

Auf die Künstliche Intelligenz allein kommt es nicht an

In den vielen Diskussionen gerade geht es viel zu oft nur um die technologischen Aspekte der KI. Ein Experiment zeigt, wovon ihr Erfolg tatsächlich abhängt.

Von Johannes Lohse und Mario Mechtel

Die Leistungsfähigkeit generativer Künstlicher Intelligenz wächst in einem Tempo, das selbst für das digitale Zeitalter außergewöhnlich ist. Sprachmodelle, denen man gerade noch mit gefühlter Überlegenheit beim Scheitern an den einfachsten Aufgaben zusehen konnte, strukturieren inzwischen komplexe Argumente, fassen anspruchsvolle Texte zusammen und schreiben als agentische KI sogar eigenständig Computerprogramme. Mit sich häufenden Berichten über die Substitution menschlicher Arbeitskraft durch KI wächst auch die Sorge über die Sicherheit des eigenen Jobs.

Die tatsächliche Verbreitung und Nutzung von KI-Systemen hält mit ihren stetig wachsenden technischen Möglichkeiten indes keineswegs überall Schritt. Und das, obwohl sich die wöchentlichen Nutzerzahlen des meistgenutzten Sprachmodells ChatGPT allein innerhalb des letzten Jahres fast verdoppelt haben.

Werkzeuge wie ChatGPT, Claude oder Gemini sind in manchen Bereichen längst fester Bestandteil des Arbeitsalltags. In anderen spielen sie bislang jedoch noch immer nur eine Nebenrolle. So deutet etwa eine Auswertung von Nutzerdaten durch das KI-Unternehmen Anthropic aus der jüngeren Vergangenheit darauf hin, dass selbst in Aufgabengebieten, in denen sich KI-gestützte Automatisierung besonders anbietet, etwa in der Programmierung, bislang weniger als 40 Prozent des technischen Potentials ausgeschöpft werden.

In der öffentlichen Debatte geht es oft darum, was technisch schon möglich scheint – und nicht, was gesellschaftlich, organisatorisch und institutionell erst noch entstehen muss. Das Potential neuer Technologien wird dabei auf kurze Sicht schnell überschätzt, die noch notwendigen Schritte zu ihrer tatsächlichen Verbreitung dagegen unterschätzt.

Beispielsweise zeigt die Forschung der Ökonomen Daron Acemoglu oder Erik Brynjolfsson, dass technologischer Fortschritt vom Buchdruck bis zum Internet nicht von selbst Wohlstand schafft. Seit Jahrzehnten wird in der Wirtschaftswissenschaft immer wieder das sogenannte Produktivitätsparadoxon diskutiert. Zur Erinnerung: Süffisant hatte der Nobelpreisträger Robert Solow einmal formuliert, dass man das Computerzeitalter überall sehen könne, außer in den Produktivitätsstatistiken.

Entscheidend ist also nicht immer die Technologie im abstrakten Sinn, sondern die Art und Weise, in der sie sich in bestehende Prozesse und Institutionen integrieren lässt. Das Muster ist vertraut: Schon lange bevor wir über KI diskutiert haben, war es möglich, zahlreiche Alltags- und Arbeitsprozesse zu digitalisieren. Doch zwischen Möglichkeit und Wirklichkeit lag schon damals ein weiter Weg – der vielerorts bis heute nur teilweise zurückgelegt worden ist.

Ein in der aktuellen Debatte um die Auswirkungen von KI-Sprachmodellen unterschätzter Faktor ist der Nutzer selbst. Die Produktivitätsgewinne der KI sind nicht einfach ein Produkt ihrer technologischen Möglichkeiten. Sie hängen wesentlich davon ab, wer sich die neuesten Modelle aneignet und wie diese dann genutzt werden. An diesem Punkt setzt eine aktuelle experimentelle Studie an, die wir mit Mark Praet vom Fraunhofer-Institut für Angewandte Informationstechnik durchgeführt haben.

Wir wollten nicht nur wissen, ob der Zugang zu einem großen Sprachmodell die Leistung von Studierenden verbessert. Uns interessierte vor allem, wie groß der Abstand zwischen Zugang und tatsächlicher Nutzung ist, wie stark dieser Abstand je nach Aufgabe variiert und was daraus

für die Messung von KI-Produktivitätseffekten folgt.

Für unsere Studie haben 210 Studierende Aufgaben aus drei Bereichen bearbeitet, die einem in fast jedem akademischen Berufsbild begegnen: kreative Problemlösung, Textverständnis und quantitatives Denken.

