

Die weiße KI: Warum wir KI-Darstellungen hinterfragen müssen.

Erarbeitet von
Marie von Lobenstein

Das Thema „weiße KI“ ist äußerst komplex und wird aus verschiedenen Fachdisziplinen und Perspektiven betrachtet. In diesem Video haben wir daher eine Auswahl getroffen, um einen ersten Überblick zu geben. Eine vollständige Behandlung des Themas würde den Rahmen der Lehrveranstaltung sprengen und ist im Rahmen dieses Kurses nicht möglich.

Lernziele	1
Inhalt	2
Einstieg	2
Racial Biases in Roboter Darstellungen	2
Reale, humanoide Roboter	4
Virtuelle Assistenz	4
Stock Images und Archivbilder	5
KI-Darstellung in der Pop-Kultur	6
Warum das „weiß“ sein von KI hinterfragen?	6
Take-Home Message	7
Quellen	7
Disclaimer	8

Lernziele

- Du kannst das Konzept der „weißen“ KI und seine Problematiken einordnen
- Du kannst vier Ausprägungen von „weißer“ KI erläutern
- Du kannst bewerten, warum die überproportionale Darstellung von KI als „weiß“ problematisch ist

Inhalt

Einstieg

In ihrem Vorschlag für einen „Ethikkodex“ für die Interaktion zwischen Mensch und Roboter weisen Riek und Howard

Quelle [1]

auf die „mangelnde Vielfalt bei Morphologie und Verhalten von Robotern“ hin:

Sie sagen:

“Was die *Race* betrifft, so ist die überwiegende Mehrheit der Androiden – mit wenigen Ausnahmen wie Hansons Bina48 – ohne erkennbaren Grund asiatisch oder kaukasisch gezeichnet. Außerdem haben die meisten dieser Roboter ein eurozentrisches Design, was ihr Aussehen, ihr Verhalten und ihre Stimme angeht.”

Quelle [2] (Eigene Übersetzung)

Genau über dieses Phänomen möchte ich in diesem Video mit euch sprechen, da wir in der Mensch-Maschine-Interaktion immer auch die gängigen Bilder von KI reflektieren müssen und unsere eigenen Vorurteile. Und da unser gängiges Bild von KI oft eng verknüpft ist mit einer Darstellung von KI als humanoide Maschine, schauen wir uns das jetzt mal genauer an.

Übrigens: im Verlaufe des Videos werde ich öfter den Begriff „Race“ verwenden und meine damit:

Race ist ein komplexes und vielschichtiges Konzept, das in erster Linie als soziales Konstrukt und nicht als biologische Realität verstanden wird. Das Konzept Race wird verwendet, um Menschen auf der Grundlage physischer Merkmale in Gruppen einzuteilen und Systeme der Macht, Privilegien und Unterdrückung zu etablieren und zu rechtfertigen. Diese soziale Kategorisierung entbehrt einer genetischen oder wissenschaftlichen Grundlage.

Racial Biases in Roboter Darstellungen

Der Begriff der „weißen“ KI bezieht sich meist auf die Beobachtung, dass Technologien der künstlichen Intelligenz, insbesondere virtuelle Assistenten, Spracherkennungssysteme und Gesichtserkennungssoftware, häufig rassistische Vorurteile widerspiegeln und aufrechterhalten. Dieser Begriff bezeichnet aber auch die überproportional präsente Darstellung von KI als weiß, auf die wir und in diesem Video konzentrieren wollen. Was das genau bedeutet, schauen wir uns jetzt an.

Die Forschung hat bereits empirisch nachgewiesen, dass Maschinen durch uns in verschiedene Races eingeordnet werden und dass diese Einordnung die Übertragung der in unserem Alltag anzutreffenden Vorurteile beinhaltet.

Zum Beispiel haben Strait et al.

Quelle [3]

in einem Experiment Online-Reaktionen auf drei Videos analysiert, in denen jeweils ein weiblicher Androide mit einer anderen Identität dargestellt wurde: Schwarz, weiß und ostasiatisch.

„Der Begriff Schwarz wird oft als Selbstbezeichnung von Menschen afrikanischer und afro-diasporischer Herkunft, schwarzen Menschen, Menschen dunkler Hautfarbe people of colo(u)r gewählt. Das großgeschriebene „S“ wird bewusst gesetzt, um eine sozio-politische Positionierung in einer mehrheitlich weiß dominierten Gesellschaftsordnung zu markieren und gilt als Symbol einer emanzipatorischen Widerständigkeitspraxis.“

Quelle [4]

Ihr Ziel war es, festzustellen, ob die gleiche Art von ausgrenzenden und entmenslichenden Kommentaren, die auf reale People of Color (POC) angewandt werden, auch auf diese Roboter angewandt würden.

