

MARGINALIEN

*Wer die Sprachmodelle beherrscht,
beherrscht auch die Politik*

Von Hannes Bajohr

Die Welt Künstlicher Intelligenz denkt groß und zugleich simpel, und zwar von Anfang an. Als im Sommer 1956 am Dartmouth College jener Workshop stattfand, der den Begriff und das Feld der Artificial Intelligence ins Leben rief, lautete die selbstgesetzte Aufgabe, herauszufinden, »wie man Maschinen dazu bringen kann, Sprache zu verwenden, Abstraktionen und Begriffe zu formen, Probleme zu lösen, die bisher Menschen vorbehalten sind, und sich selbst zu verbessern«. ¹ Angesetzte Arbeitszeit: zwei Monate.

Fast siebzig Jahre später – am 22. März 2023 – erschien auf der Website des Future of Life Institute ein Offener Brief, der zum Zeitpunkt der Niederschrift dieses Artikels um die dreißigtausend Unterschriften zählte. ² Darunter befinden sich Persönlichkeiten wie Elon Musk und viele renommierte KI-Forscher, die ein Moratorium für die Entwicklung großer Sprachmodelle (*Large Language Models*, LLMs) für mindestens sechs Monate fordern.

- 1 John McCarthy/Marvin L. Minsky u. a., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. In: *AI Magazine*, Nr. 27/4, 2006.
- 2 *Pause Giant AI Experiments: An Open Letter*. In: *Future of Life Institute* vom 22. März 2023 (futureoflife.org/open-letter/pause-giant-ai-experiments/).

Systeme wie ChatGPT seien inzwischen zu mächtig und zu gefährlich geworden, und es bestünden »fundamentale Risiken für die Gesellschaft und die Menschheit«, die von »auf Menschenniveau agierender KI« ausgingen. Bis man sich nicht darauf geeinigt habe, wie das zu regulieren sei, sollten alle KI-Labore auf weitere Forschung verzichten.

Unterschätzte man in Dartmouth noch spektakulär, als wie schwierig sich die Automatisierung von Intelligenz herausstellen würde, zieht der Offene Brief mit ähnlich viel Bombast die falschen Konsequenzen aus den Möglichkeiten gegenwärtiger Sprachtechnologie.

Denn erstens ist auch heute das Ziel von Dartmouth nicht erreicht – allen Erfolgen zum Trotz agieren ChatGPT und Co. nicht »auf Menschenniveau«. ³ Solche Fantasien sind Teil des Hype um KI – eine Tendenz, die die Branche seit ihren Anfängen plagt und die am Ende vor allem den Firmen dient, die sie entwickeln. ⁴ Was beweist die Macht eines Entwicklers

- 3 Melanie Mitchell/David C. Krakauer, *The Debate Over Understanding in AI's Large Language Models*. In: *Proceedings of the National Academy of Sciences*, Nr. 120/13 vom 21. März 2023; Luciano Floridi, *AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models*. In: *Philosophy & Technology*, Nr. 36, 2023.
- 4 Dieuwertje Luitse/Wiebke Denkena, *The Great Transformer: Examining the Role of Large Language Models in the Political Economy of AI*. In: *Big Data & Society*, Nr. 8/2, Juli 2021.

mehr als seine Fähigkeit, ein Produkt zu vertreiben, das potentiell die Welt zerstören könnte? Insider spekulierten ohnehin bald, hier gehe es darum, die in der Branche lange geltende Regel offener Forschung zu unterlaufen und in den sechs Monaten im Geheimen weiterzuarbeiten. Und tatsächlich kündigte Musk am 14. April 2023 die Gründung seines eigenen KI-Unternehmens mit dem Namen X.AI an; seine Unterschrift schien schon bald nicht mehr viel zu gelten.

Zweitens, und viel wichtiger, spricht aus dem Brief auch ein katastrophales Verständnis vom Zusammenwirken von Technik und Politik, und zwar sowohl was ihre Gefahren wie was deren Bekämpfung betrifft. Während die Befürchtung, KI-generierter Text könnte Informationskanäle mit Unwahrheiten und »propaganda-as-a-service« überschwemmen,⁵ durchaus berechtigt ist, ist der Brief ansonsten von apokalyptischen Fantasien über die völlige Ersetzung des Menschen durch Maschinen und den »Kontrollverlust über unsere Zivilisation« getragen.