Die Versuchsanordnung war dabei bewusst einfach. Eine Hälfte der Teilnehmenden hatte Zugriff auf ChatGPT, die andere nicht. Anders als in vielen bisherigen Onlinestudien war die Kontrollgruppe tatsächlich eine Kontrollgruppe. Wer in ihr saß, konnte während der Lösung der Aufgaben nicht nebenbei doch noch auf KI zurückgreifen. Zugleich konnten wir in der KI-Gruppe genau nachvollziehen, ob und auf welche Weise das bereitgestellte Werkzeug tatsächlich genutzt wurde. In einem zweiten Schritt wurden die anonymisierten Antworten dann von 1176 externen Personen bewertet, um festzustellen, wie gut die Aufgaben bearbeitet wurden.

Unser erster Befund ist eindeutig und kaum überraschend: Zugang zu generativer KI verbessert im Durchschnitt die Qualität der Antworten. Wichtiger ist jedoch der zweite Befund: Diese Produktivitätseffekte verteilen sich keineswegs gleichmäßig. Die deutlichsten Zugewinne zeigen sich im Bereich der sprachintensiven Aufgaben – also überall dort, wo Informationen verdichtet, geordnet und in eine klare Form gebracht werden müssen. Auch bei quantitativen Fallstudien, bei denen Annahmen getroffen, Größenordnungen plausibel abgeschätzt und Gedankengänge neu strukturiert werden müssen, wirkt sich der Einsatz positiv aus.

Merklich geringer fallen die Effekte dagegen bei einer Aufgabe aus, mit der wir die Kreativität der Teilnehmenden messen wollten. Hier musste eine möglichst neuartige und nützliche Erfindung erdacht werden. Nahezu ganz verschwinden sie dort, wo visuelle Informationen aus einem Diagramm in einen kurzen Text übersetzt werden sollten. Schon das relativiert die Vorstellung eines einheitlichen Produktivitätsschubs durch KI.

Noch aufschlussreicher sind deshalb unsere Daten zur tatsächlichen Nutzung. Selbst unter Bedingungen, die den Einsatz von KI eher begünstigen dürften, also bei einem direkten Zugang und sogar ausdrücklicher Aufforderung zur Nutzung, griffen längst nicht alle Teilnehmenden tatsächlich auf das angebotene Werkzeug zurück. Nur 29 Prozent setzten die KI in allen vier Aufgaben ein, 16 Prozent nutzten sie dagegen kein einziges Mal. Auch zwischen den Aufgabentypen zeigten sich deutliche Unterschiede. Beim Textverständnis griffen fast drei Viertel der Teilnehmenden auf KI zurück, in den Diagrammaufgaben dagegen nicht einmal die Hälfte.

Hinzu kommen verschiedene Selektionsmechanismen. In unserem Experiment griffen vor allem jene häufiger zur KI-Unterstützung, die schon Vorerfahrung mit solchen Werkzeugen hatten. Auch individuelle Fähigkeitsprofile spielten eine Rolle. Damit wird deutlich, dass die wirtschaftliche Wirkung generativer KI nicht allein von den Fähigkeiten des Modells abhängt, sondern – wenig erstaunlich – natürlich auch von den Fähigkeiten der Nutzer.

KI ist kein Werkzeug, das man nur bereitstellen muss. Ihr Nutzen hängt von komplementären Kompetenzen ab. Diese Kompetenzen zu entwickeln, wird zunehmend auch Aufgabe universitärer Hochschullehre. Im Rahmen des von der Stiftung Innovation in der Hochschullehre geförderten Projektes LAICA (Lüneburg AI Campus) werden wir uns in den kommenden Jahren gemeinsam mit Kolleginnen



und Kollegen der Frage widmen, wie solche Kompetenzen in unterschiedlichen Studiengängen systematisch aufgebaut werden können.

Unsere Ergebnisse rücken eine klassische Einsicht der Ökonomik wieder ins Zentrum: Technologien entfalten ihren Ertrag nicht im luftleeren Raum. Sie brauchen Organisation, anpassungsfähige Institutionen und Zeit. Das galt für frühere Automatisierungswellen, und es gilt auch heute.

Die Vorstellung, das bloße Vorhandensein einer Technologie werde automatisch breite Produktivitätsschübe erzeugen und Massenentlassungen nach sich ziehen, war schon in früheren Innovationsphasen oft zu kurz gegriffen. Im Falle der generativen KI dürfte dies auch gelten.

Was folgt daraus? Wer verstehen will, wie KI Produktivität und Beschäftigung verändert, sollte nicht nur auf das schauen, was die neuesten Modelle theoretisch leisten können. Entscheidend ist auch, wer diese Werkzeuge tatsächlich nutzt und wofür.

