

MeMo:KI - Factsheet Nr. 2 - August 2020

Künstliche Intelligenz und Diskriminierung

Diskriminierung durch Künstliche Intelligenz (KI) wird in der Bevölkerung allgemein nur als mäßiges Risiko wahrgenommen. Wenn es jedoch um negative wirtschaftliche Konsequenzen geht, wird der Einsatz von KI eher kritisch betrachtet. Viele Bürger*innen wünschen sich eine stärkere Regulierung von KI.

Hintergrund

Dass Diskriminierung ein zentrales gesellschaftliches Problem ist, wurde nicht zuletzt durch die weltweit stattfindenden Black Lives Matter Demonstrationen deutlich. Doch nicht nur People of Color sind von Diskriminierung betroffen – tagtäglich werden Menschen aufgrund ihres Geschlechts, ihres Glaubens, ihrer Sexualität oder ihres familiären Hintergrunds benachteiligt. Zuletzt gerieten auch immer wieder Systeme mit Künstlicher Intelligenz im Zusammenhang mit Diskriminierung in den Fokus. So gab es in der Vergangenheit bereits Beispiele von diskriminierenden KI-Anwendungen in der Personalrekrutierung, in der Polizeiarbeit oder beim Einsatz von Chatbots (Einen Überblick finden Sie hier: [Plattform Lernende Systeme, 2019](#)).

Wie im [Dashboard](#) des [MeMo:KI] abzulesen ist, steht die deutsche Bevölkerung dem Einsatz von KI zumindest in einigen Anwendungsbereichen durchaus positiv gegenüber. Unklar ist jedoch, ob sich deutsche Bundesbürger*innen des Risikopotenzials von KI-Technologien bewusst sind und in welchen konkreten Anwendungsbereichen Bürger*innen überhaupt Diskriminierungspotenzial durch KI vermuten. Darüber hinaus stellt sich die Frage, wie Diskriminierung verhindert werden kann. Die [EU-Kommission](#) hat in diesem Zusammenhang Leitlinien für vertrauenswürdige KI aufgestellt, die diverse Gegenmaßnahmen zur Bekämpfung von Diskriminierung durch KI vorschlagen. Für eine politische Legitimation dieser Gegenmaßnahmen ist es deshalb auch relevant zu erfassen, inwieweit Bürger*innen diese als wirksam erachten und welche am dringlichsten umgesetzt werden sollten.

In der neuesten Erhebung des [MeMo:KI] wurden 1.022 deutsche Bundesbürger*innen zum Themenkomplex „KI und Diskriminierung“ befragt. Die zentralen Ergebnisse werden in diesem Factsheet präsentiert. Da für eine Bewertung der Fragen ein grundlegendes Verständnis von Künstlicher Intelligenz vorausgesetzt ist, wurden in den folgenden Analysen nur Bürger*innen berücksichtigt, die angegeben haben, dass sie wissen, was der Begriff bedeutet bzw. diesen sogar erklären können. Dies betrifft in dieser Befragung 915 Personen (89,5% der Befragten).

Methodensteckbrief

Methode:	Onlinebefragung
Ausführendes Institut:	forsa Politik & Sozialforschung GmbH
Grundgesamtheit:	Deutsche Bevölkerung ab 18 Jahren, die mindestens gelegentlich das Internet nutzt
Stichprobe:	gewichtete Zufallsstichprobe; N=1022
Gewichtungskriterien:	Alter, Geschlecht und Region (Bundesland)
Erhebungszeitraum:	Kalenderwoche 31 (27.07.-31.07.)
Weitere Informationen:	Ausführlicher Methodensteckbrief

EIN PROJEKT VON:



CENTER FOR
ADVANCED
INTERNET STUDIES



Heinrich Heine
Universität
Düsseldorf

GEFÖRDERT DURCH:

