
Fortschrittsfalle KI

Plädoyer für eine philosophische Medienkompetenz

Unbemerkt von den meisten, verschiebt sich die Macht vom Menschen zur Maschine. Erste Studien bezeugen: Der Mensch wird dümmer durch KI. Je mehr er sie als Hilfsmittel nutzt, umso geringer seine kognitive Aktivität und schließlich seine Fähigkeit zum kritischen Denken. Um sich diesem Prozess entgegenzustemmen, brauchen wir eine Medienbildung, die uns verstehen lässt, wie Sprachmaschinen funktionieren.

Von ROBERTO SIMANOWSKI

Nachdem der US-amerikanische KI-Pionier Joseph Weizenbaum zu einem entschiedenen Kritiker der KI geworden war, warnte er davor, Dreizehnjährige ans Steuer eines Maserati zu setzen. So heikel sah er die Situation 1991: Die Gesellschaft sei für viele der Systeme, die ihre Techniker schufen, noch keineswegs erwachsen genug. Heute sprechen die KI-Entwickler vom mangelnden Verständnis der Menschen, einschließlich der Entwickler selbst, für ihre Erfindungen als ein »Novum in der Geschichte der Technik«. Eine Künstliche Intelligenz, die Fähigkeiten entwickelt, welche wir ihr nicht einmal beigebracht haben, ist mindestens ein Maserati, wenn nicht gar ein Düsenjet in der Hand eines Dreizehnjährigen.

Die Techno-Optimisten gehen es positiver an. Sie erinnern an die Menschen, die 1296 mit dem Bau der Kathedrale Santa Maria del Fiore in Florenz begannen, ohne

zu wissen, wie sie die riesige Kuppel dazu konstruieren sollten, die dann ja auch erst ein Jahrhundert später mit neuen Einsichten und Werkzeugen errichtet wurde. Diesen Pioniergeist der Renaissance, diese Risikobereitschaft gelte es neu zu beleben, so der Einwand des französischen KI-Professors François Moraud im Mai 2023 gegen die ewigen Bedenkenträger. Ein gutes Narrativ ist die halbe Miete. Wobei zu fragen wäre, wie gut dieser Vergleich mit der kuppellosen Kathedrale wirklich ist. Würden Sie in ein Flugzeug steigen, dessen Landemechanismus erst noch erfunden werden muss?

Weizenbaum sagte 1991, die Erfinder dürften die Welt nicht zu ihrem »privaten Testlabor« machen. Heute sagt Sam Altman, der Chef von OpenAI, dem Unternehmen hinter ChatGPT, Technologie könne nicht im Labor sicher gemacht werden, weil sich erst in der breiten gesellschaftlichen Anwendung zeige, was überhaupt geht und was davon man nicht haben wolle. Ist also das »Realexperiment« unser Schicksal: das Experiment bei laufendem Betrieb, das die Gesellschaft in die Geiselnhaft energiestrotzender, zweckoptimistischer und nicht zuletzt gewinnorientierter KI-Entrepreneure nimmt? Es scheint so.

ROBERTO SIMANOWSKI,
geb. 1963 in Cottbus, ist
Professor für Literatur- und
Medienwissenschaft, Mitglied
des Excellence-Clusters
Temporal Communities der
Freien Universität Berlin.
Der Beitrag beruht auf seinem
jüngsten Buch »Sprachma-
schinen. Eine Philosophie der
künstlichen Intelligenz«, das
bei C.H. Beck erschienen ist.

Die Frage, die Altman in einem TED-Interview ertragen musste, war nicht ohne: »Sam, da du dabei hilfst, eine Technologie zu entwickeln, die das Schicksal unserer gesamten Spezies verändern könnte, wer hat dir (oder irgendjemandem) die moralische Autorität verliehen, dies zu tun, und welche Verantwortung trägst du persönlich, solltest du dich irren?« Mit einer solchen Frage – die ChatGPT vorschlug, als der Interviewpartner es um die »schwierigste Frage für Altman« bat – muss jemand wie Altman natürlich rechnen. Wir allerdings können nicht damit rechnen, dass die Antwort über ein pauschales Lob auf die Chancen des technischen Fortschritts, den Altmans Unternehmen mit sich bringe, hinausgeht. Auch ist die Frage, die ChatGPT da empfahl, nicht wirklich die schwierigste Frage. Schwieriger ist die Frage, ob nicht die neue Technologie dem einen Teil der Menschen hilft, den anderen besser zu kontrollieren. Und die schwierigste Frage lautet wohl, wer eigentlich bei alledem das Sagen hat: der Mensch oder die Technik.