Das sind die Sorgen sogenannter Longtermists – einer libertären, transhumanistischen und ultra-utilitaristischen Denkschule, die möglichen zukünftigen Menschen ein ungleich größeres moralisches Gewicht einräumt als realen gegenwärtigen.⁶ Ihre Vertreter, denen auch Musk

nahesteht, denken in Jahrtausenden und in Begriffen einer reinen Nutzenmaximierung. Aus diesem Grund beunruhigt sie die Bedrohung durch hyperintelligente Maschinen – »Wir werden alle sterben«,⁷ heißt es in einem besonders alarmistischen Meinungsartikel – viel mehr als beispielsweise die unmittelbaren Schäden des Klimawandels, der sozialen Ungerechtigkeit oder der globalen Armut – Probleme, die für sie entweder überhaupt keine sind oder durch eben jene KI gelöst werden, die sie als existentielle Bedrohung wahrnehmen.

Was aber mit großen Sprachmodellen wie ChatGPT tatsächlich auf uns zukommt, ist keine *technische* Katastrophe bössartiger Computer. Viel konkreter drohen Sprachmodelle ein *demokratisches* Desaster zu werden – durch die Privatisierung von Sprachtechnologien als zukünftigem Ort politischer Öffentlichkeit und den rein technokratischen Ansatz zur Lösung dieser Probleme. Genau an dieser Stelle kommen Politik und Zivilgesellschaft ins Spiel, und auch die Demokratietheorie ist hier gefragt.

Die technische Entwicklung der letzten Jahre hat gezeigt, dass ein KI-System umso leistungsfähiger wird, mit je mehr Daten man es füttert – aber umso teurer ist es auch in der Entwicklung. Es ist zwar schwierig, künftige Trends vorherzusagen, es ist aber nicht ausgeschlossen, dass der Wettbewerb um immer größere Modelle dazu führen könnte, dass nur noch eine Handvoll Unternehmen im Rennen bleibt,

5 Eugene Bagdasaryan/Vitaly Shmatikov, *Spinning Language Models: Risks of Propaganda-As-A-Service and Countermeasures*. In: 2022 IEEE Symposium on Security and Privacy (arxiv.org/abs/2107.10443).

6 Emile P. Torres, *Against Longtermism*. In: *Aeon* vom 19. Oktober 2021 (aeon.co/essays/why-longtermism-is-the-worlds-most-dangerous-secular-credo).

7 Eliezer Yudkowsky, *Pausing AI Developments Isn't Enough: We Need to Shut it All Down*. In: *Time* vom 29. März 2023 (time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/).

wie OpenAI/Microsoft, Googles Deepmind oder Anthropic.⁸ Und obwohl es Open-Source-Bestrebungen gibt, Sprachmodelle zu demokratisieren, müssen sie sich im Vergleich zu Big Tech erst noch als erfolgreich erweisen (schließlich werden die meisten Menschen nicht ihre eigenen LLMs auf ihren Heimcomputern trainieren, sondern sich auf das fertig verpackte und gewartete Produkt eines großen Unternehmens verlassen). Und zumindest im Moment spielen kleinere nichtkommerzielle Unternehmungen und Universitäten beim Erreichen immer neuer Größenrekorde so gut wie keine Rolle mehr.⁹

8 James Vincent, *AI Is Entering an Era of Corporate Control*. In: *The Verge* vom 3. April 2023 (www.theverge.com/23667752/ai-progress-2023-report-stanford-corporate-control).

9 Zum Zeitpunkt der Abfassung dieses Essays gibt es zumindest einige Gerüchte, dass sich die großen Akteure über Open-Source-Modelle Sorgen machen. Es ist noch schwer zu sagen, ob diese Bedenken gerechtfertigt sind. Es ist zwar richtig, dass HuggingChat oder LLaMA einige ermutigende Ergebnisse hervorgebracht haben, aber es gibt zwei Tatsachen, die einen skeptisch gegenüber einer goldenen Open-Source-Zukunft machen sollten: Die erste ist, dass dieses LLM von der Firma Meta geleakt wurde und somit nur in zweiter Linie ein Produkt der freien Entwicklung ist, sondern vielmehr auf den Schultern von Riesen entstanden, die letztendlich allein in der Lage sind, die notwendigen technischen Grundlagen zu schaffen. Zweitens sind »foundation models« zunehmend Teil einer Lizenzwirtschaft, in der die Verantwortung und die Wartung für das zugrundeliegende Modell zum Verkaufsargument werden, das große Anbieter gegenüber Open-Source-Anbietern bevorzugt. In der Anhörung des US-Senats zur KI-Regulierung am 16. Mai 2023 stimmte Sam Altman, CEO von OpenAI, zu, dass »es eine relativ kleine Anzahl von Anbietern geben wird, die Modelle herstellen

Zwei Probleme stehen hier im Vordergrund. Was uns möglicherweise bevorsteht, und das ist das erste Problem, ist ein Oligopol, das die Sprachtechnologien in den Händen einiger weniger Privatunternehmen konzentriert. Diese mächtigen Akteure üben nicht die Vorherrschaft über ein beliebiges Produkt aus. Vielmehr wird, um es dramatisch zu formulieren, die Zukunft der politischen Meinungsbildung und Deliberation in LLMs entschieden. Das stellt eine direkte Herausforderung für die demokratische politische Theorie dar.

