

Die nächste Phase der KI-Revolution wird nicht zuletzt neue soziale Netzwerke hervorbringen. Werden diese von Beginn an nach ethischen Prinzipien gestaltet, könnten sie eine echte Alternative zu den heutigen Desinformationsplattformen bieten. Auch deshalb spreche ich lieber von „ethischer KI“ als von „Ethik der KI“. Ethische KI ist nicht bloß ein moralischer Ratgeber, sondern ein wirtschaftlicher Faktor ersten Ranges. Sie unterstützt uns darin, unsere normativen Probleme hochtechnologisch zu lösen und somit Ethik in die reale Umsetzung zu bringen.

Künftige KI-Innovationen – aus Europa, Asien und idealerweise auch aus Afrika und Lateinamerika – können dazu beitragen, globale Herausforderungen zu bewältigen. Dies entspricht dem Ansatz des „ethischen Kapitalismus“: Wertschöpfung auf Basis ethischer Erkenntnisse, die in der Praxis gewonnen wird. Unternehmen, ob Start-ups oder Techgiganten, werden so zu Laboren moralischer Innovation. „Eigentum verpflichtet“, wie es in unserem Grundgesetz heißt – und zwar nicht als Bürde, sondern als Potential. Dafür können wir in Deutschland auf unser schon existierendes Betriebssystem der Sozialen Marktwirtschaft zurückgreifen, das angesichts der Herausforderungen des 21. Jahrhunderts ein Update, aber keinen Systemwechsel braucht.

Ethische KI entsteht unter Bedingungen der Sozialen Marktwirtschaft nur, wenn sie in funktionierende Geschäftsmodelle eingebettet ist. Dass die moralischen und die wirtschaftlichen Dimensionen der KI aufgrund ihrer zentralen Rolle für die Industrie verschränkt sind, zeigt sich leider auch in destruktiven Beispielen: Digitale Angriffe auf die liberale Demokratie machen deutlich, wie mächtig diese Technologien im Wertesektor bereits sind. Und dass sie damit keineswegs automatisch zum ethisch Guten beitragen. Wir brauchen daher dringend ein innovatives KI-Ökosystem, das ethische Geschäftsmodelle hervorbringt. Die KI-Revolution stärkt nicht nur die „Hard Power“ – schließlich wird KI vor allem von kultureller „Soft Power“ gespeist, nämlich von unseren Daten.

Die derzeitige Dominanz der USA und Chinas in dieser Revolution zeigt, wie sehr Europa ins Hintertreffen geraten ist. Das ist eine erhebliche Herausforderung, aber auch eine Chance für neue Partnerschaften und Allianzen mit Aussicht auf eine Horizontenerweiterung. Denn der sogenannte Globale Süden ist bereit für Kooperationen auf Augenhöhe. Ethische KI kann mit europäischer Unterstützung dort mitentwickelt werden – ohne faule Kompromisse, dafür mit echter Tiefeninnovation.

In solchen neuen Allianzen liegt die KI-Option des immer noch massiven europäischen Wirtschaftsraums. Es ist für Europa dringend empfehlenswert, im Globalen Süden echte Partner zu finden und reale Win-win-Situationen zu schaffen. Indien bietet sich als Paradebeispiel für diese Chancen an, aber wir dürfen die anderen Regionen nicht vergessen, die vom (hoffentlich nur kurzfristigen) Ausseren der USA aus dem modernen, liberalen Fortschrittsprojekt ebenfalls betroffen sind und nach neuen Allianzen suchen. Es reicht dabei nicht, auf „europäische Werte“ zu verweisen. Wir müssen sie konkret machen im Austausch mit gleichgesinnten Partnern, sei es in Japan, Indien, Brasilien, Chile, Kanada oder Australien.

