

08.06.2021 – 09:25 Uhr

Künstliche Intelligenz / Menschen gehen mit Maschinen rücksichtslos um

München (ots) -

- Wissenschaftler der LMU und der Universität London haben untersucht, ob sich Menschen im Umgang mit Systemen der Künstlichen Intelligenz (KI) genauso kooperativ verhalten wie gegenüber ihren Mitmenschen.
- Von Künstlicher Intelligenz wird erwartet, dass sie Rücksicht auf den Menschen nimmt. Doch umgekehrt kann davon keine Rede sein.
- Online-Experimente belegen, dass Menschen Maschinen ausnutzen und dabei keine Schuldgefühle entwickeln.

Ein Team von Wissenschaftlerinnen und Wissenschaftlern der LMU und der Universität London hat untersucht, ob sich Menschen im Umgang mit Systemen der Künstlichen Intelligenz (KI) genauso kooperativ verhalten wie gegenüber ihren Mitmenschen. Eine Thematik, die in Zukunft eine besondere Relevanz in unserem Alltag haben wird, wenn wir an das Autonome Fahren denken.

"Kooperation hält unsere Gemeinschaft zusammen. Sie erfordert Kompromiss- und Risikobereitschaft, da das Vertrauen, das wir anderen entgegenbringen, immer auch ausgenutzt werden kann", erklärt Jurgis Karpus, Ph.D., Mitarbeiter am Lehrstuhl für Philosophy of Mind der LMU. "Der Autoverkehr ist dafür ein gutes Beispiel: Wir verlieren etwas Zeit, wenn wir jemandem Vorfahrt gewähren, und sind verärgert, wenn andere es uns nicht gleichtun."

Im Rahmen der Studie wurden in Online-Experimenten verschiedene Situationen mit Methoden der verhaltensorientierten Spieltheorie modelliert, in denen Mensch und Maschine zusammentreffen.

"Unsere Studie zeigt, dass Menschen Maschinen zunächst dasselbe Vertrauen entgegenbringen wie ihren Mitmenschen: Die meisten gehen davon aus, auf Kooperationsbereitschaft zu treffen", sagt Karpus. Doch dann beginnen die Unterschiede: "Menschen sind sehr viel weniger bereit, sich einer KI gegenüber reziprok zu verhalten als gegenüber einem Menschen. Sie beuten sogar die 'Gutmütigkeit' der Maschine zum eigenen Vorteil aus. Im Autoverkehr würde ein Mensch einem menschlichen Fahrer die Vorfahrt gewähren, nicht jedoch einem selbst fahrenden Auto."

Im Laufe der Experimente erwies sich dieses Muster als so konsistent, dass in der Studie die Rede von einer "Ausbeutung von Algorithmen" ist. "Dieser Widerwillen zur Kooperation mit Maschinen ist eine Herausforderung für die zukünftige Interaktion zwischen Mensch und KI", sagt Jurgis Karpus.

Kontakt:

Jurgis Karpus, Ph.D.

Lehrstuhl für Philosophy of Mind, Ludwig-Maximilians-Universität München

Tel.: +49 89 218078611

E-Mail: jurgis.karpus@lmu.de

Web: <https://ots.de/rdnZES>

Publikation:

Jurgis Karpus u.a.: Algorithm exploitation: humans are keen to exploit benevolent AI

iScience 2021

<https://www.sciencedirect.com/science/article/pii/S2589004221006477>

Pressekontakt:

Claudia Russo
Leitung Kommunikation & Presse
Ludwig-Maximilians-Universität München
Leopoldstr. 3
80802 München

Phone: +49 (0) 89 2180-3423
E-Mail: presse@lmu.de

Diese Meldung kann unter <https://www.presseportal.ch/de/pm/100057148/100872212> abgerufen werden.