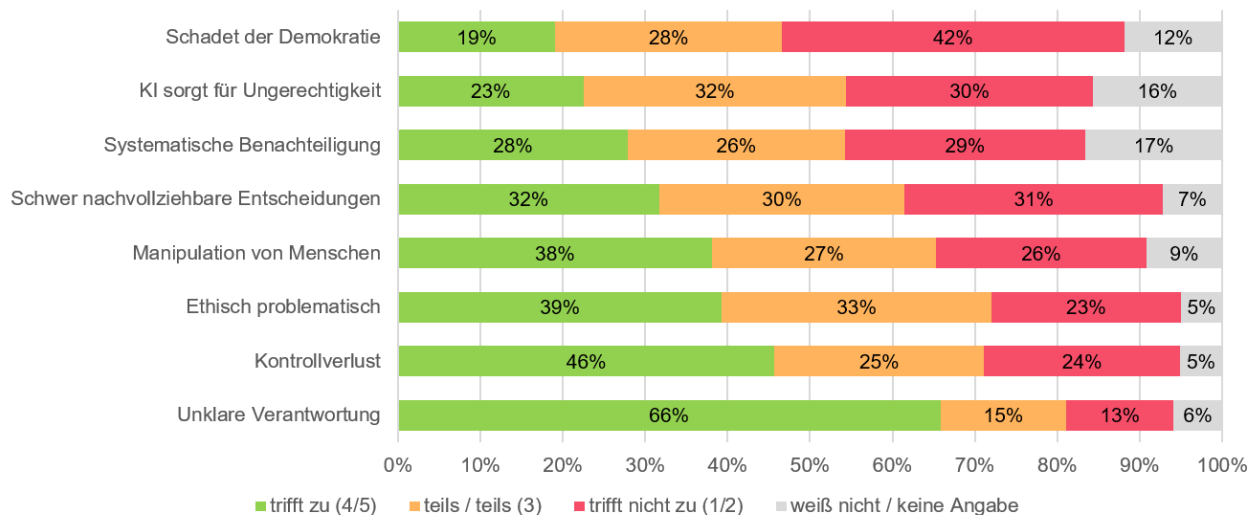
Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen



Die Befragten sehen nur mittlere Risiken in Diskriminierung durch KI

Diskriminierung ist eines von vielen möglichen Risiken, die KI-Technologie mit sich bringen kann. Wir haben die deutschen Bundesbürger*innen gefragt, wie stark sie verschiedene Probleme beim Einsatz von KI für die Gesellschaft einschätzen. Sie konnten ihre Meinung auf einer fünfstufigen Skala abstimmen, wobei für (1) „trifft nicht zu“ und (5) für „trifft voll und ganz zu“ stand. Für die Auswertung wurden die Abstufungen (1) und (2) für „kein risikobewusstsein“ und (4) und (5) für „risikobewusstsein“ zusammengefasst.

