

SWR2 Wissen: Aula

Roboterethik – Haben Maschinen Moral?

Gespräch mit Catrin Misselhorn

Sendung vom: Sonntag, 30. Mai 2021, 8.30 Uhr

Erstsendedatum: Sonntag, 22. April 2019, 8.30 Uhr

Redaktion: Ralf Caspary

Produktion: 2019

KI, künstliche Intelligenz, scheint immer intelligenter zu werden, und vielleicht werden eines Tages Maschinen besser und effizienter entscheiden können als wir Menschen. Das wirft die Fragen auf: Wie verändert das unsere Moralvorstellungen, brauchen wir eine Maschinenethik?

Bitte beachten Sie:

Das Manuskript ist ausschließlich zum persönlichen, privaten Gebrauch bestimmt. Jede weitere Vervielfältigung und Verbreitung bedarf der ausdrücklichen Genehmigung des Urhebers bzw. des SWR.

SWR2 können Sie auch im **SWR2 Webradio** unter www.SWR2.de und auf Mobilgeräten in der **SWR2 App** hören – oder als **Podcast** nachhören.

Die SWR2 App für Android und iOS

Hören Sie das SWR2 Programm, wann und wo Sie wollen. Jederzeit live oder zeitversetzt, online oder offline. Alle Sendung stehen mindestens sieben Tage lang zum Nachhören bereit. Nutzen Sie die neuen Funktionen der SWR2 App: abonnieren, offline hören, stöbern, meistgehört, Themenbereiche, Empfehlungen, Entdeckungen ...

Kostenlos herunterladen: www.swr2.de/app

MANUSKRIFT

Anmoderation:

Mit dem Thema: „Roboterethik- haben Maschinen Moral?“. Am Mikrofon: Ralf Caspary.

Ein Ball hüpfte auf die Straße, Sie fahren mit dem Auto auf ihn zu, und Sie werden bremsen, weil Sie aufgrund der Erfahrung wissen, dass nach dem Ball höchstwahrscheinlich ein Kind folgen wird. Was macht in diesem Fall eine künstliche Intelligenz, die den Wagen steuert? Hat man sie so programmiert, dass die erst den Ball als Ball erkennt, dann annimmt, es folgt ein Kind?

Solche Szenarien zeigen, dass es oft um die Frage geht, brauchen wir eine neue Ethik für Roboter oder wie kann man Maschinen Moral beibringen? Welche Konzepte kommen da die Frage. Diese Fragen bearbeitet seit Jahren die Philosophin Professor Catrin Misselhorn, Ihr neues Buch dazu heißt „Maschinenethik“.

Ich habe mit ihr über diese Ethik gesprochen, meine erste Frage war, wie sie das einschätzt: Wenn wir über KI reden, imaginieren wir uns immer eine Superintelligenz, die uns kontrollieren will?

Interview:

Catrin Misselhorn:

Das sind eigentlich ja zwei Ängste: die erste Angst ist, es wird die Superintelligenz kommen, und die zweite Angst ist, sie wird sie wird uns beherrschen und uns in irgendeiner Form versklaven oder unterjochen.

Zunächst muss man sich anschauen, was ist denn die Grundlage solcher Szenarien? Häufig argumentiert man mit dem sogenannten Moore'schen Gesetz aus den 60er-Jahren, es ist eigentlich eine empirische Beobachtung: Man hat kein Naturgesetz, sondern man hat gesehen, okay, so alle ein bis zwei Jahre verbessert sich die Hardware so stark, dass sich die Informations-Verarbeitungsgeschwindigkeit etwa verdoppelt. Zudem wird Rechenkapazität immer billiger. Was hat früher ein Rechner gekostet, was kostet er heute – gar kein Vergleich. Und daraus schloss man dann direkt, dass sich das ihm immer weiter fortsetzen wird.

