W02_Ethik_Verantwortung

Verantwortung

Erarbeitet von
Marie von Lobenstein

Lernziele 1
Inhalt 1
Einstieg 1
Was ist Verantwortung? 2
Beispiel: öffentliche Verwaltung 3
Take-Home Message 3
Quellen 3
Disclaimer 4

Lernziele

- Verantwortung erklären können
- Beispiele für KI bezogene Verantwortung nennen können
- Erklären können warum Verantwortung ein wichtiges Thema ist

Inhalt

Einstieg

Ich vermute, du hast auch schon von autonomen Fahrzeugen gehört, also fahrerlosen Fahrzeugen. Im amerikanischen Bundesstaat Arizona wurde 2018 eine Fußgängerin von einem autonom fahrenden Auto getötet. Ein Testfahrzeug von Uber erfasste eine Frau, die mit ihrem Fahrrad die Straße überquerte. Die Testfahrerin an Bord des Autos konnte nicht schnell genug eingreifen und die Fußgängerin, Elaine Herzberg, starb durch den

Projekt „KI für alle"
HeiCAD – Heine Center for Artificial Intelligence and Data Science

CC BY

Zusammenstoß. Herzbergs Tod war der erste tödliche Fußgängerunfall, an dem ein selbstfahrendes Auto beteiligt war.

Quelle [1]

Was ist Verantwortung?

Johanna Wagner schreibt in „Künstliche Intelligenzen als moralisch verantwortliche Akteure?“:
Wenn ein Schaden entstanden ist, steht in der Regel die Frage im Raum, wer für den Schaden verantwortlich gemacht werden kann. Dabei geht es zumeist nicht nur um die Frage, wer im rechtlichen Sinn zur Verantwortung gezogen werden kann, sondern auch um die Frage, wer aus moralischer Sicht dafür geradestehen muss.“

Quelle [2]

Begründen kann man Verantwortung aus verschiedenen Perspektiven, juristisch und moralisch, aber auch zum Beispiel marktwirtschaftlich, etc. (Kroll 2020)

Quelle [3]

Darüber hinaus stellt sich die Frage, auf welche Einheit wenden wir diese Verantwortung an. Nehmen wir das Beispiel vom Unfall in Arizona: Was betrachten wir, wenn wir uns über Verantwortung Gedanken machen? Betrachten wir eine einzelne Komponente, also zum Beispiel die Algorithmen des KI-Systems des Fahrzeuges. Betrachten wir ein größeres System, also zum Beispiel das Programmierteam von Uber oder betrachten wir die gesamte Gesellschaftsstruktur.

In der Moralphilosophie der Gegenwart ist Verantwortung ein wichtiger Begriff. Mit Verantwortung wird meist eine Beziehung beschrieben, in der Informationen an eine Identität gemeldet werden und im Gegenzug Lob, Missbilligung oder, wenn angemessen, Konsequenzen folgen. Wir müssen bei dem Thema Verantwortung also bestimmte Rollen in dieser Beziehung zwischen Information und an die zu meldende Identität bestimmen. Genauer gesagt: Wer oder was muss für wen oder was Verantwortung übernehmen?

Quelle [4]

Sagen wir mal ein KI-System hat entschieden, dass ich mich nicht für die Stelle eigne, für die ich mich beworben habe. Einige Wochen später kommt raus, dass dieses KI-System Frauen aus dem Bewerbungsprozess ausgeschlossen hat, weil das System aus verzerrten Trainingsdaten falsche Schlüsse gezogen hat und nicht, weil ich unterqualifiziert war oder meine Bewerbung schlechter war als die der männlichen Bewerber. Wer, also die Personen, die das KI-System für die Bewerbung eingesetzt haben, die Programmierer*innen des KI-Systems oder das KI-System selbst, sind jetzt dafür verantwortlich, dass ich nicht weiter gekommen bin in dem Bewerbungsprozess? Ist es die Firma, die diese KI benutzt oder das KI-System selbst, die Programmierer*innen oder gesellschaftliche Strukturen, die verzogene Trainingsdaten hervorbringen? Muss es überhaupt eine verantwortliche Instanz geben, bei diesem Beispiel?

Lasst uns Verantwortung an einem weiteren Beispiel ansehen

CC BY

Beispiel: öffentliche Verwaltung

2019 wurde in Schweden ein Computersystem eingeführt, das automatisch überprüfen sollte, ob Menschen in einer bestimmten Form des Arbeitslosenbezuges ihre Verpflichtungen erfüllten. Das KI-System wurde nicht nur genutzt zu bewerten, ob diese Verpflichtungen erfüllt wurden, sondern konnte dazu Verwarnung schicken und Zahlung zurückhalten. Bei genauerem Hinsehen wurden Unregelmäßigkeiten bemerkt. Nach einer Überprüfung des KI-Systems stellte sich heraus, dass 10−15 % seiner Entscheidungen falsch waren. Über 70.000 Antragssteller*innen erhielten aufgrund des fehlerhaften Algorithmus zu Unrecht ihre Leistungen nicht. Das Programm wurde daraufhin aus dem Verkehr gezogen. Bis heute ist unklar, wo der Fehler im System lag und wie er entstehen konnte.

Quelle [5][6]

Wer ist verantwortlich dafür, dass die Antragsteller*innen ihre Leistungen nicht erhalten haben? Wenn alle nach bestem Wissen und Gewissen das KI-System verwendet haben, kann dann überhaupt jemand verantwortlich gemacht werden? Denken wir etwas weiter in eine Zukunft. Können KI-Syteme selbst moralisch verantwortliche Akteure werden und vertrauen wir ihnen an, diese wichtigen Entscheidungen zu treffen?

Take-Home Message

Im besten Fall führen KI-Systeme korrekt die Aufgaben aus, für die sie angewandt werden. In der Realität machen KI-Systeme noch marginale bis existentielle Fehler, die wir diskutieren sollten und für die wir festlegen müssen, wer verantwortlich ist.

Quellen

Quelle [1] Fernandez, E. (2020). *Who Is Responsible In A Crash With A Self-Driving Car?* Forbes. https://www.forbes.com/sites/fernandezelizabeth/2020/02/06/who-is-responsible-in-a-crash-with-a-self-driving-car/

Quelle [2] Wagner, J. (2020). *Künstliche Intelligenzen als moralisch verantwortlich Akteure? begriffliche Möglichkeiten und pragmatische Notwendigkeiten*. Brill, mentis.

Quelle [3] Kroll, Joshua A. "Accountability in computer systems." *The oxford handbook of ethics of AI* (2020): 181-196.

Quelle [4] Kroll, Joshua A. "Accountability in computer systems." *The oxford handbook of ethics of AI* (2020): 181-196.

Quelle [5] Wills, Tom (2019): Schweden: Fehlerhafter Algorithmus stoppt Zahlungen für über 70.000 Arbeitslose. AlgorithmWatch.

Quelle [6] Wagner, J. (2020). *Künstliche Intelligenzen als moralisch verantwortlich Akteure? begriffliche Möglichkeiten und pragmatische Notwendigkeiten*. Brill, mentis.

Disclaimer

Transkript zu dem Video „W02_Ethik_Verantwortung“, Marie von Lobenstein.
Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz CC-BY 4.0 veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.