



Wenn KI die Fäden zieht

Bereits heute erhalten Mitarbeitende weltweit Arbeitsanweisungen von Algorithmen. Noch misstrauen viele Menschen der Künstlichen Intelligenz (KI). Wissenschaftler untersuchen Chancen und Risiken der Nutzung KI-gesteuerter Führungstools.

► Stellen Sie sich eine Führungskraft vor, und wie diese Ihnen Arbeitsaufgaben und Deadlines zukommen lässt und Ihnen Feedback auf Ihre Arbeitsweise und -ergebnisse gibt. Wie würden Sie sich fühlen, wenn diese Führungskraft kein Mensch, sondern eine Künstliche Intelligenz (KI) wäre? Dieser Gedanke mag erst einmal abstrus wirken, aber er ist möglicherweise naheliegender, als Sie denken.

KI-Führungskräfte: Fiktion oder Alltag?

Künstliche Intelligenz wird immer breiter eingesetzt – von sogenannter generativer KI wie ChatGPT oder Dall-E bis hin zu leistungsstarken Machine-Learning-Modellen, die Vorhersagen zu Ergebnissen von Fußballspielen, Aktienkursen oder sogar Krebsleiden

erstellen. So kommen auch immer mehr Mitarbeitende und Führungskräfte mit verschiedenen Formen dieser neuartigen Technologie in Berührung. Für Unternehmen ergeben sich zahlreiche Möglichkeiten, KI so einzusetzen, dass sie unternehmensförderlich wirken kann. Eine beachtliche Entwicklung findet hier insbesondere im Führungskräftekontext statt. Millionen von sogenannten „Gig-Workern“ (beispielsweise Fahrer und Fahrerinnen bei Firmen wie Uber) erhalten ihre täglichen Arbeitsanweisungen und Evaluationen bereits ausschließlich von Algorithmen. Teilweise übernimmt KI sogar eigenständig die Einstellung und Kündigung von Mitarbeitenden. Neben solchen „einfachen“ Managementtätigkeiten sagt das Weltwirtschaftsforum voraus, dass KI bis 2026 bereits Rollen in Aufsichtsräten großer

Unternehmen einnehmen wird. Umgesetzt hat dies bereits der kolumbianische Rumhersteller Dictador. Die Edel-Rummarke hat vor Kurzem bekannt gegeben, dass die KI-Software Mika die Rolle des Chief Executive Officers (CEO) im Unternehmen übernehmen wird. Während unklar ist, inwiefern solche Entwicklungen auch zum Teil Marketingaktionen sind, verdeutlichen sie eins: Führung mit und durch KI wird in (naher) Zukunft viele Unternehmen begleiten – und sie tut dies bereits für mehrere Millionen Beschäftigte weltweit.

Aversion oder Wertschätzung gegenüber KI?

Dies führt zu der Frage, wie Menschen mit solchen KI-Führungskräften umgehen werden. Werden sie deren Anweisungen Folge leisten? Neueste Forschung scheint darauf hinzudeuten, dass Menschen Führung durch KI häufig noch skeptisch gegenüberstehen. So werden Führungsanweisungen und Personalentscheidungen durch KI oftmals als reduktionistisch sowie weniger zufriedenstellend und vertrauenswürdig empfunden – selbst wenn die Entscheidungen identisch zu denen von Menschen sind oder sich als vorteilhaft für die menschlichen Mitarbeitenden erweisen würden. Forschende führen dies auf die sogenannte „Algorithm Aversion“ (übersetzt: Aversion gegenüber Algorithmen) zurück, die vom amerikanischen Forscher Berkeley Dietvorst (University of Chicago) erstmals benannt wurde. Er zeigte in einer Reihe von Studien, dass Menschen eine generelle Ablehnung gegenüber Algorithmen an den Tag legen. Weitere Studien anderer Arbeitsgruppen deuten darauf hin, dass die Ablehnung von KI insbesondere in ethischen Fragestellungen, aber auch in HR-Kontexten verbreitet ist.