Das ist natürlich ein viel zu einfacher Gedankengang. Erstens gibt es physikalische Grenzen, zu kleine Bauteile schmelzen dann eben irgendwann; und zweitens kann man von einem solchen exponentiellen Wachstum in der Vergangenheit eben nicht auf die Zukunft schließen. Eigentlich ein ganz klassisches philosophisches Problem: das Induktionsproblem. Man kann nicht einfach aus einer bestimmten Menge von Einzelfällen schließen, dass es immer so weitergehen wird. Beispiel beim Würfeln: Manchmal hat man Glück und würfelt fünf Mal eine sechs. Aber es wäre falsch, daraus zu schließen, beim sechsten Mal würde man auch mit höherer Wahrscheinlichkeit eine sechs würfeln.

Und in diesem Fall: Ein Kollege von mir, Luciano Floridi vom Oxford Internet Institute,

hat es ganz schön veranschaulicht am Beispiel der Truthähne. Man hat beobachtet, dass seit 1929 die Größe der Truthähne um 150 Prozent gewachsen ist. Wenn man das natürlich weiterdenken würde, dann würden sie irgendwann so groß wie Menschen und irgendwann die gesamte Welt überziehen. Also kurz: Das ist natürlich auch ein Fehlschluss und insofern sollte man solchen Szenarien nicht zu einfach Glauben schenken. Ich glaube nicht, dass die Superintelligenz kommen wird. Und dann wäre natürlich die Frage, und wenn sie käme, hätte sie dann tatsächlich ein Interesse uns zu unterjochen.

Man kann natürlich sagen, es gibt Szenarien, auch im Science-Fiction. Beispielsweise übernimmt in „I Robot“ eine künstliche Intelligenz das Kommando und schränkt die Freiheiten der Menschen ein, weil sie der Auffassung ist, die Menschen sind eben irrational. Wenn man sagt, okay, Intelligenz ist maximale Rationalität, könnte man daraus schließen, insofern der Mensch sich irrational verhält, müsste man ihn eben in seinen Freiheiten einschränken. Das wäre aber eigentlich ein moralischer Hintergrund. Es wird da auch der Vergleich gezogen mit Kindern. Jetzt muss man sagen, wir beanspruchen aber natürlich auch, Dummheiten machen zu können, obwohl wir intelligent sind. Und ein solches Szenarium, setzt auch wieder voraus, dass ein solches System sehr ähnlich funktioniert wie ein Mensch oder vielleicht noch eigene Machtgelüste entwickelt – was dann noch auf einem ganz anderen Blatt steht, denn dazu müsste die Intelligenz einen eigenen unabhängigen Willen irgendwie haben, einen Willen zur Macht, müsste man letztlich sagen, und auch das sehe ich in dieser Form nicht gegeben.

Ralf Caspary:

Das Interessante bei diesem Horrorszenario, dass die Intelligenz uns irgendwann unterbuttern will, das basiert ja aus meiner Perspektive auf einem Menschenbild, das so sehr eingeschränkt und reduziert ist, da sagt man ja, die menschliche Intelligenz ist weder perfekt noch besonders gut, sie kann jederzeit übertrumpft werden. Ist das nicht so ein Selbstbild, das da drinsteckt, was eher negativ ist?

Catrin Misselhorn:

Ja, also das kann man schon sagen. Aber zunächst einmal auch vielleicht eine Quantifizierung von Intelligenz, die es so ohne weiteres gar nicht gibt. Also ich meine, was heißt, etwas ist intelligenter oder superintelligent? Kann die Intelligenz besser rechnen oder was genau ist der Hintergrund? Hat sie mehr Daten? All solche Sachen können natürlich – zumindest beim Menschen – auch nach hinten losgehen. Wer zu viel auswendig lernt, der erkennt die Zusammenhänge nicht, also diese Quantifikation von Intelligenz finde ich an sich schon problematisch. Dann steckt sicherlich auch ein Stück weit ein auf die Aufklärung zurückgehendes Bild von Rationalität darin, was einfach darin besteht, das Eigeninteresse immer zu maximieren in irgendeiner Form. Und auch das muss man natürlich so nicht übernehmen.

Ralf Caspary:

Was ist für Sie Intelligenz?