Genau hier deutet unsere Studie schon auf erhebliche Unterschiede hin, sowohl zwischen verschiedenen Nutzergruppen als auch zwischen unterschiedlichen Aufgabentypen. Auch andere Untersuchungen kommen zu ähnlichen Ergebnissen und zeigen deutliche Ungleichheiten, etwa zwischen Frauen und Männern.

In einer Metastudie der Harvard Business School lag die durchschnittliche Wahrscheinlichkeit, dass Frauen KI im Beruf nutzen, um 22 Prozent unter der von Männern. Wenn also KI, wie in unserer Studie, auf individueller Ebene zu Produktivitätssteigerungen führt, dann kann ungleiche Nutzung bestehende Ungleichheiten verstärken oder aber auch nivellieren, wenn sie insbesondere von denjenigen eingesetzt wird, die ansonsten eine Aufgabe eher schlecht erfüllen würden. Erste Befunde, die sich auch in unserer Studie bestätigen, deuten darauf hin, dass KI häufiger dort hilft, wo Fähigkeiten schwächer sind, und eben nicht den Abstand der ohnehin Starken immer weiter erhöht.

Neben der bei vielen Menschen ausgeprägten Neugierde beim Ausprobieren der neuesten Sprachmodelle schwankt die öffentliche Wahrnehmung zunehmend zwischen Faszination und Ermüdung. Aus FOMO, also „The Fear of Missing Out“, ist für viele Angestellte längst FOBO geworden, also „The Fear of Becoming Obsolete“, die Angst überflüssig zu werden.

Mit Blick auf diese Sorgen dürften die beschriebenen Erkenntnisse relativierend wirken. Denn am Ende hat man es zumindest teilweise selbst in der Hand, KI produktiv im eigenen Job einzusetzen.

Lehrreich ist zudem eine lange bestehende Einsicht, welche ihren Ursprung schon in der Zeit der industriellen Revolution hat. Würden Dampfmaschinen einmal alle menschliche Arbeit ersetzen, weil sie stärker und schneller als Muskelkraft sind? Das sogenannte Jevons-Paradox besagt, dass Beschäftigung nicht unbedingt immer dort zurückgehen muss, wo eine bessere Technologie eingeführt wird. Senkt diese Technologie nämlich Kosten, wie etwa KI aktuell in der Softwareentwicklung, kann die dadurch ausgelöste zusätzliche Nachfrage am Ende auch zu mehr Beschäftigung führen.

Prof. Dr. Johannes Lohse ist Juniorprofessor für Law & Economics an der Leuphana Universität Lüneburg.

Prof. Dr. Mario Mechtel ist Professor für Empirische Mikroökonomik an der Leuphana Universität Lüneburg und leitet den Wissenschaftsraum „Verhaltensökonomik und gesellschaftliche Transformation“, ein Forschungscluster von sieben niedersächsischen Universitäten und dem Fraunhofer FIT.

Frühere technologische Umbrüche am Arbeitsmarkt haben überwiegend männlich geprägte Branchen wie die Schwerindustrie und die Fertigung betroffen. Generative KI funktioniert anders. Sie automatisiert keine körperlichen Tätigkeiten, sondern zuerst Büro- und Verwaltungstätigkeiten. Also die klassischen Aufgaben des Sekretariats, der Sachbearbeitung, der Buchhaltung, im Kundenservice oder Callcenter. In diesen Funktionen arbeiten in den entwickelten Volkswirtschaften überproportional viele Frauen.

Nach Berechnungen der Internationalen Arbeitsorganisation (ILO) sind Berufe, die von Frauen dominiert werden, international fast doppelt so häufig generativer KI ausgesetzt wie Berufe, in denen Männer in der Überhand sind. In einkommensstarken Ländern fallen 9,6 Prozent der weiblichen, aber nur 3,5 Prozent der männlichen Beschäftigten in die höchste Risikoklasse.

Eine aktuelle Brookings-Analyse quantifiziert das Risiko für die Vereinigten Staaten: Von den rund sechs Millionen Beschäftigten mit hoher KI-Exposition und zugleich geringer Anpassungsfähigkeit sind 86 Prozent Frauen, und überwiegend solche in Verwaltungs- und Büroberufen.

Die Stellenpläne der Großunternehmen folgen dem Muster. Die Reederei

Wie Frauen in Bürojobs der KI-Automatisierung entkommen

Generative KI betrifft vor allem die Stellen, in denen besonders viele Frauen arbeiten. Es gibt schon messbare Verschiebungen – was jetzt zu tun ist.