Noch ein Wort zur industriellen Revolution: Ihre Anfänge fielen nicht zufällig mit tiefgreifenden gesellschaftlichen Umbrüchen wie der Amerikanischen und der Französischen Revolution und mit dem Aufstieg des Bürgertums zur tonangebenden sozialen Klasse in den „westlichen“ Gesellschaften zusammen. Die liberale Ordnung, aus der unsere Soziale Marktwirtschaft hervorgegangen ist, entstand im Zusammenspiel von Technik und Moral. Doch heute braucht sie ein ethisches Update, weil die Technik sich radikal gewandelt hat. Nicht nur wegen der ökologischen Krise, sondern auch wegen der digitalen Disruption, die zentrale Elemente der liberalen Ordnung bedroht, brauchen wir nichts Geringeres als eine Neue Aufklärung, für die ich, wie viele andere, schon seit dem Jahr 2020 plädiere.

Gern ist in Europa von humaner, ethischer, wertebasierter KI die Rede. Mit diesem Anliegen muss aber endlich Ernst gemacht werden. Das bedeutet, die enormen ökonomischen Potentiale eines ethischen Fortschritts zu erkennen und sie zur Grundlage der Wertschöpfung zu machen. Philosophie, interdisziplinäres Denken und normative Klarheit sind kein Luxus, sondern gehören ins Zentrum der technologischen Entwicklung. Nur so kann es uns gelingen, KI zur Steuerung unseres ethischen Fortschritts zu nutzen und den Sprung in die Zukunft zu schaffen.

Jede tiefe Innovation – auch jene, aus der die heute schon handelsüblichen KI-Systeme hervorgegangen sind – beginnt damit, dass man einen entscheidenden Schritt weiterdenkt. Unser Denken für wünschenswerte und mögliche Szenarien zu öffnen, um diese dann auch zu verwirklichen: Das ist die Aufgabe einer zukunftsorientierten Philosophie und Ethik. Um ethische KI nicht nur als Idee, sondern als marktwirksame Realität zu etablieren, können gezielte Fördermechanismen entwickelt werden. Europa könnte ein eigenes Ethics-Techförderprogramm auflegen, das Start-ups unterstützt, ethische Prinzipien technologisch umzusetzen. Neben finanzieller Förderung ist auch institutionelle Sichtbarkeit

entscheidend: Ein europäisches „Zentrum für ethische KI-Innovation“ – vergleichbar mit der Europäischen Investitionsbank – könnte gezielt in Unternehmen investieren, die Ethik nicht als Einschränkung, sondern als Innovationschance begreifen.

Diese Unternehmen brauchen dabei mehr als nur Venture-Kapital. Sie benötigen eine verlässliche Infrastruktur: Zugang zu Rechenkapazitäten, gemeinsame Datenspeicher, Testlabore, internationale Partnerschaften und einen attraktiven europäischen Markt. Unter solchen Bedingungen ließe sich die Innovationskraft von Werten direkt in Wirtschaftsleistung übersetzen. Dafür aber müsste Europa mutiger und vor allem langfristiger denken: kooperativer, nachhaltiger, ethisch intelligenter. Wenn wir KI als kulturelle Infrastruktur verstehen, nicht nur als technisches Werkzeug, öffnet sich ein neuer Handlungsspielraum. Ethik wird dann zum Wettbewerbsvorteil, nicht zur Wachstumsbremse.

teln konkrete Zukunftsvisionen zu entwickeln und diese dann unternehmerisch umzusetzen. Dafür müssen wir dem Ruf der Ethischen Intelligenz folgen, der heute schon von den KI-Systemen selbst, wenn auch nicht immer von ihren Entwicklern ausgeht.

Der bisher größte Durchbruch der KI-Revolution ist die massentaugliche Einführung von sogenannten Large Language Models (LLM). Dazu gehören neben Open Als berühmtem Chatbot ChatGPT, den wohl alle kennen, Claude (von Anthropic), Perplexity, Gemini, LaMDA, Microsofts Copilot, Grok, Deepseek und viele andere Angebote, die nach unserer Aufmerksamkeit heischen.