Catrin Misselhorn:

Beim Menschen ist nicht recht klar: Gibt es die eine Intelligenz? An sich nicht, es werden verschiedene Arten unterschieden, beispielsweise mathematische Intelligenz, aber auch räumliche Intelligenz, sogar emotionale Intelligenz. Intelligenz

hat einfach sehr viele Facetten. Früher sagte man, künstliche Intelligenz liegt dann vor, wenn ein künstliches System Aufgaben lösen kann, die bei einem Menschen Intelligenz erfordern. Das klingt ja für den Anfang nicht schlecht, ganz so einfach ist es auch wieder nicht. Nehmen wir das Beispiel Taschenrechner. Rechnen erfordert beim Menschen natürlich Intelligenz. Die Tatsache, dass ein Taschenrechner vieles schneller und akkurater berechnen kann als ein Mensch, würde man nicht als übermäßiges Zeichen von Intelligenz werten. Insofern, denke ich, ist das auch wieder zu hoch gegriffen, aber es geht in irgendeiner Form um Problemlösungsfähigkeiten.

Was unterscheidet dann eigentlich künstliche Intelligenz von menschlicher Intelligenz? In der Frühphase war es das Ziel der KI, eine bestimmte Intelligenz zu schaffen: eine allgemeine Problemlösungsfähigkeit, also ein General Problem Solver. Und das ist eben ein Kennzeichen: KI, so wie wir sie jetzt kennen, ist eigentlich immer auf bestimmte Spezialgebiete beschränkt, kann eine Aufgabe sehr gut erledigen, besser auch als Menschen. Aber eben in anderen Feldern gar nicht. Wir Menschen können unsere Intelligenz prinzipiell auf alle Gebiete anwenden. Dazu gehört sicherlich auch die Lernfähigkeit, also dass Menschen lernen können, und auch die Fähigkeit, sich auf neue Situationen einstellen zu können.

Lernfähigkeit und künstliche Intelligenz werden manchmal auch gleichgesetzt mit maschinellern Lernen. Das haben Maschinen auch, allerdings in einer anderen Form, als sie beim Menschen derzeit vorliegt, nämlich es geht vor allem um Mustererkennung aufgrund großer Datenmengen, und Menschen lernen so nicht. Das habe ich an unserer kleinen Tochter etwa gesehen. Sie lernt nicht aus Billionen von Sätzen die natürliche Sprache, sondern der Input ist, das hat der Linguist Noam Chomsky beobachtet: Der Input, den Kinder beim Spracherwerb haben, ist spärlich. Also haben sie irgendwie die Fähigkeit, zusammen mit angeborenen Fähigkeiten dann aufgrund von recht geringem Input Sprache zu lernen. Das funktioniert beim Menschen nicht einfach durch Mustererkennung

Ralf Caspary:

Ich verstehe: Die Maschinen brauchen sehr großen Input, die brauchen sehr große Daten, und das menschliche Gehirn bildet anhand weniger Daten schon Regeln, auch beim Spracherwerb ist das so. Lassen Sie uns das konkretisieren: Wir besprachen neulich in unserer Wissenschaftsredaktion den Fall, dass eine künstliche Intelligenz, ein Algorithmus Millionen Röntgenbilder auswerten und danach relativ treffsicher sagen kann, auf diesem Röntgenbild der Lunge ist ein Krebs zu sehen. Ist das intelligent?

Catrin Misselhorn:

Wenn wir die erste Definition zugrunde legen, die ich angegeben habe, dann ja. Die klassische Intelligenz bezieht sich auf die Lösung von Problemen, für die der Mensch Intelligenz braucht. Das trifft hier zu. Ich habe jetzt kein Problem damit, in dem Fall von künstlicher Intelligenz zu sprechen, würde aber sagen, man muss sich im Klaren darüber sein, was der damit verbundene Anspruch ist. Und da gibt es die Unterscheidung zwischen „starker“ und „schwacher“ KI, also schwache KI hat wirklich nur diesen Anspruch, Problemlösungsfähigkeiten zu schaffen, aber nicht den Anspruch, menschliche Intelligenz selber nachzubilden. Das wäre der Anspruch der starken KI. Deswegen würde ich sagen, im schwachen Sinn ist es kein Problem, von künstlicher Intelligenz zu spielen.