Maersk hat im vergangenen Jahr weltweit 1000 Verwaltungsstellen gestrichen, der Konsumgüter-Konzern Procter & Gamble und der Internet-Konzern Amazon haben Entlassungswellen vor allem in den unterstützenden Funktionen vorgenommen. In Deutschland kündigte ERGO im Februar an, bis zum Jahr 2030 Stellen in Callcentern, Schadenbearbeitung und Schriftgutbearbeitung abzubauen. Die Commerzbank legte im Mai weitere Streichungen in Operations

und Zentralfunktionen nach, Allianz Partners und ING haben vergleichbare Programme aufgelegt.

All dies sind Branchen mit traditionell hohem Frauenanteil in der Sachbearbeitung in Banken, Versicherungen und Dienstleistungsunternehmen. Parallel dazu sind die Stellenausschreibungen in Branchen mit hohem Frauenanteil überproportional zurückgegangen.

Frauen haben ihren Rückstand in der KI-Nutzung zwar etwas aufgeholt, set-

zen die Technik im Vergleich zu Männern für berufliche Zwecke aber immer noch seltener ein und sind zudem in typischen MINT-Berufen immer noch unterrepräsentiert. Damit stehen sie in der KI-Welle mehrheitlich auf der passiven Seite des Arbeitsmarktes. Fachleute der UN warnen, dass der KI-Trend ohne Gegensteuerung bestehende Ungleichheiten am Arbeitsmarkt weiter verschärfen könnte.

Die überproportional hohe Exposition der Frauen gegenüber KI lässt sich kurzfristig nicht beseitigen, aber Frauen können gegensteuern. Im Zentrum steht die Weiterbildung, also das Erlernen komplementärer KI-Kompetenzen. Denn die ILO sieht Transformation, nicht Verdrängung als das wahrscheinlichste Szenario. Wer mit der Künstlichen Intelligenz arbeitet, zählt mit höherer Wahrscheinlichkeit zu den Gewinnern. Im Global AI Jobs Barometer wurde ein schnell steigender Lohnaufschlag von 56 Prozent für Beschäftigte mit KI-Kenntnissen ermittelt. Löhne wachsen in KI-exponierten Branchen doppelt so schnell wie in unexponierten, auch in den automatisierbaren Berufen.

Im Kern geht es um die Verbindung aus Domänenwissen mit KI. Eine Personalfachfrau mit KI-Werkzeugen schlägt

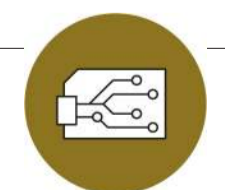
einen Programmierer ohne HR-Erfahrung, eine Steuerassistentin mit DATEV plus KI eine Fachkraft ohne Branchenkenntnis. Die vertikale KI macht jahrelang gewachsene Berufserfahrung zum Engpassgut, nicht zum Auslaufmodell. Hier liegt der größte Hebel für Frauen in exponierten Berufen.

Eine aktuelle Auswertung des Ideed Hiring Lab zeigt den Mechanismus: Während der Gesamtmarkt für Stellenanzeigen seit 2022 schrumpft, ist der Anteil KI-bezogener Anzeigen in Deutschland im Verlauf des Jahres 2025 von rund zwei auf 3,5 Prozent gestiegen. Der Anstieg findet vor allem außerhalb der klassischen Tech-Rollen statt: Im Personalwesen nahm die KI-Nachfrage 2025 um 138,7 Prozent zu, im Marketing um 123,2, im Projektmanagement um 117,1, im Bank- und Finanzwesen um 100,9, im Kundenservice um 99 Prozent. Personalchefs stellen im Jahr 2026 nicht in der Breite ein, sondern gezielt jene Kräfte, welche die Produktivität durch KI-Anwendung heben.

Interessant könnte auch der Wechsel in komplementäre Rollen sein. Mit dem Inkrafttreten des EU-AI-Acts entstehen viele neue Stellen in KI-Governance, KI-Compliance, Datenschutz und KI-Ethik. Der IAPP-Salary-Report nennt

für die Doppelqualifikation Datenschutz und KI-Governance ein mittleres Jahresgehalt von rund 170.000 Dollar. Diese Funktionen verlangen juristisches, organisatorisches und ethisches Urteilsvermögen, weniger Programmierung.

HOLGER SCHMIDT



DIGITALWIRTSCHAFT

Alles Wichtige zu Künstlicher Intelligenz, Plattform-Ökonomie und Digitalisierung finden Sie in unserem Premium-Newsletter und auf FAZ.NET.

