Foto: Kai R. Joachim

«In manchen Kontexten sollten wir keine autonomen Systeme einsetzen»

Digitale Ethik // Oliver Bendel forscht und publiziert seit Jahren über Informations-, Roboter- und Maschinenethik und wurde als Sachverständiger mehrfach in den Deutschen Bundestag und ins Europäische Parlament eingeladen. Wir haben mit ihm über die Chancen von KI, die Moral in der Maschine und Schuld und Verantwortung von autonomen Systemen gesprochen.

Prof. Dr. Oliver Bendel

Professor für Informationsethik, Roboterethik und Maschinenethik, FHNW

Oliver Bendel studierte Philosophie und Germanistik sowie Informationswissenschaft an der Universität Konstanz und promovierte an der Universität St. Gallen über Conversational Agents. Seit 2009 ist er Professor an der FHNW mit den Schwerpunkten Informationsethik, Roboterethik und Maschinenethik und untersucht aus diesen Perspektiven Roboter und KI-Systeme. Er war mehrfach Sachverständiger und Vortragender im Deutschen Bundestag und im Europäischen Parlament, die von ihm herausgegebenen Bücher «Pflegeroboter» (2018), «Handbuch Maschinenethik» (2019), «Maschinenliebe» (2020) und «Soziale Roboter» (2021) sind Standardwerke.

Herr Bendel, macht Ihnen die Entwicklung der KI Angst? Welche sind für Sie die grössten Gefahren von künstlicher Intelligenz?
Ich beschäftige mich seit den 1990er-Jahren mit künstlicher Intelligenz. Angst habe ich seitdem nie verspürt. Angst ist auch ein schlechter Ratgeber. Ich kann nicht allgemein von Gefahren von KI sprechen, da es so viele und so unterschiedliche Systeme und Anwendungen gibt. Wir beschreiben in unseren Arbeiten etwa Text-, Bild- und Videogeneratoren und arbeiten die Chancen und Risiken heraus. Auch Gesichts- und Stimmerkennung schenken wir unsere Aufmerksamkeit.

Welche ethischen Aspekte machen Ihnen im Zusammenhang mit KI die grössten Sorgen?
Einige Ausprägungen von KI sind sehr mächtig, etwa im Bereich der generativen KI. Man kann mit ihnen neue Gift- und Kampfstoffe erfinden. Allerdings kann man ebenso neue Medikamente entwickeln. In der Zukunft kann man mit KI-Systemen etliche Krankheiten besiegen. Meta will mit Hilfe von KI die biologischen und chemischen Prozesse im menschlichen Körper beobachten, messen und analysieren. Es ist wichtig, dass wir in der Ethik solche Chancen herausarbeiten.

Was ist künstliche Intelligenz? Sind die Verarbeitung von Big Data und das statistische Ausrechnen von Wahrscheinlichkeiten per Definition intelligent? Ist ChatGPT intelligent?
Der Begriff ‹Künstliche Intelligenz› steht für einen eigenen wissenschaftlichen Bereich der Informatik, der sich mit dem menschlichen Denk-, Entscheidungs- und Problemlösungsverhalten beschäftigt, um dieses durch computergestützte Verfahren ab- und nachbilden zu können. Die Intelligenz von Maschinen selbst kann ebenfalls mit dem Begriff gemeint sein, also die künstliche Intelligenz als Gegenstand und Ergebnis. Unter diesen weiten Begriff fallen auch herkömmliche, regelbasierte Chatbots. Und natürlich ChatGPT, wobei hier Machine Learning angewandt wird, genauer gesagt Reinforcement Learning from Human Feedback. Es geht nicht darum, dass Maschinen im menschlichen Sinne intelligent sind. Es geht darum, dass sie menschliche Intelligenz mit technischer Hilfe ab- und nachbilden.

Bedeutet die Tatsache, dass KI nur menschliche Intelligenz simulieren kann, dass sie nie intelligenter als Menschen werden kann?
KI-Systeme können auch tierische Intelligenz simulieren. Oder eine vorgestellte Intelligenz, wie es im Falle der Superintelligenz sein könnte. Und ohne Zweifel können KI-Systeme menschliche Fähigkeiten übertreffen. Das tun sie ja bereits seit Jahren, denken wir nur einmal ans Schachspielen. Wenn wir Intelligenz als komplexes System bestimmter Lebewesen ansehen, kann man natürlich sagen, dass KI diese Intelligenz nicht erreichen kann. Dafür müsste sie ja wie das Lebewesen werden. Aber wenn wir bestimmte Ausprägungen dieser Intelligenz nehmen, werden wir täglich von der KI übertroffen.