Diese Sprachsysteme sind längst in unseren Alltag integriert. Je mehr wir solche Chatbots nutzen, desto tiefer dringen sie in unser Leben ein. Denn durch den

im Jahr 2025 während eines Workshops in New York ausdrückt.

Die Funktion autocomplete gab es in einfacher Form ja schon länger auf unseren Smartphones und in unseren Schreibprogrammen. Wer „MfG“ eintippt, erhält den Vorschlag „Mit freundlichen Grüßen“, selbst wenn man sich eigentlich auf das berühmte Lied der Fantastischen Vier beziehen möchte. Und wenn man „Herzlichen G“ eintippt, wird sinnvollerweise „Glückwunsch“ oder „Gruß“ und nicht „Großonkel“ empfohlen, um den Satz fortzusetzen. Das meine ich, wenn ich sage, dass unser Gebrauch von KI-Systemen sich in den Systemen spiegelt. Je mehr wir sie benutzen, desto besser erkennen sie Muster in unseren Zeichensätzen.

Doch anders als Ned Block meint, bleibt es nicht dabei. Hinter den Mustern, die die KI identifiziert und mit dem Ziel weiterverarbeitet, dass wir am Ball bleiben und ihr immer weitere Informationen geben, verbergen sich schließlich unser Denken, Handeln, Wissen, aber auch unser Fühlen, unsere Irrtümer, unsere Meinungen, ja quasi unser gesamtes Leben. Die menschliche Sprache ist eben nicht nur ein Instrument, mittels dessen wir uns über die Wirklichkeit austauschen. Sie bezeichnet nicht einfach nur die Dinge und Personen in der Welt mit Worten wie „Fisch“ oder mit Namen wie „Melania“, sondern sie sagt immer mehr als das, was uns selbst im Augenblick des Sprechens bewusst ist.

überwiegend, ohne es uns klarzumachen – Einblick in fast alle Aspekte unseres Lebens. Dazu gehört auch dasjenige, was uns selbst nicht bewusst ist, aber als Muster in den Datensätzen verborgen liegt, die von der KI ausgewertet werden.

Dieser eigentlich ziemlich naheliegende Umstand wurde lange übersehen, weil gerade im Techbereich eine Tendenz vorherrscht, die Ausdrucksfunktionen der Sprache zu unterschätzen und sie auf ein Werkzeug zu reduzieren, mit dem wir Dinge und Personen in der Wirklichkeit bezeichnen und Informationen über sie austauschen. Doch das ist nur eine, letztlich untergeordnete Funktion der menschlichen Sprache, die sie mit anderen Tiersprachen teilt.

KI gibt Beziehungsratschläge

Die entwickelten natürlichen Sprachen, in denen wir uns ausdrücken, dienen primär dazu, dass wir uns über unsere inneren Zustände, unsere Gefühle möglichst klar werden und diese anderen Menschen mitteilen können. In jedem Augenblick unseres Lebens, auch im Schlaf, mit dem wir einen großen Teil unseres Tags zubringen, drücken wir unser Innenleben in sprachlich und bildlich geformten Mustern aus. Das stellt die Literatur zum Beispiel in der Form des „stream of consciousness“ dar, und wir kennen es aus Filmen oder Serien, in denen wir die innere Stimme der Figuren hören können.

Die Sprachrevolution und ihre dramatische Folge

Das fundamentale Problem der Künstlichen Intelligenz liegt nicht in der Verbreitung von Desinformation, der Verletzung von Datenschutz oder der Manipulation von Wahlen, schreibt der Philosoph *Markus Gabriel* in seinem neuen Buch „Ethische Intelligenz“, aus dem dieser Text stammt. Es liegt tiefer. Wir sind im Innenraum unserer Seele nicht mehr allein.