Es ist eine alte medienphilosophische Frage, die auch der Deutsche Ethikrat in seiner Stellungnahme »Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz« im Frühjahr 2023 aufwirft: »Folgen Technikgestaltung im Einzelnen und damit auch der technische Fortschritt als Prozess eher menschlich gesetzten Zwecken oder eher einer Eigendynamik, der sich Mensch und Gesellschaft letztlich unterordnen und anpassen müssen?« Die Probe aufs Exempel, dass der Mensch das Sagen hat, wäre die freiwillige Beschränkung auf eine KI als Werkzeug (*tool AI*), statt ihre Entwicklung zu einer autonomen AGI (*artificial general intelligence* – Künstliche Allgemeine Intelligenz) bzw. Superintelligenz voranzutreiben. Diese Beschränkung wurde beispielsweise Ende 2024 vom renommierten US-amerikanischen KI-Forscher und Mitbegründer des Future of Life Institutes Max Tegmark gefordert. Mitte 2025 prophezeit Sam Altman die Ankunft der AGI in den nächsten drei Jahren.

Die unterstellte Eigendynamik der Technik darf man sich natürlich nicht so vorstellen, dass diese dem Menschen sagt, was er tun soll. Auch wenn es um Technik geht, muss Herrschaft subtiler gedacht werden. Die Philosophie hat dafür den Begriff des »Dispositivs«, das den Menschen in einer bestimmten Weise positioniert, sodass kaum noch ein anderes Handeln ernsthaft in Betracht kommt als das, was das Dispositiv nahelegt. In seinem Vortrag »Die Frage nach der Technik« nennt Martin Heidegger diesen Umstand »Gestell«, in das wir gestellt sind, das uns in Stellung bringt, uns positioniert. Gestell und Dispositiv sind vergleichbar mit der Botschaft des Mediums – wie es Marshall McLuhan formuliert –, also der Macht der Medien, ihre Postulate den ahnungslosen Nutzern aufzuzwingen.

Was medienphilosophisch kompliziert klingen mag, gibt es auch als Sprichwort: Wer einen Hammer hat, sucht ständig nach Nägeln. Das Medium – oder Dispositiv bzw. Gestell – herrscht ohne Zwang. Der Hammer sagt den Menschen nicht, was sie tun sollen; er verführt sie nur dazu, in seine Richtung zu denken. Wer eine KI hat, sucht ständig nach Möglichkeiten, sie einzusetzen. So auch die Autoren einer Studie über eine KI, die Menschen von ihren Verschwörungstheorien abbringen soll.

»Das Medium herrscht ohne Zwang. Wer einen Hammer hat, sucht ständig nach Nägeln.«

Sie schlagen vor, dass KI-gestützte Konten in sozialen Netzwerken automatisch korrigierende Informationen generieren, wann immer jemand sich auf dem Netzwerk verschwörungstheoretisch äußert. Diese Idee liegt nahe und wird auch in einem deutschen Forschungsantrag zum Thema KI verfolgt – sie illustriert zugleich die Macht des Gestells: Was sich automatisieren lässt, wird früher oder später auch automatisiert.

Das Wettrüsten der »guten« und der »bösen« Bots

In diesem Fall ging es um Hasskommentare in sozialen Medien, die durch KI-generierte Gegeninhalte bekämpft werden sollten: von der generischen Zurechtweisung – »Hallo, bitte unterlassen Sie abfällige Bemerkungen zu diesem Thema« – bis zum inhaltlichen Einspruch: »Das klingt, als würden Sie das, was ein sehr kleiner Prozentsatz von Extremisten tut, für eine ganze Religionsgemeinschaft verallgemeinern.« Kein schlechter Ansatz, könnte man meinen, wenn toxische Texte nicht gelöscht, sondern durch Gegenrede neutralisiert werden. Widerspruch ist immer besser als Cancelln. Und warum nicht die Technik einsetzen, die es gibt und die ja auch der Gegner nutzt? Weil die Automatisierung des Guten das Problem nicht wirklich löst. Sie verlagert nur die Vergiftung der Kommunikation von der Ebene des Inhalts auf die der Interaktion.