Menschen projizieren gerne anthropomorphe Konzepte wie Moral und Gefühle auf andere Entitäten. Bei der Geburt eines Kindes ist die Katze dann schnell einmal ‹eifersüchtig›, und Chatbots sind unmoralisch, weil sie lügen. Ist es falsch, menschliche Konzepte auf Maschinen übertragen zu wollen, und vielleicht auch zu idealistisch, für Maschinen höhere ethische Standards zu definieren als für Menschen?
Dass wir anthropomorphisieren, ist korrekt. Das machen wir bereits bei Autos, denen wir Namen geben und mit denen wir sprechen. Es ist ein natürlicher Zug des Menschen. Aber manchmal projizieren wir nicht nur etwas in Maschinen hinein, sondern wir pflanzen etwas in sie hinein. Das kann eben künstliche Intelligenz sein, künstliche Moral oder künstliches Bewusstsein. Das ist die andere Seite der Anthropomorphisierung: Wir wählen menschliche Eigenschaften aus und bilden sie computerbasiert nach. Wir bauen seit vielen Jahren im Kontext der Maschinenethik sogenannte moralische Maschinen. Diese befolgen bestimmte moralische Regeln, die man ihnen gibt. Auch hier gilt, dass die Moral simuliert wird. Es sind schlicht und ergreifend Maschinen, ohne Bewusstsein, ohne freien Willen. Aber sie können in gewisser Hinsicht mehr als Tiere – sie können unterschiedliche Optionen abwägen und dann die richtige Entscheidung treffen, die sich hoffentlich als eine gute erweist. Das bedeutet nicht, dass die Aktionen der Maschinen gut oder böse im eigentlichen Sinne sind.

Können Sie ein Beispiel einer solchen moralischen Maschine geben?
Wir haben 2016 den Lügenbot gebaut. Er hat sieben unterschiedliche Strategien, mit denen er Unwahrheit erzeugt. Er sucht nach einer Eingabe des Benutzers nach einer richtigen Antwort und manipuliert sie dann. Er macht dies mit einer Zielorientiertheit, die man generell Softwareagenten zubilligt. Ich halte es für unproblematisch, hier vom Lügen zu sprechen. Wenn wir solche Metaphern nicht zulassen und nicht zulassen, dass sie zu Termini technici erstarren, können wir kaum noch über die Dinge reden, vor allem nicht über die neuen. Ein paar Wissenschaftler wollen höhere Standards für Maschinen definieren. Zum Beispiel sollen sich diese, anders als wir, durchgehend und verlässlich an moralische Regeln halten. Die Maschinen sollen uns sogar als Vorbild dienen. Susan L. Anderson und Ronald C. Arkin vertreten solche Ideen. Ich achte und verstehe die Ansätze der beiden Maschinenethiker. Aber ich verfolge eine andere Richtung. Mir geht es darum, individuelle moralische Überzeugungen auf die Maschinen zu übertragen. Das machen wir unter anderem mit Hilfe von Moralmenüs. ▶

Der Begriff der Entfremdung spielt in der europäischen Geistesgeschichte eine wichtige Rolle, sei es die Entfremdung der Natur oder der Arbeit. KI und IoT haben das Potenzial, uns noch mehr von unserer direkten Umgebung, selbst vom eigenen Körper zu entfremden – zum Beispiel wenn die Smartwatch einen erinnert, mal wieder zwei Schlucke Wasser zu trinken, da es schliesslich 28.6 Grad bei 56 Prozent Luftfeuchtigkeit sind. Bereitet Ihnen das Sorgen?
Werkzeuge und Medien können sich zwischen uns und die Realität drängen, die Realität verändern und unsere Vorstellung von der Realität. Wir sind biologische Entitäten und zugleich kulturelle Artefakte. Wenn wir zu Cyborgs werden, wird die Verschmelzung von Natürlichem und Künstlichem noch offensichtlicher. Ab und zu sollten wir uns auf unsere Instinkte verlassen.

Beeinflusst KI unsere Vorstellung von Verantwortung und Schuld, wenn Handlungen von autonomen Maschinen Schaden anrichten?
Autonome Systeme können keine Verantwortung tragen. Sie haben ebenso wenig Schuld. Ich sage seit Jahren, dass wir Maschinen auch kaum etwas verzeihen können und wollen, nicht so, wie wir Menschen etwas verzeihen. Diese Idee hat Richard David Precht im Herbst 2023 auf einem Summit in Düsseldorf aufgegriffen. Die Verantwortung liegt immer beim Menschen. Allerdings ist die Frage, welcher Mensch das ist. Bei einem KI-System kann es Erfinder, Hersteller, Manager und hunderte Entwickler geben, dazu kommen die Betreiber und die Benutzer. Ich denke, dass man in manchen Kontexten einfach keine autonomen Systeme einsetzen sollte. Dann wird es nicht zuletzt mit der Zuweisung von Verantwortung einfacher.

