigen menschliche Weltmeister, daß die Spielstärke der besten Backgammon-Programme inzwischen höher als die der besten Spieler ist. Künstliche Intelligenz ist besser als natürliche Dummheit, aber natürliche Intelligenz ist bei sehr vielen Aufgabenstellungen besser als künstliche Intelligenz. Wenn Unsicherheit, Vagheit und wenig Erfahrungswerte eine Problemstellung dominieren oder soziale und emotionale Faktoren eine wesentliche Rolle bei der Problemlösung spielen, haben heutige KI-Systeme kaum Chancen. So könnte kein heutiges KI-System einem Kind das Fahrradfahren beibringen. Warum auch? Das sollte weiterhin eine schöne Herausforderung für die Eltern mit all ihrer Empathie bleiben. Trotz der Fehlbehauptungen von Kurzweil muss man die Ängste als Reaktion auf dessen Thesen aber sehr ernst nehmen und die Risiken von selbstlernenden Systemen offen adressieren.

CW: Wie könnte das aussehen?

WAHLSTER: Erstens muss natürlich mit Hochdruck an der Zuverlässigkeit, Sicherheit und der Resilienz von KI-Systemen gearbeitet werden, wenn sie für kritische Steuerungen oder

Entscheidungen eingesetzt werden. Cyber-Angriffe auf KI-Systeme sind besonders gefährlich, weil aufgrund der extremen Komplexität dieser Softwaresysteme Verhaltensanomalien nicht sofort auffallen müssen. Zweitens gilt bei KI-Systemen natürlich auch, dass diese umfangreichen Softwaresysteme fast immer Programmierfehler enthalten, so dass man sich nie zu 100 Prozent darauf verlassen sollte. Die Forschung hat in den letzten Jahren allerdings sehr leistungsfähige Deduktionssysteme entwickelt, welche zumindest für kleine, aber sicherheitskritische Module helfen, Programmierfehler nachweisbar auszuschließen. Und drittens ist die Gefahr des kriminellen Missbrauchs von KI-Technologien natürlich ein permanentes Risiko, beispielsweise im Bereich des Terrorismus oder der Finanzunterwelt.

CW: Was bedeutet KI für die Interaktion zwischen Mensch und Maschine zum Beispiel im Umfeld Industrie 4.0?

WAHLSTER: In der Fabrik der Zukunft werden sich Mensch und Maschine komplementär ergänzen. Mit der neuen Generation von kollaborativen Leichtbaurobotern werden die Werker in der Smart Factory zusammenarbeiten. Allerdings wird dabei eine klare Rollenverteilung bestehen: Der Mensch nutzt den Roboter als Assistenten beispielsweise bei der Überkopfarbeit im Automobilbau. Während die bisherigen Roboter in Sicherheitskäfigen Schwerstarbeit wie Punktschweißen oder Lackieren verrichteten, sind die neuen kollaborativen Roboter für den direkten Kontakt mit Menschen ausgelegt und können diesen auf keinen Fall verletzen. Sie weichen dem Menschen sogar sensorgesteuert bei der gemeinsamen Arbeit aus. Wie gute Assistenten reichen diese Roboter dem Menschen die richtigen Bauteile und Werkzeuge im passenden Moment an und helfen dem Facharbeiter. Dazu müssen sie KI-Verfahren zur Erkennung der Handlungsabsicht des Menschen nutzen, denn der Werker soll ja in Industrie 4.0 selbst den Arbeitstakt entscheiden und nicht von einem Roboter gegängelt werden. ▶



Zur Person

Wolfgang Wahlster ist Professor für Informatik an der Universität des Saarlandes und leitet seit 1988 als Direktor und Vorsitzender der Geschäftsführung das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI GmbH) in Saarbrücken, Kaiserslautern, Bremen und Berlin. Das DFKI arbeitet derzeit im Rahmen des vom Bundesforschungsministerium geförderten Projekts Hysociatea daran, wie künftig mit Hilfe von Softwaresystemen hybride Teams aus Mitarbeitern und verschiedenen Robotern für gemeinsame Problemlösungen im Umfeld von Industrie 4.0 koordiniert werden können. Hier geht es auch darum, wie ein solches hybrides Team die verschiedenen Teilaufgaben an die jeweils geeignetsten Akteure verteilt und den gemeinsamen Fortschritt verfolgt. Wahlster sagt dazu: „Es ist also in der künstlichen Intelligenz keinesfalls das Ziel, einen digitalen Homunkulus zu erschaffen. Dagegen sollen die Sinneswahrnehmung, die Motorik, die Lernfähigkeit, die Verhaltensplanung und die Inferenzfähigkeit des Menschen komplementär dort ergänzt werden, wo es im Hinblick auf Assistenzfunktionen für den Menschen sinnvoll ist.“