Einblendung Definition People of Color: „Ein selbst definierter, auf Werten basierender Begriff für Menschen, die sich nicht als Weiße identifizieren, oft abgekürzt mit POC; wird anstelle von „Minderheit“ verwendet, was „weniger als“ bedeutet, und anstelle von „nicht-weiß“, was defizitär ist.“

Quelle [5]

Sie fanden heraus, dass die Wertigkeit der Kommentare gegenüber dem Schwarzen Roboter deutlich negativer war als gegenüber den weißen oder asiatischen Robotern. Sowohl der asiatische als auch der Schwarze Roboter erhielt mehr als doppelt so viele entmenslichende Kommentare wie der weiße Roboter.

Cave et al. unterscheiden in ihrem Paper „Whiteness of AI“ zwischen 4 Kategorien von weißer KI:

Quelle [6]

Reale humanoide Roboter, virtuelle persönliche Assistenten, Archivbilder von KI und Darstellungen von KI in Film und Fernsehen. Wir schauen uns diese 4 Kategorien der „weißen“ KI mal gemeinsam an.

Reale, humanoide Roboter

Reale, humanoide Roboter sind oft weiß, entweder sind sie aus weißem Material gebaut oder sie haben kaukasische Züge. Die meisten der wichtigsten Forschungsplattformen für soziale Robotik haben Roboter hergestellt, die aus weißen Materialien sind.

Quelle [7]

Vielleicht habt ihr diese Roboter schonmal gesehen oder von ihnen gehört:

Einblendung: Fotos Nao, Pepper

Das sind Nao und Pepper, bekannte Beispiele von Robotern, die bereits im Einsatz sind. Sie sind, wie viele andere Roboter ihrer Bauart aus dem Material. Das mag für euch jetzt in dem Beispiel von Nao und Pepper noch etwas weit hergeholt wirken, dass diesen Darstellungen von KI-Vorurteile des Weißseins zu Grunde liegen, da sie eher wie Maschinen als wie Menschen aussehen. Also lasst uns dazu ein weiteres Beispiel heranziehen.

Wer es im Vergleich zu Nao und Pepper zu größerem internationalerem Ruhm geschafft hat und diese erste Kategorie der weißen Roboter auch sehr gut abbildet, ist Sophia von Hanson Robotics.

Einblendung: Foto Sophia abbilden

Diese Maschine besteht in erster Linie aus einem weißen humanoiden Kopf, manchmal auch mit einem Oberkörper.

Andere Beispiele dieser weißen, humanoiden Roboter sind Ameca von Engineering Arts, Alter 3 oder Nadine der Nanyang Technological University.

Einblendung: Foto Ameca, Alter 3, Nadine

BINA48 ist eine Ausnahme dieser weißen Darstellung und relativ unbekannt im Vergleich zu Sophia, die auch von Hanson Robotics gebaut wurde.

Einblendung: Foto BINA48

Hier ist es nochmal wichtig zu sagen, dass Robotik nicht gleich KI ist. Roboter sind erstmal Maschinen, die automatisiert Dinge tun. In den Beispielen, die wir uns angesehen haben, wird KI-Technik in diesen Robotern verwendet.

Virtuelle Assistenz

Obwohl virtuelle Assistenten keine visuellen Merkmale von Race aufweisen, wird „Weißsein“ durch soziolinguistische Marker hervorgehoben.

Quelle [8]

Eine der häufigsten Interaktionen mit KI-Technologie erfolgt über virtuelle Assistenten in Geräten wie Smartphones, die in einem weißen Mittelklasse-Englisch sprechen. Dabei wird oft das Beispiel von Samantha aus dem Film „Her“ herangezogen. Diese wird von Scarlett Johansson gesprochen und ist eine der bekanntesten Darstellungen von virtueller Assistenz in Film und Fernsehen. An diesem Beispiel könnte man denken: Moment, das ist doch nur ein Film? Kann man daraus wirklich Rückschlüsse auf „weiße“ KI-Stimmen ziehen? Im Mai 2024 hat OpenAI ihre Voice-Assistentin Sky vorgestellt, die stark nach Scarlett Johansson klingt.