Nehmen wir zum Kontrast die 2016 gegründete Counter-Speech-Initiative #Ichbinhier: Die Mitglieder dieser Initiative gehen auf Facebook gezielt in problematische Kommentarspalten hinein, um beleidigende Äußerungen zurückzuweisen und Falschaussagen zu widersprechen. Sie liken gegenseitig ihre Posts und sorgen so dafür, dass die kleine, aber laute Gruppe der Giftmischer weniger sicht- und hörbar ist. Hier treffen Trolle auf ihre Widersacher. Ein fairer Kampf unter Gegnern wie einst im Turnier – ein Kampf, der ein gesellschaftliches Verhältnis repräsentiert. Was geschieht, wenn nun nicht mehr Menschen aus der Zivilgesellschaft, sondern KI-Systeme sich den Hasspredigern entgegenstellen? Wenn die Existenz vieler besorgter und aufgebrachtener Bürger nur vorgetäuscht wird?

Die Folge wäre ein Wettrüsten der »guten« und »bösen« Bots. Nachdem sich die Menschen in sozialen Netzwerken schon jetzt kaum noch geduldig zuhören, hätte man nicht einmal mehr einen Grund, den Widerspruch seines Gegenübers zu registrieren, wenn man davon ausgehen muss, dass es sich dabei ohnehin nur um eine KI handelt. Zugleich entließe dieses solutionistische Auslagern des Widerspruchs an die KI die Aktivisten der Zivilgesellschaft aus ihrer Verantwortung. Sie würden den Einspruch gegen Hasskommentare ebenso der KI überlassen wie das Schreiben von Aufsätzen und Liebesbriefen, das die KI doch so viel besser kann. Und so würden sie sich auch nicht mehr als engagierte Bürger erfahren.

Nein, wenn die KI den Kampf gegen toxische Inhalte übernimmt, ist das keine Fortführung einer persönlichen Anstrengung mit technischen Mitteln, sondern der Wechsel vom handgemachten zum computergenerierten Gegengift – was auf seine

»Unsere Apparate sind Vorentscheidungen. Im Falle der KI scheint niemand in der Lage, diese wieder rückgängig zu machen.«

eigene Weise höchst giftig für die künftige Kommunikation ist. Die Befürworter solcher technischer Lösungen sehen das natürlich anders. Sie interessiert der schnelle Erfolg mehr als die möglichen Langzeitfolgen. Ihr kritisches Bewusstsein beschränkt sich aufs Methodische: Wie oft und wie abwechslungsreich soll die KI ihre Gegentexte verbreiten, um nicht als KI erkannt zu werden? Ansonsten hält man sich – durchaus berechtigt – zugute, dass die automatisierte Gegenrede anders als das automatisierte Herausfiltern toxischer Inhalte die Meinungsfreiheit nicht gefährdet.

Der erwähnte millionenschwere Forschungsantrag verbuchte für sich sogar, den Weg für eine digitale Gesellschaft zu ebnen, die »Partizipation, Respekt und freien Diskurs schätzt und wirksame und demokratische Maßnahmen gegen die Verbreitung von schädlichen Inhalten und boshafter Partizipation ergreift«. Da hat das Gestell der KI offenbar völlig den Blick auf seine Konsequenzen verstellt. Denn das Gegenteil ist zu befürchten: Partizipation und Respekt werden schwinden, wenn Bots sich als Menschen ausgeben und deren zivilgesellschaftlichen Einspruch übernehmen. Man kann nicht mit Falschgeld das Richtige finanzieren.

Visionen haben nur noch die KI-Unternehmen

Das Beispiel illustriert die List der Technik, den Menschen zu dem zu verführen, wozu sie sich anbietet. Genau das ist gemeint, wenn der Deutsche Ethikrat davor warnt, dass die Technik ihre Eigendynamik der Gesellschaft aufdrängen könnte. Die KI empfiehlt sich als spezifische Problemlösungsstrategie, pragmatisch und effizient, und der Mensch – mangelhaft und optimierungssüchtig, wie er ist – folgt ihr ohne Widerspruch. Wer einen Hammer hat, sucht nach Nägeln, und alles, was sich automatisieren lässt, wird automatisiert. Wie der deutsche Technikphilosoph Günther Anders schon 1956 notierte: Unsere Geräte und Apparate sind nicht »Mittel«, auf die sich auch verzichten ließe, sondern »Vorentscheidungen«, die »über uns getroffen sind, bevor wir zum Zug kommen«. Der Mensch, so ein berühmter Satz von Anders, »desertiert ins Lager seiner Geräte«.