Interessanterweise zeigen neuere Erkenntnisse jedoch auch, dass Menschen in manchen Situationen einer KI sogar mehr Vertrauen entgegenbringen als einem Mitmenschen. Dieses Phänomen wurde von Jennifer Logg (Georgetown University) und Kolleginnen und Kollegen als „Algorithm Appreciation“ (übersetzt: Wertschätzung von Algorithmen) bezeichnet. So zeigte zum Beispiel eine Studie von amerikanischen Forschenden auf, dass Mitarbeitende bei der Arbeit lieber von Algorithmen als von Menschen überwacht und bewertet werden möchten. In mehreren Studien fanden zudem Florian Pethig (Universität Mannheim) und Julia Krönung (Fernuniversität Hagen) heraus, dass Frauen ihre Bewerbungsunterlagen lieber von einem Algorithmus als von einem männlichen, menschlichen Personal bewertet lassen möchten. Auch wenn die Bewerberinnen dem Algorithmus

durchaus Entscheidungsverzerrungen und Fehler zutrauten, dachten sie, dass die KI im Vergleich zu einem potenziell von Vorurteilen geprägten männlichen Entscheider das „kleinere Übel“ darstelle.

Unethische Anweisungen von KI-Führungskräften

In unseren eigenen Studien haben wir mit rund 1700 US-Amerikanern untersucht, wie Menschen auf unethische Anweisungen von KI-Führungskräften

Die Ablehnung von KI insbesondere in ethischen Fragestellungen, aber auch in HR-Kontexten scheint verbreitet zu sein.

reagieren. Wir konfrontierten Teilnehmende mit einer fiktiven Situation, in der ihre Führungskraft sie dazu aufforderte, eine Person finanziell zu benachteiligen – nur auf Basis der Information, dass diese alleinerziehend sei. Das Szenario unserer Studie war inspiriert durch eine fragliche Überlegung der Firma H&M, über die kürzlich in der Presse berichtet wurde. Wir teilten die Teilnehmenden zufällig einer von zwei Experimentalgruppen zu: Die einen erhielten die Anweisungen von einer menschlichen

Wessen unethische Anweisungen werden häufiger befolgt?

Experimentalgruppe: Prozentuale Befolgung der Anweisung



Quelle: Lanz/Binker/Seppelt

In der Experimentalgruppe mit menschlicher Führungskraft setzen die Teilnehmenden die Anweisung, das Gehalt der alleinerziehenden Person zu kürzen, durchschnittlich zu 31 Prozent um. Bei der KI-Führungskraft waren die Teilnehmenden mit circa 23 Prozent deutlich zurückhaltender.

STUDIE KOMPAKT

Forschungsfrage: Wem folgen Mitarbeitende bei unethischen Anweisungen mehr – KI oder menschlichen Führungskräften?

Forschungsansatz: Vier experimentelle Untersuchungen mit 1701 US-Amerikanern mit Berufserfahrung.

Ergebnisse: Mitarbeitende befolgen diskriminierende, unethische Anweisungen signifikant weniger, wenn diese von einer KI kommen, nicht von einer menschlichen Führungskraft. Eine vermenschlichte Darstellung von KI lindert diesen Effekt nicht. Machine-Learning-Analysen deuten zudem darauf hin, dass dieser Effekt vermutlich damit zu tun hat, dass KI als weniger intelligent und emotional versiert wahrgenommen wird.

Führungskraft, die anderen von einer KI-Führungskraft. Die Ergebnisse: Die Teilnehmenden folgten den Empfehlungen signifikant mehr, wenn es sich um eine menschliche statt einer KI-Führungskraft handelte. Die Abbildung zeigt diese Effekte, welche wir in mehreren weiteren Studien replizieren konn-

Machine-Learning-Tools können leistungsstarke Hilfsmittel sein, um Potenziale von Führungskräften zu entdecken und zu fördern.