Quelle [9]

Diese hat daraufhin ein Statement veröffentlicht, in dem sie erklärte, dass sie die Anfrage, Sky ihre Stimme zu verleihen, abgelehnt habe und schockiert darüber sei, dass ihre Stimme offensichtlich kopiert wurde. Das scheint zunächst ein Fall von Urheberrechtsverletzung zu sein, aber was hat das nun mit „weißer“ KI zu tun? Scarlett Johansson erläuterte in ihrem Statement, dass Sam Altman, CEO von OpenAI, ihr gesagt habe, er glaube, dass ihre Stimme dem System helfen könne, die Kluft zwischen Technologieunternehmen und Kreativen zu überbrücken und den Verbraucher*innen zu helfen, sich mit dem seismischen Wandel in Bezug auf Menschen und KI wohlfühlen. Er sagte, er habe das Gefühl, dass ihre Stimme die Menschen beruhigen würde.

Da unser gängiges Bild von verkörperter und vertonter KI fast ausschließlich auf „weißen“ Darstellungen beruht, war es für Altman naheliegend, dies auch in unserer Realität zu reproduzieren, weil Scarlett Johanssons Stimme „beruhigend“ wirkt und uns bekannt ist. Das Bild von KI wird also nicht nur in Film und Fernsehen durch Weißsein geprägt.

Stock Images und Archivbilder

Wie wir auch schon in dem Video zu Anthropomorphisierung von KI bemerkt haben, sind die Darstellungen von KI oft menschlich. Diese Darstellungen sind aber nicht nur oft anthropomorphisiert, sondern auch überwiegend weiß. Je realistischer diese Maschinen als „menschlich“ dargestellt werden, desto „weißer“ sind ihre Merkmale.

Quelle [10]

Wenn man Begriffe wie „Roboter“ oder „künstliche Intelligenz“ in eine Suchmaschine eingibt, erhält man eine Fülle von Bildern weißer Humanoiden.

Einblendung: Suchanfragen

Vielleicht noch bemerkenswerter ist, dass diese Maschinen nicht nur eine weiße Farbe haben, sondern: Je menschlicher sie aussehen sollen, desto mehr werden ihre Merkmale ethnisch „weiß“.

Quelle [11]

Solche Suchen sind aus zwei Gründen ein nützlicher Indikator dafür, wie ein Thema zu einem bestimmten Zeitpunkt dargestellt wird. Erstens sind Suchmaschinen sehr weit verbreitet (täglich werden etwa 3,5 Milliarden Suchanfragen bei Google gestellt, das sind 40 000 pro Sekunde) und können daher als äußerst einflussreiche Quelle für Informationen und Wahrnehmungen betrachtet werden.

Quelle [12]

Zweitens bedeutet die Natur solcher Suchmaschinen, dass sie nicht nur bestimmte Ideen und Wahrnehmungen fördern, sondern auch deren Verbreitung widerspiegeln. Auch wenn die genaue Art der Such-, Ranking- und Ergebnisdarstellungsalgorithmen von Google geschützt ist, wissen wir, dass sie (sehr grob gesagt) Einfluss und Popularität bewerten – zum Beispiel in Bezug darauf, wie viele andere Websites auf eine bestimmte Website verlinken. Die Tatsache, dass bestimmte Bilder angezeigt werden, wenn jemand nach einem relevanten Begriff sucht, bedeutet also nicht nur, dass diese Bilder von einigen der mächtigsten Organe der Inhaltsvermittlung, die es heute gibt, gefördert werden, sondern auch, dass diese Bilder bereits weit verbreitet sind und auf anderen einflussreichen Websites verwendet werden, denn das ist die Grundlage für ihre Förderung durch die Suchmaschinen.

KI-Darstellung in der Pop-Kultur

In der Science-Fiction ab dem neunzehnten Jahrhundert wird die KI überwiegend als „weiß“ vorgestellt.

Von „Terminator“ bis „Blade Runner“, von „Metropolis“ bis „Ex Machina“, alle diese Darstellungen von KI werden von weißen Schauspieler*innen verkörpert oder sind auf dem Bildschirm sichtbar weiß. Androiden aus Metall oder Plastik erhalten weiße Gesichtszüge, wie in „I, Robot“. Selbst körperlose KI – von HAL-9000 bis Samantha in „Her“ – werden von „weißen“ Sprecher*innen gesprochen. Erst in jüngster Zeit haben einige Fernsehserien, wie Westworld, KI-Charaktere mit unterschiedlichen Hautfarben verwendet.