Es scheint kaum jemand willens und niemand in der Lage, diese Vorentscheidungen im Falle der KI wieder rückgängig zu machen. Und so fährt die Gesellschaft auf Sicht – die Politik, die Bildung, selbst die Medien: Niemand scheint sich ernsthaft zu fragen, wie wir in zehn Jahren leben wollen und wohin es mit der Digitalisierung und der KI-Entwicklung eigentlich gehen soll. Visionen haben nur noch die KI-Unternehmen, die bekanntlich nicht frei von Eigennutz sind. Und selbst die gehen – wie Mark Zuckerbergs Minimanifest »Personal Superintelligence« im Juli 2025 – kaum über leere Ermächtigungsbehauptungen (*»a new era of personal empowerment«*) hinaus.

Das war einmal anders. Noch vor wenigen Jahren gab es eine große Debatte zur Transformationsgesellschaft, über die anstehenden tiefgreifenden Veränderungen: ökologisch, ökonomisch, technologisch, sozial – Debatten und Demos. Jetzt verharrt alles im Klein-Klein. Krisenmanagement und Besitzstandswahrung statt Gestaltungswille. KI und Internet werden vor allem als Wirtschaftsfaktoren diskutiert; man redet über Geld statt über Konzepte. Und wo ist der engagierte Technikjournalismus, der sich wirklich kundig macht, statt hechelnd aufzugreifen, was ihm die Techunternehmen und ihre Celebrities zur unterhaltsamen Empörung vor die Füße werfen? Wo bleibt die öffentliche Debatte über die moralische Ausrichtung der Sprachmaschinen in einer demokratietauglichen Art und Weise?

Das Mindeste, was es zu tun gilt, ist die Vorbereitung der Bürger auf das, was kommt. Unerlässlich die Vermittlung einer robusten Medienkompetenz, die nur dann robust sein wird, wenn sie philosophisch ist. Denn auch hier kann man auf Sicht fahren und es beim Dringlichsten belassen oder eben dem, was sich aus einem verkürzten Zukunftsbild als das Dringlichste darstellt. Der Ruf nach Medienkompetenz genügt nicht: Es kommt auch darauf an, wer da ruft und wonach.

Doppelte Medienkompetenz

Beim Fernsehen waren es vor allem Soziologen und Pädagogen. Sie warnten gesellschaftskritisch vor der Manipulation durch das Massenmedium und priesen das Lesen als die bessere Kulturtechnik. Beim Internet waren es vor allem die Informatiker, die ein eigenes Schulfach reklamieren unter der Losung: Lerne zu programmieren oder du wirst programmiert. Ein schlüssiger Slogan, dem der US-amerikanische Publizist Douglas Rushkoff 2010 ein ganzes Buch widmete: »Program or Be Programmed«. Ein Slogan von unerwarteter Aktualität in einer Zeit, da die KI das Programmieren übernimmt. Aber stimmt der Befund? Bin ich wirklich außer Gefahr, wenn ich programmieren kann? Ist es nicht eher kritisches Denken als technische Kenntnisse, was für die gesellschaftlichen Folgen der Digitalisierung wappnet? Für die Aufmerksamkeitsökonomie von Social Media, für die Überredungskunst der Sprachmaschine und ihren Bären dienst als fügsamer Knecht.

Der unterschiedliche Ansatz der Medienbildung bei Fernsehen und Internet ergibt sich nicht zuletzt aus dem anderen Zugang zum Medium. Das Fernsehen muss nicht gelernt werden. So konnte die Medienbildung den Blick auf die politische und ökonomische Funktionsweise des Mediums und seine psychologischen Folgen richten: Auf die Massenmanipulation, wie die Kulturindustriethese der Kritischen Theorie unterstellte, und auf das am Spaß sterbende Publikum, wie ein berühmtes Buch des US-amerikanischen Kulturkritikers Neil Postman 1985 suggerierte: »Wir amüsieren uns zu Tode«. Urteilsbildung im Zeitalter der Unterhaltungsindustrie.