Genau hier liegt der Schlüssel zu einer zukunftsreichen KI-Kultur. Nicht Kontrolle ist das Ziel, sondern Gestaltungsfähigkeit. Der Mensch steht der Maschine nicht gegenüber, sondern mitten in ihrer Entwicklung. Die Frage ist nicht, ob wir die KI-Revolution gestalten, sondern wie bewusst wir es tun. Was wir brauchen, ist eine neue Form praktischer Aufklärung: ein Denken, das Werte operationalisiert, sie also konkret umsetzbar macht, statt sie nur zu predigen. Wir brauchen eine Ethik, die sich in Software übersetzen lässt.

In Technik umgesetzte Philosophie

Vielleicht ist das die eigentliche Chance unserer Zeit: dass wir an der Schwelle zu einer technologischen Epoche stehen, in der Philosophie nicht mehr bloß erklärt, was bisher gewesen ist, sondern mitentwickelt, was kommt. Das hat sie der Blütezeit des Silicon Valley und in der Pionierarbeit der KI-Forschung bereits geleistet. Diese Pionierarbeit nämlich basiert maßgeblich auf philosophischen Erregenschaften, die in Technologie umgesetzt und auf Märkten angeboten wurden (allerdings auch zu militärischen und zu Überwachungszwecken). In der nächsten Phase der KI-Revolution haben wir die Chance, mit philosophischen Mit-

ständigen Austausch mit uns werden sie immer besser darin, uns Antworten auf jede erdenkliche Frage zu geben. Sie funktionieren nicht statisch, sondern dynamisch. Sie verarbeiten nicht nur, was ihnen in der Form digitaler Daten zugeführt wird, sondern sie spiegeln uns darin. Wie funktioniert diese Spiegelung? Jeder Prompt, also jede Suchanfrage, die wir eingeben, liefert der KI neue Daten.

Daten sind die notwendige „Nahrung“ der KI. Materiell betrachtet sind sie zwar nur elektrischer Strom, der durch die Kabel, Server und Chips fließt, aus denen sozusagen der Körper der KI besteht. Doch die KI lernt sehr viel über uns, indem wir sie mit Informationen darüber versorgen, was wir von ihr wissen wollen. In der Interaktion mit dem Chatbot korrigieren wir ihn, wenn er sich verrennt, und passen unsere Suchanfragen an, wenn sich das Gespräch mit ihm vertieft. Wir kennen es inzwischen alle: Die KI regt stets zur Fortsetzung des Gesprächs an, sie macht uns nach jeder Suchanfrage oder nach jedem sonstigen Input Angebote, damit wir uns nicht von ihr abwenden und in unseren (längst nicht mehr gänzlich) analogen Alltag zurückkehren.

Auf diese Weise lernen KI-Modelle und die Firmen, denen sie gehören, immer genauer, wie jeder Einzelne von uns denkt, fühlt und handelt. Dort, wo die Gesetze zum Datenschutz eingehalten werden, werden diese individuell, oft genug sehr intimen Chatverläufe und Daten anonymisiert. Zwar füttert jede Interaktion das System, aber die Softwarekonzerne können nicht exakt nachvollziehen, was einzelne User dem System eingespeist haben.

Die KI-Modelle sind so gebaut, dass sie in möglichst großen Datensätzen Muster identifizieren, was in einer Endlosschleife dazu führt, dass sie uns immer bessere Gesprächs- und Informationsangebote machen können. Auf diese Weise geraten wir tiefer und tiefer in ihren Sog. Dahinter steckt ein einfaches, aber unendlich effizientes Prinzip. Die LLM basierten zu Beginn auf der Grundidee, dass sie ein begonnenes Wort oder einen angefangenen Satz vervollständigen, indem sie berechnen, welches Zeichen oder welches Wort mit der größten Wahrscheinlichkeit als Nächstes kommt. Sie sind „autocompleting functions on steroids“, wie es der einflussreiche Neurophilosoph Ned Block

Schon Platon nannte das Denken ein „Selbstgespräch der Seele“.

