

Das Gewissen der Künstlichen Intelligenz

Dieser Artikel gibt den Stand zum Zeitpunkt seiner Veröffentlichung wieder. Der Inhalt kann inzwischen veraltet sein.

07/02/2020 Professor Christoph Lütge leitet eines der ersten Forschungsinstitute der Welt für Ethik in der Künstlichen Intelligenz. An der TU München sollen sich interdisziplinäre Teams mit den Folgen der Entscheidungen von Software beschäftigen. Ein zentrales Thema wird das autonome Fahren sein.

Was wäre, wenn ich jetzt einfach nichts mehr täte? Diese Frage schoss Professor Christoph Lütge durch den Kopf, als er 2016 zum ersten Mal in einem hochautomatisierten Auto auf der A9 von München nach Ingolstadt fuhr. Die Teststrecke war ein Jahr zuvor eröffnet worden, und auf ihr dürfen Fahrzeuge verkehren, die selbstständig beschleunigen, bremsen und lenken können. Wenn ein Warnsignal ertönt, hat der Fahrer zehn Sekunden Zeit, wieder die Kontrolle zu übernehmen. Und wenn er das nicht tut? Welche Kriterien nutzt der Bordcomputer für die Entscheidung, wie es dann weitergehen soll? Welche Prioritäten setzt er? Diese Fragen ließen Lütge nicht mehr los. Er war auf ein neues, brandaktuelles Forschungsfeld gestoßen. Professor Christoph Lütge leitet eines der ersten Forschungsinstitute der

Welt für Ethik in der Künstlichen Intelligenz. An der TU München sollen sich interdisziplinäre Teams mit den Folgen der Entscheidungen von Software beschäftigen.

Seit neun Jahren ist der 49-jährige Professor für Wirtschaftsethik an der Technischen Universität München und erforscht, wie Wettbewerb die soziale und ethische Verantwortung von Unternehmen fördert. Mit Künstlicher Intelligenz (KI) hatte er sich vor seiner Testfahrt auf der A9 nur beiläufig beschäftigt. Danach las er Studien, recherchierte, sprach mit Herstellern. Dabei wurde ihm schnell klar, dass KI eine ganze Reihe ethischer Fragen aufwirft: Wer haftet, wenn etwas schiefgeht? Wie nachvollziehbar sind die Entscheidungen der intelligenten Systeme? Auch die Transparenz der KI-Algorithmen ist derzeit nicht ausreichend: Noch kann man oft nicht nachvollziehen, aufgrund welcher Kriterien sie ihre Entscheidungen treffen – die KI wird zu einer Blackbox. „Diesen Herausforderungen müssen wir uns stellen. Egal ob KI zur Diagnose von medizinischen Befunden, zur Verbrechensbekämpfung oder zur Steuerung von Autos einsetzt wird“, so Lütge. „Mit anderen Worten: Wir müssen uns mit den ethischen Fragen rund um die Künstliche Intelligenz beschäftigen.“

Denn die Vorstellung, dass unser Leben in Zukunft von Maschinen bestimmt wird, die nur Logik, aber keine Ethik kennen, erzeugt bei vielen Menschen ein ungutes Gefühl. In einer Umfrage des World Economic Forum (WEF) in 27 Ländern gaben 41 Prozent von insgesamt 20.000 Befragten an, sie seien beunruhigt über den Einsatz von KI. 48 Prozent wünschen sich eine stärkere Regulierung von Unternehmen, 40 Prozent mehr Restriktionen beim Einsatz von Künstlicher Intelligenz durch Regierungen und Behörden.

Ethik des autonomen Fahrens

Ein besonders aktuelles und schwieriges Feld ist das autonome Fahren, weil es hier sehr schnell um Menschenleben geht. Was soll der KI-Algorithmus beispielsweise tun, wenn die Bremsen versagen und das voll besetzte Fahrzeug entweder mit einer Betonbarriere kollidieren oder in eine Gruppe Fußgänger fahren kann? Welche Prioritäten soll die KI in diesem Fall setzen? Soll sie das Leben der Insassen höher bewerten als das der Passanten? Sollen Opfer unter Kindern eher vermieden werden als unter alten Menschen? „Das sind typische Dilemmata, die von Sozialwissenschaftlern erforscht werden“, so Lütge. Sie spielten auch in der vom Bundesministerium für Verkehr und digitale Infrastruktur eingesetzten Ethik-Kommission zum automatisierten Fahren, der Lütge angehört hat, eine wichtige Rolle. „Wir waren uns dort einig, dass es eine Diskriminierung aufgrund von Alter, Geschlecht oder anderen Kriterien nicht geben darf. Denn sie ist nicht vereinbar mit dem Grundgesetz.“ Erlaubt sei jedoch eine Programmierung, die die Zahl der Personenschäden minimiert.