Abbildung 1: Risikobewusstsein gegenüber KI-Technologie



Anmerkung: N=913

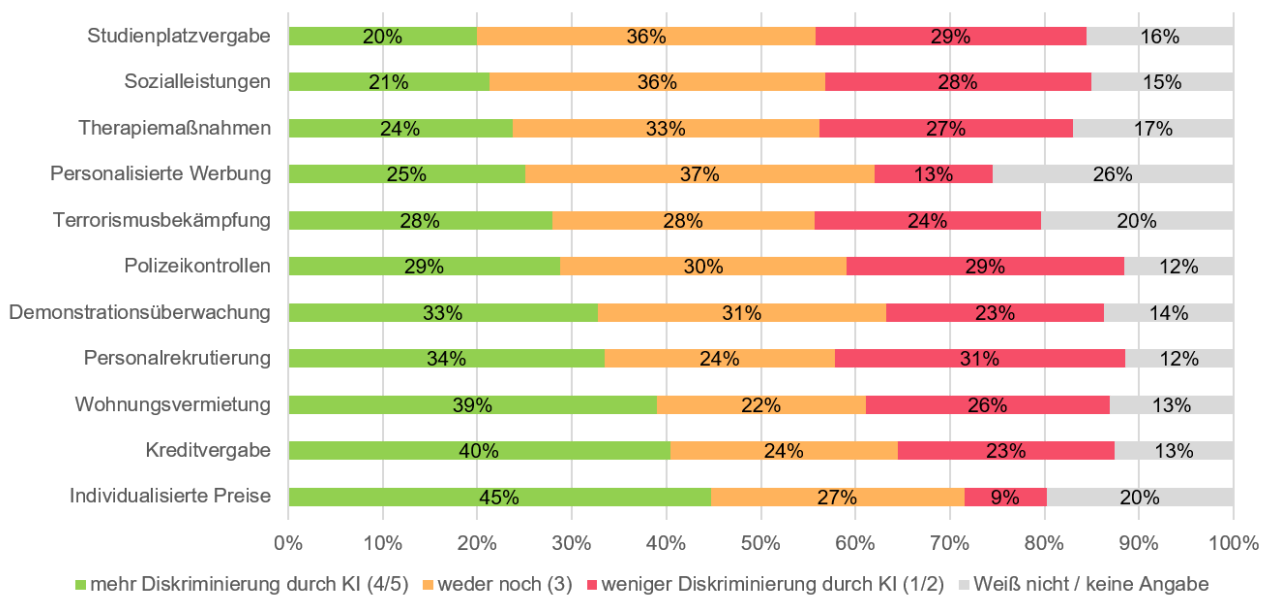
Fragestext: Jetzt geht es auch um Ihre persönliche Meinung zum Einfluss von Künstlicher Intelligenz auf die Gesellschaft. Bitte beurteilen Sie, inwieweit die folgenden Aussagen auf KI zutreffen oder nicht zutreffen.

Die Daten zeigen, dass Diskriminierung (Antwortoptionen „systematische Benachteiligung“) nur von circa 28% der Befragten als großes oder sehr großes Risiko gesehen wird. Der Anteil derer, die hingegen (eher) keine Diskriminierungsgefahr im Einsatz von KI sehen, überwiegt mit 29% leicht. Stark wahrgenommene Risiken sind dagegen die unklare Verantwortlichkeit beim Einsatz von KI-Technologien (66% Zustimmung) oder ein möglicher Kontrollverlust (46% Zustimmung). Dass KI jedoch für Ungerechtigkeiten sorgen kann (23% Zustimmung) oder gar der Demokratie schadet (19% Zustimmung) wird hingegen als geringes Risiko eingeschätzt. Insgesamt zeigt sich – gerade im Vergleich zu den anderen Risikopotenzialen – dass Diskriminierung durch KI von den Befragten eher als ein geringes Problem eingestuft wird.

Sorge vor Diskriminierung bei wirtschaftlichen Themen

Diskriminierung findet in unterschiedlichen Gesellschaftsbereichen statt. Die Antidiskriminierungsstelle des Bundes hat 2017 einen Bericht über den Status Quo von Diskriminierung in Deutschland veröffentlicht, in dem unter anderem problematische Bereiche aufgelistet werden – in vielen dieser Bereiche werden heute auch KI-Systeme eingesetzt. Darüber hinaus wurde im Auftrag des Europarates ein rechtswissenschaftliches Gutachten zu Diskriminierungspotenzialen durch den Einsatz von KI herausgegeben. Für unsere Befragung haben wir relevante Anwendungen aus beiden Papieren identifiziert und die Bürger*innen gefragt, ob ihrer Einschätzung nach der Einsatz von KI in diesen Bereichen zu mehr oder weniger Diskriminierung führt. Die Befragten konnten ihre Meinung auf einer fünfstufigen Skala angeben, wobei der Wert (1) für „deutlich weniger Diskriminierung“ durch den Einsatz von KI und (5) für „deutlich mehr Diskriminierung“ durch KI steht. Wie zuvor wurden die beiden oberen und unteren Antwortkategorien jeweils zusammengefasst.

Abbildung 2: Diskriminierungseinschätzung nach Anwendungsbereichen



Anmerkung: N=905

Fragetext: Was glauben Sie: sorgt KI für mehr oder weniger Diskriminierung in den folgenden Bereichen? Mit Diskriminierung ist gemeint, dass eine Person oder eine Gruppe aufgrund besonderer persönlicher Merkmale gegenüber anderen bevorzugt oder benachteiligt behandelt wird.