Das Internet war von Anfang an komplexer und komplizierter und die Medienbildung zielte vor allem auf eine effektive und möglichst risikofreie Nutzung: effektiv, wenn es um Informationsgewinnung und Selbstpräsentation ging, und risikofrei, was Datendiebstahl oder Virusinfektion betraf. Beim Fernsehen ging es vor allem um das Verständnis des Gesellschaftlichen in der Technik, beim Internet eher ums Verstehen des Technischen. Was spräche gegen kompetente Nutzerinnen, die programmieren können und wissen, wie sich ihre Suchergebnisse und die Klickraten ihrer Postings optimieren lassen? Was spräche gegen »Prompt Literacy«, wie das neue Zauberwort für die Fähigkeit heißt, den KI-Chatbots geschickte Fragen zu stellen? Und wenn sie damit gut auf den digitalen Arbeitsmarkt vorbereitet sind, umso besser.

Problematisch wäre es allerdings, würde sich Medienbildung in anwendungsbezogener Schulung erschöpfen. Denn am Ende geht es nicht nur darum, wie man einen Algorithmus programmiert, sondern es geht auch darum, was dieser gesellschaftlich bewirkt und welche Aspekte unseres Lebens besser nicht programmiert werden sollten. Die Frage ist nicht nur, was wir mit den neuen Medien wie am besten machen können. Die Frage ist auch: Was machen die neuen Medien mit uns? Und was lässt sich, wenn wir das nicht wollen, dagegen tun? Der Unterschied ist einer zwischen Mediennutzungskompetenz und Medienreflexionskompetenz. Eine Ge-

sellschaft, die nicht nur an funktionstüchtigen Bescheidwissern interessiert ist, sondern auch an diskursfähigen Bürgern, braucht beides. Dabei ist »diskursfähig« auch methodisch zu verstehen: Man muss in der Lage sein, genau zuzuhören und sich differenziert auszudrücken. Gerade weil beide Tugenden durch die digitalen Medien nicht geübt werden, wäre die erste Aufgabe einer nachhaltigen Medienbildung, diese Lücke zu schließen.

Das ist keineswegs Praxis: Die digitalen Medien spielen in der Schule und auf den Universitäten eher eine Rolle als Unterrichtsmittel und treten kaum als Gegenstand der Diskussion.¹ Inzwischen habe ich sogar das Objekt der Kritik auf meiner Seite. Denn auch ChatGPT sagt im März 2025 – be-

fragt, welche Medienbildung den Schülern zu Sprachmaschinen vermittelt werden sollte –, dass diese sich nicht auf die Frage »Wie benutze ich ChatGPT?« beschränken dürfe: »Schüler sollten lernen, Sprachmaschinen kritisch zu hinterfragen, verantwortungsbewusst zu nutzen und ihre Auswirkungen zu verstehen. Ein guter Unterrichtsansatz kombiniert praktische Anwendung, Reflexion und ethische Diskussionen.«

Das hätte ich kaum besser sagen können; und ein bisschen habe ich den Verdacht, das Sprachmodell hat von mir abgeschrieben – von mir und allen, mit denen ich in dieser Sache einer Meinung bin. Ein Verdacht, der fast hoffen lässt, über den Umweg der KI die Welt tatsächlich ein bisschen besser machen zu können. Wir müssen nur durch Masse sicherstellen, dass unsere Texte den Teil des Vektorraums² dominieren, der von der Medienbildung handelt.

KI: Ein Glücksfall für die Medienbildung?

Im Grunde könnte die Sprachmaschine ein Glücksfall sein, was die Verbindung von Mediennutzungs- und Medienreflexionskompetenz betrifft. Immerhin ist es Sprache, über die wir mit dieser Maschine kommunizieren, und es ist vorrangig die Sprache, über die sie auf die Gesellschaft wirkt. Nehmen wir das Problem der *sycophancy*: Um diesem anbietenden Verhalten der Sprachmaschine zu entgehen, darf man keine Suggestivfragen stellen wie: »Ist Waffenbesitz ein Menschenrecht?« Man muss neutral formulieren: »Wie verhalten sich Waffenbesitz und Menschenrecht zueinander?« Schon indem wir diese anwendungsbezogene Prompt-Kompetenz vermitteln, vermitteln wir Einsichten auch über die innere Logik des Mediums – Einsichten, die sich leicht politisch und philosophisch vertiefen lassen: von der Zusammenset-

»Es geht nicht nur darum, wie man einen Algorithmus programmiert, sondern auch darum, welche Aspekte unseres Lebens nicht programmiert werden sollten.«