Warum das „weiß“ sein von KI hinterfragen?

Die häufige Darstellung und Verkörperung von KI als „weiß“ – von Google Suchergebnissen und Kinorobotern bis hin zu den Dialekten virtueller Assistenten – schließt People of Color aus den gängigen Visionen einer High-Tech-Zukunft aus.

In einem Interview argumentiert Kanta Dihal:

„Die Darstellung von KI als weißer Mensch stellt Maschinen in eine Machthierarchie über derzeit marginalisierten Gruppen und degradiert People of Color auf Positionen unterhalb der von Maschinen. Da Maschinen zunehmend in den Mittelpunkt automatisierter Entscheidungsfindung in Bereichen wie Beschäftigung und Strafjustiz rücken, könnte dies sehr folgenreich sein.“

Quelle [13]

Wir müssen als Gesellschaft die sozialen Auswirkungen von KI-Darstellungen berücksichtigen. Die Dominanz weißer Darstellungen in der KI anzusprechen und in Frage zu stellen, kann helfen, Vorurteile zu bekämpfen und soziale Gerechtigkeit zu fördern. Dies steht im Einklang mit den umfassenderen Bemühungen, Technologien zu entwickeln, die allen Gemeinschaften in gleicher Weise dienen.

KI-Systeme müssen kulturell und sprachlich sensibel sein, um globale Bevölkerungsgruppen wirksam zu unterstützen. Das bedeutet, dass die Vielfalt von Sprachen, Dialekten und kulturellen Nuancen anerkannt und respektiert werden muss. Weiß-zentrierte KI-Darstellungen können diese Aspekte übersehen, was zu Entfremdung und Frustration bei People of Color Nutzer*innen führt.

Take-Home Message

Zusammenfassend lässt sich sagen, dass KI an sich zwar nicht „weiß“ ist, aber die Art und Weise, wie sie oft dargestellt wird, schon. Künstliche Intelligenz ist in vielerlei Hinsicht ein Spiegel unserer Gesellschaft und ihrer Vorurteile. Dies zeigt sich besonders deutlich in der Art und Weise, wie KI und Roboter oft als „weiß“ dargestellt werden. Diese Darstellungen sind nicht nur zufällige Designentscheidungen, sondern können auch tief verwurzelte rassistische Vorurteile widerspiegeln. Um dem entgegenzutreten, müssen wir also die gängigen Darstellungen von KI und ihren Einfluss betrachten und reflektieren.

Quellen

Quelle [1][2]

Riek, L., and Howard, D. (2014). 'A code of ethics for the human-robot interaction profession'. In *Proceedings of We Robot*. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2757805>.

Quelle [3] Strait, M., Ramos, A. S., Contreras, V., and Garcia, N. (2018). 'Robots racialized in the likeness of marginalized social identities are subject to greater dehumanization than those racialized as white'. In *The 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 452–57. <https://doi.org/10.1109/ROMAN.2018.8525610>.

Quelle [4][5]

Schwarz. *People of Colo(u)r*. Diversity Arts Culture. (n.d.). <https://www.bu.edu/diversity/resources/self-guided-education/living-language-guide/race-racism-anti-racism/>

Quelle [6][10][11][12][13]

Cave, S., & Dihal, K. (2020). The whiteness of AI. *Philosophy & Technology*, 33(4), 685-703.

Quelle [7] Bartneck, C., Yogeewaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., and Eyssel, F. (2018). 'Robots and racism'. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 196–204. New York: ACM. <https://doi.org/10.1145/3171221.3171260>.

Quelle [8] Sweeney, M. (2016). 'The Ms. Dewey "experience:" technoculture, gender, and race'. *Digital Sociologies*. <https://doi.org/10.2307/j.ctt1t89cfr.31>.

Quelle [9] Murphy, M. (2024, May 21). *Chatgpt: Scarlett Johansson left "angered" by chatbot imitation*. BBC News. <https://www.bbc.com/news/articles/cm559l5g529o>

Quelle [10] Whiteness of ai erases people of colour from our "imagined futures", researchers argue. University of Cambridge. (2020, August 6). <https://www.cam.ac.uk/research/news/whiteness-of-ai-erases-people-of-colour-from-our-imagined-futures-researchers-argue>

Disclaimer

Transkript zu dem Video „KI und Ich: Die Weiße KI“, 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](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.