Auch die Haftungsfrage stellt sich völlig neu. Wenn autonom gefahren wurde, ist der Hersteller in der Pflicht – denn dann greift die Produkthaftung. Ansonsten ist der Fahrer haftbar. „Allerdings brauchen wir in Zukunft eine Art Flugschreiber im Auto“, so Lütge. „Er zeigt an, ob zum Zeitpunkt des Crashs die autonomen Fahrfunktionen eingeschaltet waren. Daraus ergeben sich natürlich wieder Fragen nach dem Datenschutz.“ Trotz aller Herausforderungen ist der Wissenschaftler aber überzeugt, dass autonome Fahrzeuge den Verkehr sicherer machen. „Sie werden besser als Menschen sein – denn

abgesehen davon, dass sie nicht müde oder unkonzentriert werden, nehmen ihre Sensoren auch mehr von der Umgebung wahr. Außerdem können sie angemessener reagieren: Autonome Fahrzeuge bremsen härter und weichen gekonnter aus. Selbst bei normalen Abläufen im Straßenverkehr werden sie darum den Menschen übertreffen."

Viel KI-Kompetenz am Standort München

Es gibt also viele spannende Fragen, mit denen sich Lütge an seinem neuen Institut für Ethik in der Künstlichen Intelligenz an der TU München beschäftigen kann. Finanziert ist es mit 6,5 Millionen Euro von Facebook. Man sei an keine Vorgaben gebunden. Aber natürlich habe Facebook Interesse an den wissenschaftlichen Ergebnissen – immerhin ist es eines der ersten Forschungsinstitute in diesem Bereich. München bot sich als Standort an, weil die TUM für ihre Kompetenz in der KI bekannt ist. Außerdem wird der Datenschutz in Deutschland besonders ernst genommen, und die Bevölkerung steht technischen Entwicklungen generell eher kritisch gegenüber. Im neuen Forschungsinstitut sollen auch diese Skeptiker Gehör finden: „Wir wollen alle wichtigen Player zusammenbringen, um gemeinsam ethische Leitlinien für konkrete KI-Anwendungen zu erarbeiten. Die Voraussetzung dafür ist, dass Vertreter von Wirtschaft, Politik und Zivilgesellschaft miteinander ins Gespräch kommen", so Lütge.

Für die Forschung zur Ethik in der KI will der Wissenschaftler interdisziplinäre Teams bilden, die untersuchen sollen, welche ethische Relevanz die neuen Algorithmen haben: „Techniker können alles programmieren", sagt Lütge. „Aber wenn es darum geht, die Folgen der Entscheidungen von Software abzuschätzen, braucht man den Input von Sozialwissenschaftlern." Darum will er künftig Tandems aus je einem Mitarbeiter aus den Technikwissenschaften und einem Vertreter der Geistes-, Rechts- oder Sozialwissenschaften bilden. Außerdem plant Lütge Projektteams, deren Mitglieder von verschiedenen Lehrstühlen oder Abteilungen kommen sollen. Sie werden konkrete Anwendungsfälle untersuchen, etwa den Einsatz von Pflegerobotern und die ethischen Fragen, die sich in diesem Zusammenhang stellen.

Dass KI in viele Lebensbereiche Einzug halten wird, hält Lütge schon jetzt für ausgemacht – denn sie biete einen enormen Mehrwert, etwa im Verkehr: „In einigen Jahren werden autonom fahrende Autos mit verschiedenem Automatisierungsgrad zum Straßenbild gehören", prognostiziert der Forscher. Nach den Regeln der Ethik-Kommission sei die Voraussetzung dafür gegeben, wenn die autonomen Fahrzeuge mindestens genauso gut sind wie ein menschlicher Fahrer – beispielsweise im Einschätzen der Verkehrslage und in ihren Reaktionen. Er persönlich freut sich darauf: „Wenn ich in ein autonomes Fahrzeug steige, spüre ich in den ersten Minuten immer eine gewisse Unsicherheit – bis klar ist, dass das Auto zuverlässig beschleunigt, bremst und lenkt. Dann kann ich die Verantwortung sehr schnell abgeben. Diesen Zustand genieße ich, denn dann habe ich Zeit zum Nachdenken." Beispielsweise darüber, was wäre, wenn autonome Fahrzeuge über Grenzen fahren: Würden dort dieselben ethischen Entscheidungsalgorithmen gelten? Oder bräuchte man an jeder Grenze ein Update? Solche Fragen werden Lütge noch lange beschäftigen.