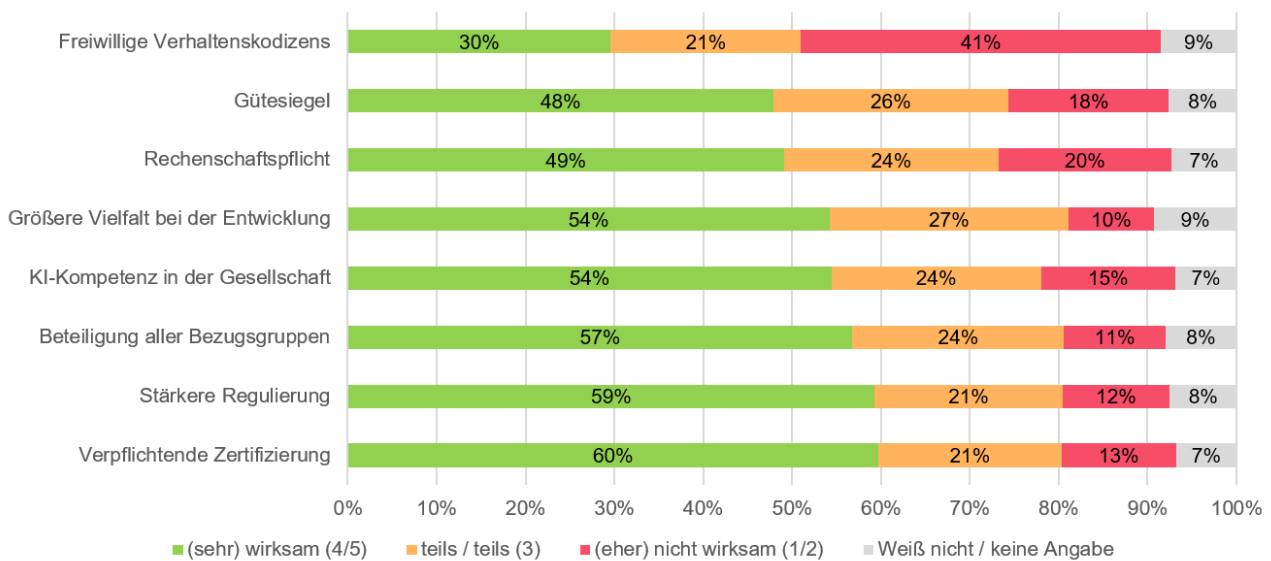
Die Ergebnisse zeigen deutliche Unterschiede in den Erwartungen des Diskriminierungspotenzials durch KI in den einzelnen Anwendungsbereichen. In sieben von elf Anwendungen befürchteten die befragten Personen, dass der Einsatz von KI zu mehr Diskriminierung führt. Allen voran bei KI-Tätigkeiten, die persönliche wirtschaftliche Konsequenzen nach sich ziehen können, wie individualisierte Preisgestaltungen, Kreditvergaben oder Wohnungsvermietungen, wird ein erhöhtes Diskriminierungspotenzial durch KI vermutet. In einigen Anwendungen sehen die Befragten jedoch auch Potenzial für eine geringere Diskriminierung durch den Einsatz von KI. Diese Anwendungen beziehen sich vor allem auf gesellschaftliche bzw. soziale Anwendungen. So zeigt sich, dass bei Therapiemaßnahmen in der Medizin, bei der Verteilung von Sozialleistungen sowie bei der Vergabe von Studienplätzen der Anteil derer überwiegt, die hier eine geringere Diskriminierung durch den Einsatz von KI erwarten. Ein Einsatz von KI in Polizeikontrollen – wie zuletzt häufiger in der medialen Diskussion aufgegriffen – wird von den Befragten ausgewogen aufgenommen; je 29% denken, dass der KI-Einsatz hier zu mehr beziehungsweise weniger Diskriminierung führt.