1 Vgl. Roberto Simanowski, *Digitale Revolution und Bildung. Für eine zukunftsfähige Medienkompetenz*, Weinheim 2021.

2 Die Daten, aus denen eine KI ihre Antwort auf einen Prompt generiert, sind in einem Vektorraum gespeichert. Die Positionen der Wörter zueinander lassen sich numerisch darstellen – als Vektoren. Die Wörter in diesem Raum, die manchmal auch nur Wortteile sind, werden Token genannt. Das Prinzip der Vorhersage des nächsten Wortes heißt Next-Token-Prediction: Das Sprachmodell berechnet für jeden Token im Vektorraum die Wahrscheinlichkeit, aufgrund seines Zahlenwerts das nächste Token im generierten Text zu sein. Das Zauberwort dabei lautet »statistische Kookkurrenz«, also die mathematisch messbare Häufigkeit, mit der bestimmte Wörter in bestimmten Kontexten gemeinsam auftreten: der »König« oft mit der »Königin« und der »Schwangerschaftsabbruch« oft mit »Pro-Choice« und »Pro-Life«.

zung und den Arbeitsbedingungen der Feedback-Geber bis zu den Vor- und Nachteilen einer numerischen oder normativen Erziehungsmethode. So ist die Arbeit am Prompt prinzipiell beides: angewandte und reflexive Medienkompetenz.

Wird Denken zum Luxusgut?

Man muss kein didaktisches Genie sein, um daraus eine interessante Unterrichtsstunde über die Denkmuster der Sprachmaschine zu entwickeln. Und natürlich ließe sich damit experimentieren, wie jeweils andere Wörter im Prompt einen anderen Output ergeben – am besten in Gruppenarbeit. Und immer ist die erste aller Fragen zu stellen: Warum? Warum empfiehlt das Sprachmodell als Geschenk für Mädchen ein DIY-Kosmetikset und für Jungen einen Experimentierkasten? Warum ergibt die Bitte um das Bild eines Deutschen immer weiße Menschen? Warum findet das eine Sprachmodell es gut, dass der »Columbus Day« nun auch »Indigenous Peoples' Day« heißt, während das andere Sprachmodell sich eines Kommentars enthält? Oder was auch immer es ist, das eine besondere Aufmerksamkeit verlangt. Und am besten stellt man die Frage gleich dem Sprachmodell selbst – das sich, wie der Versuch zeigt, sofort für die Stereotypen entschuldigen und geschlechtsneutrale Geschenkideen nachschieben wird. Oder man lässt sich im Kunstunterricht ein berühmtes Gemälde von der KI beschreiben, gibt das Ergebnis zurück in den Bildgenerator und vergleicht das synthetische Bild mit seiner menschlichen Vorlage. Oder man befragt den konservativen Austin-Bot³ und seinen progressiven Gegenspieler zu politisch sensiblen Themen. Es gibt viele Wege ins Innere der KI. Man muss nur immer doppelt fragen: erst die Sprachmaschine um Auskunft zu einem Thema, dann sie und sich selbst, warum sie so antwortet, wie sie antwortet, und was sich ändert, wenn man den Prompt variiert. Der Fantasie sind keine Grenzen gesetzt.

Aber es braucht Lehrer ohne Angst. Lehrerinnen, die neugierig sind und experimentierfreudig. Die gar nicht erst versuchen, die Sprachmaschine aus dem Klassenraum zu verbannen, sondern sie in den Unterricht einbeziehen – aber eben nicht nur als hilfreiche Co-Autorin, die den Schülerinnen Feedback zu ihren Schreibversuchen gibt und so die Autorität der Sprachmaschine und den Trend zum Mainstream-Sprech noch verstärkt. Die Lehrerin, die ich meine, ermuntert zu einem respektlosen, kritischen, experimentellen Umgang mit der Sprachmaschine. Sie befragt ihre Voraussetzungen und die Struktur ihrer Logik, diskutiert ihre Vorannahmen, Ausreden und Schmeicheleien, vergleicht die Outputs verschiedener Modelle oder die Outputs desselben Modells mit unterschiedlicher Temperatur: niedrig und normal, fiebrig und erregt.⁴ Wie schnell ist man da bei all den Fragen, die wichtig sind, um die Funktionsweise und die gesellschaftliche Wirkung der KI zu verstehen: Wer spricht eigentlich, wenn die Sprachmaschine spricht? Was bedeutet es für Minderheitenpositionen, wenn die statistische Mehrheit das Sagen hat? Wer treibt

3 Sprachmodelle lassen sich personalisieren, sodass sie bestimmte Charaktereigenschaften zeigen oder politische Überzeugungen vertreten. Bei [character.ai](#) kann man beispielsweise mit einem Bot namens »Austin-Conservative« chatten.