Leitlinien zu Ethik und KI

EU-Expertengruppe: Ethik-Leitlinien für eine vertrauenswürdige KI (April 2019)

Eine Expertengruppe für Künstliche Intelligenz hat mit den Leitlinien das Ziel verfolgt, einen Rahmen für die Verwirklichung einer vertrauenswürdigen KI zu setzen. Es soll sich dabei um eine Hilfestellung für die mögliche Umsetzung von Prinzipien in soziotechnischen Systemen handeln. Der Rahmenplan geht dabei auf Sorgen und Ängste von Bürgerinnen und Bürgern ein und soll eine Basis für die Förderung der EU-weiten Wettbewerbsfähigkeit bieten.

Ethische Grundsätze/Prinzipien im KI-Kontext

1. Achtung der menschlichen Autonomie
KI-Systeme sollen Menschen nicht auf ungerechtfertigte Weise unterordnen, nötigen oder täuschen. Sie sollen zur Schaffung sinnvoller Arbeit dienen.

2. Schadenverhütung

KI-Systeme sollen keinen Schaden verursachen. Dazu gehört der Schutz der geistigen und körperlichen Unversehrtheit. Technische Robustheit soll die Anfälligkeit für Missbrauch dezimieren.

3. Fairness

KI-Systeme sollen Chancengleichheit beim Zugang zu Bildung und Gütern fördern. Ihr Einsatz sollte nicht dazu führen, dass Nutzer in ihrer Wahlfreiheit beeinträchtigt werden.

4. Erklärbarkeit

Prozesse und Entscheidungen müssen transparent und verständlich bleiben. Dauerhaftes Vertrauen in die KI kann nur durch eine offene Kommunikation ihrer Fähigkeiten und ihrer Einsatzzwecke entstehen.

Ethik-Kommission: Automatisiertes und vernetztes Fahren (Bericht vom Juni 2017)

Die vom deutschen Bundesminister für Verkehr und digitale Infrastruktur eingesetzte interdisziplinäre Ethik-Kommission hat Leitlinien für das automatisierte und vernetzte Fahren entwickelt.

Automatisierter und vernetzter Fahrzeugverkehr: Ein Auszug aus den Regeln

Regel 2

Die Zulassung von automatisierten Systemen ist nur vertretbar, wenn sie im Vergleich zu menschlichen Fahrleistungen zumindest eine Verminderung von Schäden im Sinne einer positiven Risikobilanz versprechen.

Regel 7

Für Gefahrensituationen wird die KI technisch so programmiert, dass im Konfliktfall Tier- oder Sachschäden in Kauf genommen werden, wenn Personenschäden dadurch vermieden werden können.

Regel 9

Jede Qualifizierung von Opfern nach persönlichen Merkmalen (Alter, Geschlecht, körperliche oder geistige Konstitution) ist bei unausweichlichen Unfallsituationen untersagt.

Regel 10

Die dem Menschen vorbehaltene Verantwortung verschiebt sich bei automatisierten und vernetzten Fahrsystemen vom Autofahrer auf die Hersteller und Betreiber der technischen Systeme sowie auf die infrastrukturellen, politischen und rechtlichen Entscheidungsinstanzen.

Über die Person

Prof. Christoph Lütge ist Experte für Wirtschafts- und Unternehmensethik. Seit 2010 hat er den Peter Löscher-Stiftungslehrstuhl für Wirtschaftsethik an der TU München inne. Lütge ist auch Visiting Researcher an der Harvard University.

Info

Text: Monika Weiner
Fotos: Simon Koy

Text erstmalig erschienen im Porsche Engineering Magazin, Nr. 2/2019

Linksammlung

Link zu diesem Artikel
https://newsroom.porsche.com/de/2020/technik/porsche-engineering-kuenstliche-intelligenz-professor-christoph-luetge-19461.html

Externe Links
https://www.porscheengineering.com/peg/de/

Dr. Ing. h.c. F. Porsche Aktiengesellschaft
Porscheplatz 1
70435 Stuttgart