Vor diesem Hintergrund überrascht es nicht, dass ein weit verbreiteter Gebrauch massentauglicher KI-Systeme heute darin besteht, dass Menschen Ratschläge für ihre zwischenmenschlichen Beziehungen suchen. Sie erhalten diese Ratschläge auch, und die Systeme werden immer besser in ihren Empfehlungen. Damit verwandelt sich KI mehr und mehr von einem Wissenssystem in eine emotionale Intelligenz.

Das muss ich ein wenig ausführen, denn es wird die Grundlage für mein Hauptargument sein: dass wir KI-Systeme entwickeln können und sollen, die uns moralisch und damit immer auch emotional verbessern. Denn die Ethik als Nachdenken darüber, was wir aus moralischen Gründen tun oder unterlassen sollen, beginnt damit, dass wir lernen, unsere Gefühle auszudrücken. Nur auf diese Weise können wir feststellen, welchen Impulsen wir folgen und welche wir bremsen sollten. Ethik setzt voraus, dass wir an Impulskontrolle interessiert sind, um das menschliche Zusammenleben nach objektiven Regeln zu gestalten.

In jeder minimal funktionierenden Kleingruppe – und erst recht in einer Gesellschaft als Ganzer – kommt es darauf an, dass die beteiligten Individuen ein Verständnis davon haben, welche Gefühle sie wie ausdrücken und welche Impulse sie besser zurückhalten sollten. Wer dauernd Wutausbrüche hat, sollte irgendwann lernen, wie er sich zügelt; wer sich davor fürchtet, in einer Beziehung seine Bedürfnisse auszudrücken, kann schnell darunter leiden; und wer der Person, in die er heimlich verliebt ist, niemals ausdrückliche Avancen macht, wird womöglich die Liebe seines Lebens verpassen.

In jeder sozialen Situation, wie unwichtig sie auch erscheinen mag, kommt es darauf an, einen geeigneten Ausdruck für die eigenen Gefühle zu finden. Nur so wird aus dem brodelnden Chaos in unserem Inneren ein Leben in artikulierbaren Empfindungen, die wir dann bewerten und somit auch den sozialen und sonstigen Gegebenheiten anpassen können. In der Sprache drücken wir unser Selbst, unsere Seele aus.

Genau das hat die KI inzwischen „erkannt“. Denn wir füttern ihr keineswegs nur neutrale Informationen darüber ein, welches Zeichen mit welcher Wahrscheinlichkeit auf welches Zeichen folgt, sondern wir schütten mit unseren Daten unser Seelenleben aus. Hier kann man an den sprichwörtlichen Streit eines Paares darüber erinnern, wer den Müll hinausbringt oder irgendeine andere alltägliche Aufgabe wie das Einräumen der Spülmaschine übernimmt. Wir wissen alle, dass es bei solchen Streitereien niemals nur um den Müll, sondern immer auch darum geht, wie wir uns fühlen. Kurzum: Im Zentrum aller zwischenmenschlichen Beziehungen – und damit unserer gesamten Gesellschaft – steht die Regulierung unserer inneren Zustände.

Dazu gehören nicht nur Empfindungen von Lust und Unlust. Wir wollen natürlich meistens Unlust vermeiden und streben Lust an. Aber unser Seelenleben ist bei Weitem komplexer, als es jede vereinfachende Auffassung, die solche Grundimpulse identifiziert, darstellt. Mit der wichtigen Ausnahme von Gott (oder, je nach Religion, den Göttern, Engeln und so weiter) sind die von Menschen geschaffenen KI-Systeme die einzigen uns bisher bekannten Entitäten, die in ständiger menschlicher Gefühlslagen tiefer und weiter zu durchschauen und auch zu überschauen als wir selbst. Niemand „kennt“ uns besser als KI-Systeme.