Ralf Caspary:

Dieses Programm, das den Go-Spieler besiegt hat, ist das intelligent?

Catrin Misselhorn:

Ja, das waren ja ursprünglich zwei Programme. Daran kann man auch noch mal einen Fortschritt erläutern: Das erste Programm war AlphaGo, und das basierte auf der Analyse einer sehr großen Menge menschlicher Spiele, also Mustererkennung in dem Sinn, wie ich es eben erklärt habe. Und in dem Sinne, dass das System ein Verhalten gezeigt hat, was einigen der besten menschlichen Spieler überlegen war, würde ich sagen, ist das intelligent. Die nächste Stufe war AlphaGo Zero, und das ist deswegen interessant, weil AlphaGo Zero nicht mehr menschliche Spiele als Daten benutzt hat, sondern nur mit den Regeln ausgestattet war und gegen sich selbst gespielt hat. Es hat durch Spiele gegen sich selbst diese Daten generiert, die es brauchte, und dann ähnlich oder noch besser als AlphaGo gegen menschliche Spieler performt. Also ja, ich würde sagen in diesem ganz kleinen Ausschnitt Go-Spiel ist das ein intelligentes System.

Ralf Caspary:

Ich habe in der Süddeutschen Zeitung gelesen, es gibt einen Algorithmus, der Lieder zum Einschlafen komponiert. Das Interessante an diesem Artikel war: Er war überschrieben mit: Ein Algorithmus bekommt jetzt Rechte für einen Plattenvertrag oder so ähnlich. Da fand ich interessant, dass unsere Sprache uns schon verführt, diesen Algorithmus als Person wahrzunehmen.

Catrin Misselhorn:

Ich sehe die Gefahr weniger in der Sprache als mehr in der emotionalen Interaktion mit Maschinen, weil ich tatsächlich denke, wir neigen dazu, Maschinen zu anthropomorphisieren, wie übrigens auch andere Gegenstände, also nicht nur Maschinen, auch unbelebte Gegenstände, und zwar wenn sie gewisse menschen- oder auch tierähnliche Züge aufweisen. Das kann visuell sein, eine Ähnlichkeit in der Erscheinung, das kann aber auch stimmlich sein. Das kann auch ganz abstrakt sein. Eines der ersten Programme, das mit Nutzern interagiert hat, war das Psychotherapie-Programm „Eliza“ aus den 60er-Jahren. Die Menschen fühlten sich wunderbar verstanden von diesem Programm. Das war sehr einfach gestrickt und arbeitete mit einfachen Mustern. Es kam z.B. die Frage „Wie geht es dir“, dann antwortete man vielleicht „Gut, ich denke oft an meinen Vater“, Frage „Warum denkst du oft an deinen Vater“ usw. Und so kam ein Dialog zustande, basierend auf sehr einfachen Schlüsselbegriffen. Das Prinzip, warum sich viele Menschen verstanden fühlten, war natürlich Anthropomorphisierung. Also da sehe ich eher die Tendenz, dass wir allzu schnell bereit sind, eine Maschine zu vermenschlichen, aber mehr aufgrund dieser Emotionalität, weniger aufgrund der rein sprachlichen Ebene, würde ich sagen.

Ralf Caspary:

Wäre das ein wichtiger ethischer Imperativ, diese Dinge nicht zu vermenschlichen oder wenigstens eine klare Grenze zu ziehen zwischen Mensch und Maschine? Denn für mich fängt Anthropomorphisierung beim Navi an. Wenn es mir z.B. sagt, an der nächsten Kreuzung rechts fahren. Man denkt ja, da sitzt eine kleine Person drin, wie man früher dachte, im Gehirn sitzt wahrscheinlich auch so ein kleines Männlein, das alles steuert.

