KI als Richter? Angst vor KI weltweit im Vergleich



- Die Studie mit mehr als 10.000 Teilnehmende aus 20 Ländern zeigt erhebliche kulturelle Unterschiede in der öffentlichen Angst vor Kl, die Menschen in sechs Berufen ersetzen könnte: Ärztinnen und Ärzte, Richterinnen und Richter, Managerinnen und Manager, Pflegekräfte, Geistliche und Journalistinnen und Journalisten.

Angst entsteht, wenn es eine Diskrepanz zwischen den angenommenen Fähigkeiten der Kl und den für die Rolle erforderlichen menschlichen Eigenschaften gibt.

- Länder wie Indien, Saudi-Arabien und die USA weisen höhere Angstwerte auf, insbesondere wenn Kl als Richterin und Richter oder Ärztin und Arzt eingesetzt wird. Länder wie Japan, China und die Türkei berichten niedrigere Angstwerte, was auf kulturelle Einflussfaktoren hinweist.
- Die Forschung unterstreicht die Bedeutung einer Kl-Entwicklung, die den Erwartungen der Öffentlichkeit entspricht, und schlägt Strategien zum Umgang mit Ängsten vor.

Die Ängste gegenüber dem Einsatz von Künstlicher Intelligenz (KI) am Arbeitsplatz unterscheiden sich je nach Beruf und Land erheblich. Das haben Forschende unter Leitung des Max-Planck-Instituts für Bildungsforschung in einer repräsentativen Studie herausgefunden. Die Forschenden untersuchten die Einstellungen der Bevölkerung in 20 Ländern zur Nutzung von KI in sechs Tätigkeitsfeldern - darunter Medizin, Justiz und Journalismus. Die Ergebnisse, die in der Fachzeitschrift American Psychologist veröffentlicht wurden, können dabei helfen, vorherzusagen, wie neue KI-Entwicklungen in verschiedenen Ländern aufgenommen werden, und Bedenken auf fundierte und kulturell sensible Weise zu adressieren.

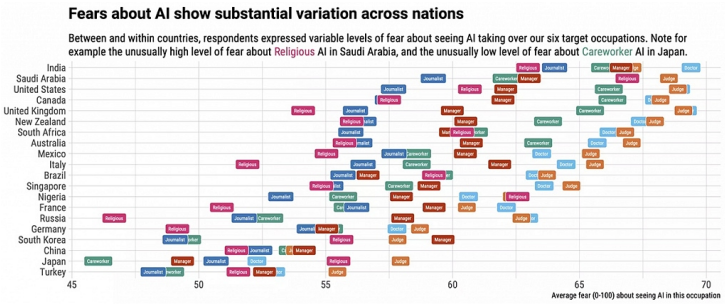
Wie würden Sie reagieren, wenn Sie eine medizinische Diagnose von einer KI erhielten? Würden Sie einem Gerichtsurteil vertrauen, das von einer KI gefällt wurde? Würden Sie Nachrichten vertrauen, die vollständig von einer Maschine verfasst wurden? Und würden Sie sich motiviert fühlen, unter einer KI-gestützten Führungskraft zu arbeiten? Diese Fragen stehen im Mittelpunkt einer aktuellen Studie, die weitverbreitete Ängste vor der Ersetzung menschlicher Arbeitskraft durch KI sowie kulturelle Unterschiede in der Wahrnehmung von KI in sechs Schlüsselberufen untersucht: Ärztinnen und Ärzte, Richterinnen und Richter, Managerinnen und Manager, Pflegekräfte, Geistliche und Journalistinnen und Journalisten.

Mehr als 10.000 Teilnehmende aus 20 Ländern - darunter die USA, Indien, Saudi-Arabien, Japan und China - bewerteten diese sechs Berufe anhand von acht psychologischen Eigenschaften: Wärme, Aufrichtigkeit, Toleranz, Fairness, Kompetenz, Entschlossenheit, Intelligenz und Vorstellungskraft. Sie bewerteten, wie gut KI diese Eigenschaften nachahmen kann, und gaben an, wie groß ihre Angst ist, dass KI diese Rollen übernehmen könnte. Die Ergebnisse deuten darauf hin, dass Menschen beim Einsatz von KI in einem neuen Beruf intuitiv die für diesen Beruf erforderlichen menschlichen Eigenschaften mit der Fähigkeit der KI, diese

nachzuahmen, vergleichen. Auffallend ist, dass das Angstniveau der Teilnehmenden direkt mit der wahrgenommenen Diskrepanz zwischen diesen menschlichen Eigenschaften und den Fähigkeiten der KI zusammenzuhängen scheint.

Die Forschenden stellten erhebliche Unterschiede in der Angst vor KI zwischen verschiedenen Ländern fest. Indien, Saudi-Arabien und die USA weisen die höchsten durchschnittlichen Angstwerte auf, insbesondere in Bezug auf den Richter- oder Arztberuf. Im Gegensatz dazu weisen Länder wie die Türkei, Japan und China die niedrigsten Angstwerte auf, was darauf hindeutet, dass kulturelle Faktoren wie historische Erfahrungen mit Technologie, Darstellungen in den Medien und KI-Politik die Einstellungen stark beeinflussen. In Deutschland sind die Befürchtungen moderat und liegen in der Mitte zwischen dem höchsten und dem niedrigsten Wert. Dies deutet auf eine vorsichtige Offenheit gegenüber der Integration von KI in die Gesellschaft hin.