In den USA darf man, glaube ich, ein autonomes Fahrzeug so programmieren, dass es nach rechts ausweicht und einen alten Mann umfährt, und nicht nach links, wo sich gerade ein ganzer Kindergarten zum Picknick niedergelassen hat. Müsste man in Europa anfangen, über solche Dinge zu reden, wenn wir die Entwicklung von autonomen Systemen ernst nehmen wollen?
Diese Diskussion haben wir in Europa zehn Jahre lang ausführlich geführt. Ich habe selbst damit begonnen bei einem Vortrag, den ich Anfang 2013 in Prag gehalten habe. Wir haben 2012 eine Formel für autonome Autos entwickelt, die quantifizieren und qualifizieren kann, und von ihrem Gebrauch abgeraten. Die Ethikkommission des damaligen Bundesverkehrsministers Alexander Dobrindt, 2016 ins Leben gerufen, kam zum selben Ergebnis. Wir bauen moralische Maschinen ja seit vielen Jahren, aber nicht alle autonomen Maschinen sollten solche sein. Zumindest nicht mit Blick auf Menschen. Für Tiere sehe ich Möglichkeiten, die ich seit 2013 ausgearbeitet habe. So könnte es Notbremsungen und Ausweichmanöver nicht nur für grosse, sondern auch für kleine Individuen geben, etwa für Kröten und Igel – wenn kein Verkehrsteilnehmer gefährdet wird.

«Auch bei KI-Systemen liegt die Verantwortung beim Menschen. Allerdings stellt sich die Frage, bei welchem Menschen – dem Erfinder, dem Hersteller, dem Manager, den Hunderten von Entwicklern oder bei den Betreibern und Benutzern?»

Denken Sie, dass KI das Potenzial hat, die Welt besser zu machen? Oder macht sie uns, wie viele Innovationen davor, einfach nur produktiver?
Ich denke, dass man mit neuronalen Netzen, mit Sprachmodellen und speziell mit generativer KI sehr viele Probleme lösen, aber auch verursachen kann. Letztlich werden sie uns einen gewaltigen Nutzen bringen, auf den wir als Homo faber nicht verzichten sollten. Mit Hilfe neuartiger KI-Systeme rückt die universelle Maschine näher. Das ist für mich ein vernetzter Roboter, der im Alltag zahlreiche Probleme bewältigen und lösen kann. Der Roboter Optimus von Elon Musk soll in diese Richtung gehen – aber um den muss sich erst einmal ein begabter Designer kümmern.

Seit der Sesshaftwerdung vor rund 10'000 Jahren hat der Mensch Maschinen entwickelt, die ihm die Arbeit abnehmen. Und jetzt hat man plötzlich Angst, dass wir keine Arbeit mehr haben. Ein Paradox?
Die Arbeit wird idealisiert, das Werkzeug wird problematisiert. Wir richten unser Leben an den Bedürfnissen von Unternehmen aus, nicht an den eigenen. Arbeit muss neu gedacht werden. Lassen wir das Werkzeug die Arbeit machen und widmen uns dem, was uns gefällt. Ich bin mir sicher, dass es plötzlich viel mehr Wissenschaftler und Künstler gibt. Man liegt also nicht auf der faulen Haut.

KI-Entwicklungen kommen in erster Linie aus den USA, Europa hinkt hinterher. Was ist zu tun?
Sie kommen auch aus China, wobei bei generativer KI sicherlich die USA die Nase vorn haben. Wir haben in Europa KI und Robotik jahrelang vernachlässigt. Wir haben Firmen verkauft und Forschung zurückgefahren. Das ändert sich derzeit. Es werden Firmen zurückgekauft und Roboter zurückgeholt. Ein Treiber dabei ist die United Robotics Group. Moderne Serviceroboter sind oftmals mit KI verbunden. Zugleich muss man, wie es Aleph Alpha macht, auf eigene Sprachmodelle setzen. Diese sind wiederum für die Robotik relevant.

Inwiefern?
Das ist derzeit eines meiner Lieblingsthemen: Wie nutzt man Sprachmodelle für Industrie- und Serviceroboter? Damit meine ich nicht nur die natürlichsprachliche Kommunikation zwischen Mensch und Maschine, sondern neue Wahrnehmungs- und Steuerungsmöglichkeiten. Der Roboter bewegt sich durch seine Umgebung und nimmt sie wahr. Die Bilddaten und Audiodaten werden in die Sprachmodelle eingebunden. Wenn man einen Prompt eingibt, indem man dem Roboter in natürlicher Sprache einen Befehl erteilt, wird er abgearbeitet, wie man es von Text- und Bildgeneratoren gewohnt ist. Aber eben nicht im virtuellen Raum, sondern im physischen – ohne vorherige aufwendige Programmierung. Es sind spannende Zeiten für unsere Wissenschaften – für die Robotik und die KI ebenso wie für die Ethik. ●

 © ti&m AG / ti8m.com