ten. Selbst wenn die KI menschlicher und nahbarer dargestellt wurde oder wenn es um tatsächliches Geld anstatt eines fiktiven Szenarios ging, blieben die Ergebnisse unverändert. Weitere Auswertungen mittels Machine-Learning-Analysen gaben Hinweise auf den Hintergrund dieser Resultate. Sie lassen vermuten, dass die stärkere Ablehnung von Anweisungen der KI-Führungskraft darauf beruht, dass Teilnehmende diese Technologie als weniger fähig einschätzen, intelligente Entscheidungen zu treffen und Emotionen zu spüren. Zusätzlich wurde deutlich, dass ältere Menschen der KI ablehnender gegenüberstanden als jüngere. Gleichmaßen zeigten Teilnehmende, die angaben, im Berufsalltag ihrer (menschlichen) Führungskraft weitgehend uneingeschränkt zu folgen, deutlich stärkeren Widerstand gegen die KI als diejenigen, die kritischer gegenüber Anweisungen ihrer Führungskraft eingestellt sind. Dies könnte damit zusammenhängen, dass solche Personen nur menschlichen Führungskräften nahezu blind Folge leisten, Maschinen gegenüber hingegen nicht. Diese Ergebnisse beruhigen dahingehend, dass Menschen unethischen Anweisungen von KI (noch) nicht blind oder in hohem Ausmaß folgen. Dennoch bleibt natürlich abzuwarten, welche Auswirkung eine tatsächliche Einführung von KI Führungskräften als alltägliche Praxis in der Arbeitswelt haben wird. Als momentanes Zwischenfazit lässt sich festhalten, dass wir Mensch und Maschine bevorzugt in einer konstruktiven Zusammenarbeit statt in einer KI-geführten Machtkonstellation sehen.

KI als Führungskraftetool und -coach

Wie genau eine solche Zusammenarbeit in der Praxis aussehen könnte, diskutierten wir vor Kurzem mit dem bekannten Führungsforscher Professor George Banks von der University of North Carolina in Charlotte bei seinem Forschungsbesuch an der Universität Maastricht. Er zeigte hierbei die Möglichkeiten auf, moderne Machine-Learning-Tools als Führungskräfte-

teentwicklungsinstrumente oder Coaches einzusetzen. In seiner Forschung nutzt er KI zum Beispiel in Videokonferenzen, um Gesichtsausdrücke, Emotionen und Verhaltensweisen automatisch kodieren zu lassen. Diese Kodierungen werden mit menschlichen Einschätzungen verglichen und kontinuierlich weiter verbessert. Dahinter steht die Idee, Führungskräften in Echtzeit Feedback darüber geben zu können, wie sie sich wann verhalten haben und wie sie dies direkt verbessern könnten. Zum Beispiel soll das Tool darüber Aufschluss geben, wie beziehungsorientiert sich eine Führungskraft verhält. Bei Auffälligkeiten kann es der Führungskraft zum Beispiel eine Mitteilung zukommen lassen, dass es für sie von Vorteil wäre, mehr Fragen zu stellen und zuzuhören statt selbst die Unterhaltung zu dominieren. Mit solchen Tools wäre es möglich, KI als Führungsunterstützung und -ergänzung statt als Ersatz für Führungskräfte zu nutzen.

Fazit

Menschen misstrauen KI-Führungskräften, insbesondere wenn es um ethische und moralische Fragestellungen geht. Um KI-Führungskräfte effektiv einsetzen zu können, sollten Unternehmen diese möglichst transparent und vertrauenswürdig gestalten. Gleichzeitig geht es nicht darum, menschliche Führungskräfte abzuschaffen, sondern deren Kompetenzen sinnvoll zu ergänzen. Neuartige Machine-Learning-Tools können leistungsstarke Hilfsmittel bieten, um Potenziale von Führungskräften entdecken und fördern zu können. Ganz essenziell ist dabei, Menschen miteinzubeziehen und die KI zu erklären. Nur so können Mensch und Maschine konstruktiv zusammenarbeiten, anstatt einander zu ersetzen. ■

AUTOREN UND AUTORIN



Roman Briker, Juniorprofessor für Organizational Behavior und Personalmanagement, Universität Maastricht, r.briker@maastrichtuniversity.nl



Fabiola H. Gerpott, Professorin für Personalführung an der WHU – Otto Beisheim School of Management, fabiola.gerpott@whu.edu



Lukas Lanz, Doktorand, WHU – Otto Beisheim School of Management, lukas.lanz@whu.edu



Mehr zum Thema:

Lanz, L./Briker, R./Gerpott, F. H.: Employees Adhere More to Unethical Instructions from Human Than AI Supervisors: Complementing Experimental Evidence with Machine Learning, Journal of Business Ethics. Advance online publication. <https://doi.org/10.1007/s10551-023-05393-1>