Catrin Misselhorn:

Es gibt Bereiche, wo ich damit weniger Probleme habe. In den 90er-Jahren gab es z.B. den damals recht bekannten „Hund“ Aibo. Wenn man sich die alten Foren, ich habe sie mir mal eine Zeit lang angeschaut, durchliest, dann sieht man, die Leute haben den wie ein Haustier gehalten und eine emotionale Bindung dazu aufgebaut. Ich finde das nicht per se schlimm. Manch einer baut zu seinem Auto eine emotionale Bindung auf – warum nicht? Im gewissen Rahmen ist das ist eine normale Tendenz, glaube ich, und auch nicht gefährlich. Gefährlich wird es aus meiner Sicht da, wo man sowas einsetzt, um Personen zu manipulieren, oder auch wo die Auswirkungen in bestimmten Bereichen einfach schwierig sind.

Ralf Caspary:

In welchen Bereichen wäre das?

Catrin Misselhorn:

Nehmen wir z.B. Call Center, da gibt es ja schon die Möglichkeit, KI einzusetzen. Und da ist es wichtig zu wissen, spricht man jetzt mit einem Menschen oder spricht man mit einer KI? Und es wäre schon eine problematische Form der Manipulation, wenn man es da mit einer netten, verständnisvollen Stimme zu tun hat, die scheinbar auf alle Besorgnisse eingeht, aber in Wirklichkeit handelt es sich um eine KI mit menschenähnlicher Stimme. Das wäre z.B. ein Kontext, wo ich denke, man sollte informiert werden.

Ralf Caspary:

Sie haben sich mit Maschinenethik beschäftigt. Was meinen Sie: Brauchen wir eine Veränderung unserer bisherigen Moral im Umgang mit künstlicher Intelligenz, unserer ethischen Standards? Oder kann alles so bleiben und wir müssen nur nachbessern?

Catrin Misselhorn:

Wo würden Sie die Ansatzpunkte für eine Veränderung sehen?

Ralf Caspary:

Zum Beispiel wird immer wieder behauptet, es wird irgendwann künstliche Intelligenz geben, die besser entscheiden kann als wir Menschen. Dann wäre für mich ein ethisches Problem, was mit meiner Entscheidungskompetenz passiert?

Catrin Misselhorn:

Ich fange ein bisschen grundlegender an. Erster Punkt wäre, um welche Art von Beziehungen geht es? Geht es darum, die Maschine trifft etwa im Bereich des autonomen Fahrens Entscheidungen über menschliches Leben? Brauchen wir dafür neue Standards? Und da würde ich sagen, nein, in diesem Fall haben wir die bereits zur Verfügung stehende Theorien. Es gibt da einen Konflikt zwischen dem Utilitarismus und der kantischen Ethik. Ein viel diskutiertes Beispiel geht zurück auf das Trolley-Problem von Philippa Foot, auch schon in den 60er-Jahren. Folgende Situation: Stellen Sie sich vor, eine Straßenbahn rast auf fünf Menschen zu, die sie töten wird, wenn Sie – Sie stehen an einer Weiche – nicht die Weiche umlenken und diese Straßenbahn auf ein Gleis leiten, wo nur eine Person steht. Klare Antwort des Utilitarismus: ja, sollte man umleiten. Warum? Für den Utilitarismus besteht das Moralprinzip darin, die Bilanz von Schmerz und Verlust zu optimieren. Wenn also fünf Menschen am Leben bleiben, ist das eine bessere Bilanz, als wenn nur einer

überlebt. Die kantische Ethik sieht das anders. Für die kantische Ethik kann man Leben in dieser Form nicht gegeneinander verrechnen, und es wäre in dem Fall die Gefahr gegeben, dass die eine Person zum Mittel zum Zweck gemacht wird, um fünf zu retten. Und das ist für Kant eine sehr problematische Entscheidung. Die Moraltheorien, die hier aufeinanderprallen, sind einfach die klassischen Moraltheorien.