KI wird in einem Großteil der Anwendungen eher mit steigender als mit sinkender Diskriminierung verbunden. Obwohl das Problembewusstsein für KI-Diskriminierung insgesamt im mittleren Bereich liegt, zeigt sich, dass in konkreten Anwendungen Menschen bzw. Institutionen eine diskriminierungsfreiere Beurteilung zugetraut wird als Maschinen.

Befragte befürworten Regulierung um KI-Diskriminierung zu verhindern

Die EU-Kommission hat in ihren Richtlinien einige Maßnahmen vorgeschlagen, wie Diskriminierung von KI-Anwendungen vorgebeugt werden kann. Diese umfassen wirtschaftliche Eingriffe (z. B. durch Regulierung), aber auch soziale Ansätze (z. B. durch Bildung). Doch für wie wirksam halten Bürger*innen diese überhaupt? Die Einschätzung der Bürger*innen ist insofern elementar, da die öffentliche Meinung politisches Gewicht haben kann, um Umsetzungen anzustoßen, zu beschleunigen oder gegenteilig zu verhindern. Wiederum konnten die Bürger*innen ihre Einschätzung bzgl. der Wirksamkeit der Gegenmaßnahmen auf einer fünfstufigen Skala angeben; für die Visualisierung der Ergebnisse wurden die Antwortoptionen (1) und (2) zu „(eher) nicht wirksam“ und (4) und (5) zu „(sehr) wirksam“ zusammengefasst.

Abbildung 3: Einschätzung der Wirksamkeit von Diskriminierungs-Gegenmaßnahmen

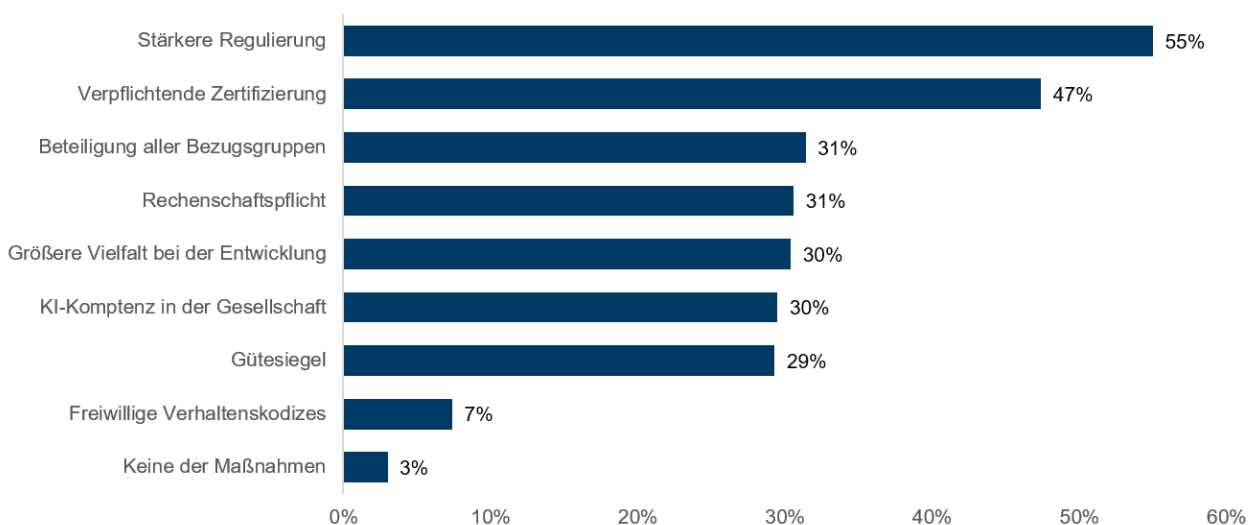


Anmerkung: N=898

Fragetext: Es werden verschiedene Möglichkeiten diskutiert, mit denen man KI-Systeme diskriminierungsfreier gestalten kann. Für wie wirksam halten Sie die folgenden Maßnahmen, um eine mögliche Diskriminierung zu verringern – auch wenn Sie selbst aktuell keine Gefahr sehen?