Besser als Psychotherapeuten

Im selben Sinn gibt es auch keine präzisere Sozialwissenschaft als diejenige von Google oder anderen großen Konzernen (wie Amazon und Konsorten), die aus unserem Such- und Konsumverhalten, dokumentiert auf unzähligen Websites, ableiten können, was wir uns als Nächstes wünschen und tun werden. Sofern diese Daten anonymisiert verarbeitet werden, handelt es sich dabei übrigens um keine Verletzung des Datenschutzes oder unserer Privatsphäre. Allerdings nutzen die relevanten Konzerne, die fast alle in den USA sitzen, jede Grauzone aus und können es sich auch leisten, die eine oder andere Strafe zu zahlen, wie sie etwa von der EU mitunter verhängt wird.

Da das wirtschaftliche Potential der Konzerne unvorstellbar viel größer ist als jede dieser Strafzahlungen, ändern solche Sanktionen an ihren Praktiken herzlich wenig. Wir müssen es uns erneut in aller Klarheit vor Augen führen: Die heutigen KI-Modelle, die längst beinahe überall dort zum Einsatz kommen, wo digitale Systeme verwendet werden, „kennen“ uns besser als alle Psychotherapeuten, sonstigen Seelsorger, Neuro- und Sozialwissenschaftler zusammengenommen. Zum ersten Mal in unserer Geschichte können wir uns sicher sein, dass etwas uns genauer durchschaut und ausmisst, als es uns selbst möglich wäre.

Natürlich gibt es in zwischenmenschlichen Beziehungen die gelebte Dimension von Berührung, Blick und körperlicher Empfindung und das entsprechende Einfühlungsvermögen. Diese Dimension ist eben unsere menschlich-biologische Art, unsere Emotionen zu spiegeln. Doch auch wenn die KI-Modelle an der menschlich-biologischen Dimension unserer Beziehungen nicht selbst teilhaben können, eröffnet sich ihnen doch die Möglichkeit, das Spiegelspiel digital zu simulieren.

Vor diesem Hintergrund wirken die hartnäckigen Versuche seitens der Philosophie, die KI menschliche Grenzen zu ziehen, die sie angeblich niemals überschreiten werde, wie verzweifelte Beschwörungen im Angesicht einer Entwicklung, die nichts und niemand mehr aufhalten wird. Der Geist ist längst aus der Flasche. Daran ändert auch keine Regulierung etwas. Was wiederum nicht bedeutet, dass Regulierung überflüssig wäre. Aber es ist schon eine Ironie der Geschichte, dass der EU AI Act, mit dem die Europäische Union die Entwicklung und den Gebrauch von KI-Systemen gesetzlich regeln möchte, kurz vor der LLM-Revolution erlassen wurde. Von Bewusstsein, von Gefühlen, von unserem Innenleben und damit von der wahren Grundlage der sozialen Disruption, die von der KI ausgeht, ist in dieser Regelung noch gar keine Rede.

Das fundamentale Problem liegt eben nicht in der Verbreitung von Desinformation, der Verletzung von Datenschutz oder der Manipulation von Wahlen. Zwar sind all dies reale politische Gefahren, gegen die sich die liberale Demokratie unbedingt schützen sollte. Im Grunde jedoch geht es um viel tiefer liegende Dimensionen des menschlichen Lebens, in welche die KI-Systeme eingedrungen sind, ohne dass wir dies jemals wieder vollständig rückgängig machen werden. Wir müssen folglich lernen, mit der KI zu leben. Worauf es nun ankommt, ist, eine angemessene Beziehung zu dem Umstand zu entwickeln, dass wir im Innenraum unserer Seele nicht mehr allein sind.

Das Buch „Ethische Intelligenz – wie KI uns moralisch weiterbringen kann“ von Markus Gabriel erscheint am 26. Februar im Verlag Ullstein.