Ein anderer Ansatz betrifft eher die Frage, wie verhält sich die Maschine gegenüber dem Menschen. Und da hatte Isaac Asimov die Idee, das muss ganz anders sein, als wenn es um Personen geht, die alle gleiche und freie moralische Amateure sind. Asimov schlug eher sowas vor – manche sagen, eine Sklavenmoral – drei Gesetze, von denen das erste besagt, ein Roboter soll keinen Menschen töten oder zu Schaden kommen lassen; das zweite: ein Roboter soll einem menschlichen Befehl gehorchen, es sei denn, er steht im Konflikt mit dem ersten Gesetz; und das dritte Gesetz besagt, ein Roboter soll sich selbst schützen, es sei denn, das steht im Konflikt mit dem ersten oder zweiten Gesetz. Dass das nicht so einfach ist, zeigt Asimov selber. Der sah sich ja nicht so sehr als Messias einer neuen Ethik, sondern er hat in seinen Kurzgeschichten gezeigt, wie schwierig es ist, diese Gesetze umzusetzen.

Eines der Probleme ist etwa, dass es eben in so Dilemma-Situationen dazu kommt, dass ein System keine Entscheidung treffen kann, und entweder in einen unendlichen Kreislauf gerät oder einfach paralysiert wird, also zusammenbricht und nichts mehr tun kann. Aus solchen Situationen ergibt sich eigentlich immer der Stoff für seine Geschichten, aus denen man nur schließen kann, so einfach ist es nicht, diese Gesetze, die zunächst ja mal recht einfach und plausibel klingen, umzusetzen.

Dritter Punkt: Brauchen wir eine neue Moral für den Umgang mit Maschinen? Das ist immer wieder Thema von Science-Fiction. Was tun wir eigentlich, wenn wir an den Punkt kommen, wo Maschinen etwa den Turing-Test bestehen und uns tatsächlich als menschliche Akteure erscheinen? Müssen wir sie dann auch als solche anerkennen oder müssen wir das nicht. Auch da kann man sagen, die klassischen ethischen Theorien haben schon Standards gesetzt, zum Beispiel müssen wir sie als empfindungsfähig auffassen – das wäre für den Utilitarismus eine relevante Frage; oder können Sie Interessen entwickeln; oder auch für die kantische Ethik: sind es also autonome Wesen, autonome rationale Wesen, die selber auch vernünftig handeln können, sich selber Ziele setzen?

Ralf Caspary:

Ich komme noch einmal ganz kurz zurück zum autonomen Fahren: Utilitarismus und die kantische Ethik, haben Sie gesagt, und Sie haben auch gesagt, mit diesen beiden Polen kommt man durch, auch in der Auseinandersetzung mit der künstlichen Intelligenz.

Catrin Misselhorn:

Was heißt, man kommt damit durch? Es gibt natürlich noch Alternativen dazu, das kann man nicht verschweigen. Eine der Alternativen ist eine auf Aristoteles zurückgehende Theorie, die besagt, wir gehen überhaupt nicht von Moralprinzipien aus, sondern wir müssen die Einzelsituation betrachten. Und auf dieser Grundlage können wir dann – vielleicht für Aristoteles auch noch mit Sozialisation vermittelt – allgemeine Daumenregeln beschließen. Aber der Ausgangspunkt hier, deswegen

Partikularismus, immer die einzelne Situation. Und das ist ein Kontrast zu diesen anderen beiden klassischen Ansätzen, die sagen, Moralprinzipien sind der Ausgangspunkt.

Ralf Caspary:

Kann man utilitaristische autonome Fahrzeuge entwickeln, die nach diesem Gesetz, das Sie beschrieben haben, handeln?

Catrin Misselhorn:

Also grundsätzlich, denke ich, sind wir da wieder beim Problem von Asimov. Die Umsetzung solche Gesetze ist sehr viel komplexer, allein schon auf der Basis, wie soll ein solches System entscheiden? Wenn man zum Beispiel nur die Anzahl zugrunde legt, dann wäre das wohl möglich, denke ich. Ein System müsste fünf Personen auf der einen Seite erkennen, eine Person auf der anderen, und das scheint nicht von, sagen wir mal, übermäßig mäßiger Komplexität zu sein, wobei klar ist, gerade das Computer-Sehen ist eine schwierige Aufgabe. Da gibt es so Beispiele: Ein System soll Katzenbilder von Bildern von Karamelleis unterscheiden, und das ist erstaunlich schwierig. In dem Bereich lauern zunächst die Schwierigkeiten. Dann natürlich die Schwierigkeiten, wie genau stellt man diese Lust-Leid-Bilanz auf. Soll das Alter da noch eine Rolle spielen? Soll der Familienstand oder die Bedeutung für die Gesellschaft eine Rolle spielen? Und natürlich auch die Frage, wie findet man das heraus? Das müsste irgendwie auf einem Chip gespeichert sein, den ein solches System dann erkennen kann. Also ich denke, die Umsetzung ist das Problem, und je mehr Parameter man mit einbezieht, desto schwieriger wird es.

Ralf Caspary:

Wir reden meistens über das autonome Fahren, weil da sozusagen die Zukunft vor der Tür steht, auch was die Entwicklung betrifft. Aber wenn wir zum Beispiel den Pflegebereich nehmen, also Pflege von alten Menschen, da denken wir auch über den Einsatz von Robotern nach. Das ist alles noch nicht so recht greifbar, da wird es ja ankommen auf das Verhältnis von Mensch und Maschine, auch noch mal mit der Frage im Hintergrund, wie sich Menschen von Maschinen unterscheiden. Wenn zum Beispiel alte Menschen im Altenheim durch Roboter angesprochen werden, aus dem Bett geholt werden, gefüttert werden etc. Da ergeben sich ja diese Fragen, die Sie eben angedeutet haben.

Catrin Misselhorn:

Genau, das ergibt sich auch nochmal. Wobei ich den aussichtsreichsten Einsatzbereich für künstliche Systeme im Pflegebereich tatsächlich in der häuslichen Pflege sehe. Ich glaube, es ist die Horrorvorstellung, man wäre in einem Heim, wo nur noch Roboter sich um einen kümmern. Sehr viel positiver ist, glaube ich, die Vorstellung, dass man länger bei sich zuhause leben kann und von einem künstlichen System dort in der häuslichen Umgebung gepflegt wird. Und da sehe ich tatsächlich auch Möglichkeiten. Ein solches System sollte die Eigenschaften haben, sich auf die individuellen Moralvorstellungen des Nutzers einzustellen. Bisher geht es um Systeme, die an Essen, Trinken und Medikamenten-Aufnahme erinnern, es geht um Systeme, die z.B. einen Arzt rufen, wenn sich jemand eine Zeitlang nicht bewegt oder vielleicht die Angehörigen verständigt. Und es geht um die Frage, soll ein System den Nutzer überhaupt rund um die Uhr überwachen. Und wie verhält sich das zu seiner Privatsphäre insgesamt? Kommt es zu Konflikten zwischen

verschiedenen Werten, etwa der Selbstbestimmung, Privatsphäre oder Sorge der Angehörigen. Die Frage ist, wie sollen die gewichtet werden? Das Wichtige für mich ist, dass man das, glaube ich, nicht einfach von vorneherein entscheiden kann, auch nicht wie als Philosophen, sondern dass Menschen, die betroffen sind, das selber entscheiden sollten. Ist es ihnen etwa wichtiger, gesundheitliche Risiken um jeden Preis auszuschalten, auch wenn das eine Einschränkung ihrer Privatsphäre bedeutet. Mancher würde das vielleicht befürworten, während es umgekehrt sicher auch Menschen gibt, die sagen, nein das möchte ich nicht, ich möchte meine Privatsphäre nicht einschränken lassen, auch wenn ich dadurch vielleicht gewisse gesundheitliche Risiken in Kauf nehmen muss. Es wäre ein großer Gewinn, wenn ein solches System sich an den Moralvorstellungen des Nutzers ausrichten würde und die ebenso gewichten würde wie der Nutzer selbst. Manchmal spreche ich auch von einem Avatar des Nutzers.