Die Studie zeigt auch berufsspezifische Unterschiede bei den Ängsten. Richterinnen und Richter wurden in fast allen Ländern als die am meisten gefürchtete KI-Berufsgruppe eingestuft, was Bedenken hinsichtlich Fairness, Transparenz und moralischem Urteilsvermögen widerspiegelt. Im Gegensatz dazu wurde der KI-gestützte Journalismus als am wenigsten bedrohlich empfunden - vermutlich, weil die Menschen nach wie vor selbst entscheiden können, wie sie mit den Informationen umgehen, die sie von Journalistinnen und Journalisten erhalten. Gerichtliche Entscheidungen bieten dahingegen weniger persönlichen Spielraum. Andere Berufe wie KI-gestützte Ärztinnen und Ärzte sowie Pflegekräfte riefen in einigen Ländern starke Ängste hervor, da KI als unfähig gilt, Empathie und emotionales Verständnis zu zeigen.



Durchschnittliche Angstniveaus in 20 Ländern (n = 500 Befragte pro Land) Über den Einsatz von KI in jedem der sechs? [mehr]© Dong, M., Conway, J. R., Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2024). Fears about artificial intelligence across 20 countries and six domains of application. American Psychologist. Advance online publication. https://doi.org/10.1037/amp0001454. Original image licensed under: CC BY-NC-ND 4.0

Dies deckt sich mit den Ergebnissen einer früheren Studie zum Einsatz von KI im Management, in der die Forschenden erste Hinweise darauf fanden, dass Menschen besonders negativ auf KI-Managerinnen und -Manager reagieren, verglichen mit KI-Kolleginnen und KI-Kollegen oder KI-Tools, die bei der Arbeit unterstützen. Diese negative Reaktion war besonders stark in Managementbereichen, die menschliche Fähigkeiten wie empathisches Zuhören oder respektvolles Verhalten erfordern (Dong, Bonnefon, & Rahwan, 2024).

"Negative Auswirkungen können immer dann auftreten, wenn KI in einem neuen Berufsfeld eingesetzt wird. Eine zentrale Aufgabe ist es, diese negativen Effekte zu minimieren, die positiven Effekte zu maximieren und eine ethisch vertretbare Balance der Auswirkungen zu erreichen", sagt Erstautorin Mengchen Dong, Wissenschaftlerin am Forschungsbereich Mensch und Maschine des Max-Planck-Instituts für Bildungsforschung. Die Studie identifiziert einen entscheidenden Zusammenhang zwischen Angst und der Diskrepanz zwischen beruflichen Erwartungen und den wahrgenommenen Fähigkeiten von KI. Dies bietet einen Rahmen für die Gestaltung einer kultursensiblen Entwicklung von KI.

Indem wir verstehen, welche Eigenschaften Menschen in menschenzentrierten Berufen wertschätzen, können Entwicklerinnen und Entwickler sowie politische Entscheidungsträgerinnen und Entscheidungsträger KI-Technologien auf eine Weise entwerfen und kommunizieren, die Vertrauen und Akzeptanz fördert. "Ein universeller Ansatz vernachlässigt entscheidende kulturelle und psychologische Faktoren und könnte potenziell Barrieren für die Einführung nützlicher KI-Technologien in verschiedenen Gesellschaften und Kulturen schaffen", ergänzt Co-Autor Iyad Rahwan, Direktor des Forschungsbereichs Mensch und Maschine am

Max-Planck-Institut für Bildungsforschung.

Die Studie zeigt auch praktische Strategien auf, um Ängste abzubauen. So könnten Bedenken über die mangelnde Aufrichtigkeit von KI-Ärztinnen und KI-Ärzten durch mehr Transparenz in Entscheidungsprozessen und die Positionierung von KI als unterstützendes Werkzeug für menschliche Fachkräfte statt als Ersatz entkräftet werden. Ebenso könnten Ängste vor KI-Richterinnen und KI-Richtern durch den Einsatz von Fairness-Algorithmen und öffentliche Aufklärungskampagnen über die Funktionsweise von KI-Systemen abgebaut werden.

Dong und ihre Kolleginnen und Kollegen setzen ihre Forschung fort, indem sie untersuchen, wie utopische und dystopische Vorstellungen von KI die heutigen Einstellungen in verschiedenen Ländern beeinflussen. Diese laufenden Untersuchungen sollen das Verständnis der Mensch-KI-Interaktion vertiefen und die ethisch und kulturell informierte Einführung von KI-Systemen weltweit begleiten.

Originalpublikation:

Dong, M., Conway, J. R., Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2024). Fears about artificial intelligence across 20 countries and six domains of application. American Psychologist. Advance online publication. https://psycnet.apa.org/fulltext/2025-56995-001.html

Dong, M., Bonnefon, J.-F., & Rahwan, I. (2024). Toward human-centered AI management: Methodological challenges and future directions. Technovation, 131, Article 102953. https://www.sciencedirect.com/science/article/pii/S0166497224000038