Es zeigt sich zunächst, dass viele der Maßnahmen, welche die EU zur Bekämpfung von Diskriminierung durch KI vorschlägt, als wirksam erachtet werden. Allen voran wird regulatorischen Eingriffen (z. B. verpflichtende Zertifizierung oder stärkere staatliche Regulierung) eine hohe Wirksamkeit attestiert. Aber auch die Stärkung von Kompetenz im Umgang mit KI und eine größere Einbindung diverser Bevölkerungsgruppen in die Entwicklung und Implementierung wird als wirksam erachtet. Die geringste Wirksamkeit wird dem Aufstellen freiwilliger Verhaltenskodizes von Unternehmerseite zugesprochen. Schließlich sollten die Befragten noch angeben, welche der genannten Gegenmaßnahmen am dringlichsten umgesetzt werden sollte. Die Befragten konnten dabei bis zu drei Gegenmaßnahmen benennen.

Abbildung 4: Priorisierung von Diskriminierungs-Gegenmaßnahmen



Anmerkung: N=898

Fragetext: Und welche der Maßnahmen sollten aus Ihrer Sicht unbedingt umgesetzt werden? Sie können bis zu drei Maßnahmen auswählen. Wenn Ihnen keine der Maßnahmen notwendig erscheint, dann können Sie auch die Antwort „Keine der genannten Maßnahmen“ auswählen.

Die Auswertung der Priorisierung der Gegenmaßnahmen zeigt, dass regulatorische Eingriffe auch den höchsten Zuspruch erfahren. Soziale Faktoren sowie freiwillige Kontrollen werden als nur bedingt dringlich wahrgenommen. Nur drei Prozent der Befragten erachten keine der genannten Maßnahmen als sinnvoll.

Insgesamt zeigt sich in der Beurteilung der Maßnahmen zur Bekämpfung von Diskriminierung durch KI ein bemerkenswertes Bild. Die Befragten sprechen sich für eine stärkere Regulierung und Zertifizierung von KI-Technologie aus, um Diskriminierung zu verhindern. Dies kann als ein Zeichen gedeutet werden, dass der europäische Weg KI mit bindenden Verordnungen zu regulieren und sich nicht allein auf Maßnahmen aus der Wirtschaft zu verlassen, von deutschen Bürger*innen als wünschenswert angesehen wird.

Zusammenfassung

Die Umfrageergebnisse liefern einen Beitrag zur aktuellen Diskussion über Diskriminierung durch KI-Technologien. Insgesamt wird deutlich, dass unter den Befragten ein eher mäßiges Problembewusstsein gegenüber KI-Diskriminierung vorliegt. Wenn Bürger*innen jedoch direkt danach gefragt werden, ob der Einsatz von KI zu mehr oder weniger Diskriminierung in einzelnen Anwendungsbereichen führt, wird bei einer Mehrheit der Anwendungen eher eine stärkere Diskriminierung erwartet – v. a. in Bereichen, die individuelle wirtschaftliche Folgen nach sich ziehen. Ein Einsatz von KI im Sozial- und Bildungswesen wird jedoch eher mit weniger Diskriminierung verbunden. Schließlich hält die Mehrheit der deutschen Bevölkerung, welche ein grundlegendes Verständnis von KI hat, die meisten der vorgeschlagenen EU-Guidelines für wirksam, um Diskriminierung durch KI zu verhindern. Interessanterweise sind die Maßnahmen, die aus Sicht der Befragten am dringendsten umgesetzt werden sollten, solche, die eine stärkere Regulierung von Technik und Wirtschaft fordern. Die grundsätzliche Bereitschaft für den europäischen Weg zu einer menschenzentrierten und gemeinwohlorientierten KI ist demnach durchaus erkennbar – wenngleich das Bewusstsein für diskriminierende Wirkung von KI-Systemen noch relativ gering ist.

Ansprechpartner: Kimon Kieslich, M. A.
Email-Kontakt: memoki@hhu.de
Website: cais.nrw/memoki