Jetzt ist natürlich die Frage, ist das nun eine gute Vorstellung oder nicht? Ich habe mal einen Cartoon gelesen, da war eine alte Dame mit einem Pflegeroboter abgebildet mit der Sprechblase: „Ja, ich kann jetzt in meinen eigenen vier Wänden leben, aber niemand besucht mich.“ Und das ist natürlich ein Problem, was da angesprochen wird. Aber dieses Problem ist, glaube ich, einfach ein gesamtgesellschaftliches Problem, eben die Frage, wie wir mit alten Menschen umgehen. Ein Großteil von uns kann sich nicht um alte Angehörige kümmern wegen eigener Berufstätigkeit oder Familie. Diese Problematik spiegelt sich in dem Cartoon wider. Aber ich glaube, sie wird nicht nur durch die Roboter geschaffen, das heißt, dass die Roboter nicht dieses Problem lösen können, zum Beispiel adäquate soziale Kontakte herzustellen. Und da kann ich auch nochmal den Kreis schließen zu Ihrer vorherigen Frage, wo die Anthropomorphisierung problematisch ist. Es wäre schon traurig, wenn etwa eine Person nur noch Beziehung zu ihrem Roboter aufbauen kann. Aber auch da ist es nicht ganz einfach, weil es vielleicht Leute gibt, die ganz zufrieden sind mit der Situation. Aber es sollte nicht per se dazu führen, dass man sagt, okay, die Pflege ist gewährleistet, damit ist Kittel geflickt. Sondern

Ralf Caspary:

Dahinter steht ja die Voraussetzung, dass Maschinen immer menschenähnlicher werden, dass sie überhaupt Empfindungen zeigen könnten. Da wäre schon zu fragen, ob eine Maschine kein Ersatz werden könnte für eine Beziehung?

Catrin Misselhorn:

Das könnte sie, glaube ich, schon, wir haben jetzt schon die Situation, dass beispielsweise Puppen ein Beziehungersatz werden können. In einem gewissen Rahmen würde man vielleicht sagen, warum nicht. Für mich ist der entscheidende Punkt, die Leute sollten einfach selbst die Wahl haben, in welcher Form sie leben wollen. Wenn tatsächlich jemand so eine Beziehung aufbaut, würde ich sagen, muss man es ihm vielleicht nicht ausreden in dieser Form, aber es sollte niemand darauf reduziert werden, eben wie die alte Dame, die sagt, sie kann zuhause leben, aber niemand kommt mehr vorbei.

Ralf Caspary:

Eine Frage zum Schluss: Stellen wir im Moment die richtigen ethischen Fragen und sind wir auf die wichtigen ethischen Probleme vorbereitet in Bezug auf künstliche Intelligenz?

Catrin Misselhorn:

Ich würde das nicht so final sehen. Ich glaube, wir stellen die richtigen Fragen, was aber natürlich nicht immer heißt, dass wir die Antworten schon komplett haben. Und diese Aussage, wir stellen die richtigen Fragen, da lege ich auch die Betonung ganz klar auf das „wir“. Als ich mich im Jahr 2007 mit der Thematik Maschinenethik begonnen habe zu beschäftigen, hat das noch kaum jemanden interessiert. Aber in den letzten fünf, sechs Jahren ist das ganz stark ins Zentrum der Öffentlichkeit gerückt. Das ist, denke ich, ein Signal in die Richtung, dass wir die richtigen Fragen stellen. Aber wir müssen sie eben nicht nur als Fach-Philosophen stellen, sondern als Öffentlichkeit. Das ist mir ein ganz großes Anliegen, dass das nicht nur im Bereich der Experten bleibt, sondern es geht um das Zusammenleben von uns allen. In Tübingen gibt es etwa eine Bürgerinitiative, die sich mit diesen Fragen beschäftigt, angeregt durch die Gründung des Cyber-Valley. Ich halte solche Initiativen für wichtig, um diesen Dialog in einer breiteren Gesellschaft zu führen und nicht nur unter Experten.

Ralf Caspary:

Dazu trägt hoffentlich auch diese Sendung bei. Ich bedanke mich ganz herzlich.

Catrin Misselhorn:

Ich bedanke mich auch.