► **CW:** *Schwierig wird es aber doch, wenn es um Entscheidungen geht – im Hochfrequenzhandel mit Wertpapieren und beim Lenken eines Flugzeugs oder selbstfahrenden Autos.*

WAHLSTER: Das sind sehr gute Beispiele, bei denen die Transparenz der vorgeschlagenen Entscheidungen für die Akzeptanz eine wichtige Rolle spielt. Gerade beim maschinellen Lernen über Massendaten sollte ein KI-System mit einer Erklärungskomponente ausgestattet sein, welche die Entscheidungsvorschläge dem Menschen zumindest plausibel macht. Letztlich kann kein Mensch die Arbeitsweise eines komplexen maschinellen Lernverfahrens über Terabytes von Trainingsdaten im Detail nachvollziehen. Man muss bei maschinellen Lernergebnissen auch immer sehr kritisch nachprüfen, ob nicht durch zu viel Trainingszyklen eine Überanpassung stattgefunden hat, die letztlich dazu führt, dass nach einer Sättigungsphase die Generalisierungsleistung wieder schlechter wird.

CW: *Das bedeutet, KI macht auch Fehler?*

WAHLSTER: Das bedeutet – wir müssen darauf bestehen, Erklärungskomponenten für selbstlernende KI-Systeme zur Verfügung gestellt zu bekommen, die in Echtzeit Lernergebnisse zumindest plausibilisieren und kritische Nachfragen des letztlich verantwortlichen Menschen beantworten können. Diese müssen in der Lage sein, auf menschlich verständliche Weise darzulegen, warum das System eine bestimmte Entscheidung vorschlägt. Beispiel: Es wird ein Störfall in der Leitwarte eines Atomkraftwerks gemeldet. Dahinter arbeitet ein automatisiertes System mit Lernmechanismen. Das empfiehlt „Achtung, wir müssen hier eine Notabschaltung betreiben und die Steuerstäbe einfahren.“ Dann muss der Mensch in der Lage sein zu fragen: „Warum soll ich das eigentlich jetzt machen? Was ist die Erklärung dafür?“ Wenn der Mensch diese Erklärung dann überhaupt nicht versteht, sollte er den Vorschlag der Maschine auch nicht ausführen. Oder

nehmen Sie ein Beispiel aus der Medizin: Ein Watson-System (ein von IBM entwickeltes selbstlernendes Softwaresystem, Anm.d.Red.) kann einem Arzt gewisse Behandlungsmethoden empfehlen und zu einer gefährlichen Operation raten. In solch einem Fall muss das System in der Lage sein, dem Arzt zu erklären, wie es auf diese Handlungsoption gekommen ist. Dabei muss Watson wie ein Kollege agieren und nicht einfach sagen, es habe bis jetzt schon zehn Millionen Datenpunkte ausgewertet. Vielmehr muss es in einer für den Arzt verständlichen Form seine Vorschläge erklären können, damit der Arzt Vertrauen in die Lösungsangebote hat. Denn letztlich trägt er die Verantwortung gegenüber dem Patienten.

CW: *Ist denn dem Urteil eines Arztes mehr zu trauen als der Auswertung von Abermillionen Datensätzen – da kann doch kein Arzt mit seinem Wissen mithalten?*

WAHLSTER: Watson kenne ich sehr gut, denn ich habe mit SmartWeb eines der Vorgängerprojekte in Deutschland geleitet. Wir dürfen solche Systeme, die auf der intelligenten Auswertung von Daten beruhen, nicht als Orakel missbrauchen, sondern sie zwar als wichtige Informationsquelle, aber mit einer notwendigen kritischen Distanz benutzen. Denn letztlich ist es der Arzt, der haftbar gemacht wird, wenn einem Patienten durch eine Fehlbehandlung Schaden zugefügt wird. Auch vor Watson konnte ein Arzt sich nicht darauf zurückziehen, dass in seinem medizinischen Lehrbuch durch einen Druckfehler eine falsche Dosierung eines Medikaments empfohlen wurde. So sieht es unser Rechtssystem vor. Und deshalb muss das System den Arzt überzeugen. Das ist technisch eine große Herausforderung. Wir müssen selbstlernende Systeme als wertvolle Hilfen bei der Entscheidungsunterstützung nutzen, ohne dauerhaft alle Entscheidungen an solche Systeme zu delegieren. Auch der Autopilot muss von einem menschlichen Piloten wegen einer klar erkennbaren Fehlfunktion abgeschaltet werden können.