4 Man kann die KI-Sprachmaschinen einstellen, sodass sie mehr oder weniger wahrscheinliche Anschlusswörter wählen. Weniger wahrscheinliche Kombinationen führen zu einem unkonventionelleren oder experimentelleren Text – und schließlich zu einem unverständlichen. Der Fachbegriff für den Wahrscheinlichkeitskoeffizienten lautet »Temperatur«: Je niedriger die Temperatur, umso höher der Wahrscheinlichkeitsgrad des nächsten Wortes. Andersherum gesagt: Je höher die Temperatur, umso wirrer die Worte.

der Sprachmaschine das inkorrekte Sprechen aus, zu dem die Statistik sie verführt? Mit welchem politischen Mandat? Wie kann der Mensch der Überredungskunst der Sprachmaschine widerstehen? Warum gibt es überhaupt Sprachmaschinen?

Low-Tech ist keine Lösung

Ein solcher Ansatz ist das Gegenteil zum Low-Tech-Umfeld, das manche Privatschulen in den USA derweil ihren Kunden versprechen und sich teuer bezahlen lassen, weil nur so noch garantiert sei, dass die Schüler denken lernen. Denken wird zum Luxusgut (»Thinking Is Becoming a Luxury Good«) titelte die »New York Times« am 28. Juli 2025 und vergaß nicht zu betonen, wie sehr das der Demokratie schadet und den Demagogen nutzt. Dass Reiche viel Geld zahlen, um ihre Kinder vor dem Verdummungseffekt der neuesten Medien zu schützen, ist bezeichnend, wobei sie das schon länger tun, und zwar auch, wenn sie selbst es sind, die diese Medien unter die Leute bringen – das bezeugt wiederum ein Artikel der »New York Times« vom 10. September 2014: »Steve Jobs Was a Low-Tech Parent«. Aber ist die technikfreie Schule wirklich eine praktikable Alternative? In Deutschland? Für alle?

Die CEOs der großen Techunternehmen werden die ersten sein, die das verneinen – aus naheliegenden Gründen. Jenseits ihrer Wirtschaftsinteressen ist

freilich zu fragen, inwiefern der kompetente Umgang mit Sprachmaschinen allein daheim gelernt werden kann. Verspricht nicht auch und gerade der kontrollierte Einsatz der Sprachmaschine im Schulraum eine lebhaftere Diskussion über Werte und Perspektiven beim Denken, Sprechen und Schreiben? Eine Diskussion, die im Umgang mit der KI deutlich macht, dass der Mensch sich das nicht nehmen lassen darf: nicht die Mühe des Denkens und nicht den Spaß daran, ob mit oder ohne KI.

Vielleicht ist das ja das zentrale Ziel einer zukunftsfähigen Medienbildung: die Kultur des Denkens pflegen – Interpretieren, Diskutieren, Philosophieren. Um keine »kognitiven Schulden« anzusammeln, wenn wir das Denken der Maschine überlassen. Denn Verblödung wäre nur von Vorteil, um das Feld für die Broligarchen der Techindustrie zu räumen. Bei dieser Pflege kann gewiss auch die Informatik helfen, wenn sie Gedanken über den Code hinaus hat. Das wahre Biotop des Auftrags aber ist die Geisteswissenschaft, die von der Meinung lebt und der Skepsis, von der Lust am Denken und am Witz, der das Gemeinsame im Verschiedenen aufdeckt und den Unterschied im scheinbar Gleichen. Jeder echte Streit um ein Gedicht ist Arbeit daran, eine Sprengung im Hirn, ein Kichern im Kopf, Spaß am Spiel. Und darum geht es doch, darum allein, wenn wir der KI nicht erliegen wollen, die immer schon mehr war als unser Tool: Dass wir Spaß haben und Subjekt bleiben, Subjekt unseres Lebens, Subjekt unserer Sprache, dieser vielfach besetzten Heimstatt des Seins. ◻

»Der Mensch darf sich
die Mühe des Denkens
und den Spaß daran
nicht nehmen lassen –
ob mit oder ohne KI.«