Warum das so ist, lässt sich anhand dessen zeigen, was bisher als größtes politisches Problem von KI-Systemen galt: ihre »Biases«.¹⁰ Sprachmodelle bilden ihre Ausgaben den Texten nach, auf die sie trainiert wurden und die größtenteils aus Dokumenten bestehen, die im Internet und anderen Quellen zu finden sind – einschließlich der Vorurteile, Rassismen und Sexismen, die einen Großteil dieser Inhalte ausmachen. Aber da »Rohdaten« ein Oxymoron ist¹¹ – Daten stehen immer in einem bestimmten Kontext und sind zu bestimmten Zwecken gedacht, und sie werden von den Werkzeugen und Systemen geformt, die zu ihrer Erzeugung, Er-

können«, was sich jedoch positiv auf eine wirksame Regulierung auswirken würde:

»Es gibt wenige von uns, die man wirklich im Auge behalten muss [...] das hat Vorteile.

10 Emily M. Bender/Timnit Gebru u. a., *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (dl.acm.org/doi/10.1145/3442188.3445922).

11 Lisa Gitelman (Hrsg.), *»Raw data« is an Oxymoron*. Cambridge/Mass.: MIT Press 2013.

fassung und Darstellung verwendet werden – und da die Realität selbst einen »world bias« aufweist¹² – Ungleichheiten in der Gesellschaft werden in Daten, die eine neutrale Haltung darstellen sollen, schlicht wiederholt und verstärkt –, sind Sprachmodelle von Natur aus ideologisch, selbst in ihrem scheinbar unschuldigen Ausgangszustand. Dies gilt auch dann, wenn ihre Schöpfer keine bewusste oder böswillige Absicht verfolgen.

Alle Versuche, ein neutrales Ergebnis zu erzielen – LLMs sozusagen zu entideologisieren –, sind immer vergeblich, unabhängig davon, an welchem Ende des Prozesses angesetzt wird. So kann man entweder den *Output* zensieren, wie es (bis zu einem gewissen Grad) mit ChatGPT und seiner Implementierung in Bing geschieht (mit dem Risiko, das Programm unbrauchbar zu machen). Oder man kann, wie ebenfalls praktiziert, den *Input* – den Datensatz – durchforsten, um unerwünschte Bestandteile zu entfernen, bevor das Training beginnt.¹³ Sowohl die Filterung der Ergebnisse als auch die Kuratierung des Datensatzes laufen darauf hinaus, das Modell auf der Grundlage einer Norm, einer Vorstellung von einer besseren Welt, zu korrigieren, was aber eine eminent politische Entscheidung ist. Die Entideologisierung der KI beinhaltet notwendigerweise die Formulierung einer gesellschaftlichen Vision – und ist damit ebenfalls ideologisch.

12 Matteo Pasquinelli, *How a Machine Learns and Fails – A Grammar of Error for Artificial Intelligence*. In: *spheres*, Nr. 5 vom 20. November 2019.

13 Erin Miller/Roudabeh Kishi u.a., *An Agenda for Addressing Bias in Conflict Data*. In: *Scientific Data*, Nr. 9/1, September 2022.

Weil GPT etwa tendenziell eher progressive Werte vertritt, haben konservative Medien sich schnell über »woke« KI ereifert.¹⁴ In Wirklichkeit aber stehen eher PR-Erwägungen hinter dieser Kuratierung: Sexistische Beleidigungen, politische Extrempositionen oder rassistische Outputs wirken sich schlicht negativ auf die Gewinnmarge der Tech-Unternehmen aus. Aber selbst wenn man davon ausgeht, dass hinter OpenAIs Entscheidungen echte Überzeugung steht, sind weder die unkuratierten noch die kuratierten Versionen von LLMs wertfrei. KI ist immer ideologisch.¹⁵

Aus diesem Grund sollte es uns beunruhigen, dass Entscheidungen über den gesellschaftlichen Entwurf, den Sprachmodelle artikulieren, in den Händen einiger weniger Unternehmen liegen, die keiner demokratischen Kontrolle unterstehen und niemandem gegenüber rechenschaftspflichtig sind außer ihren Anteilseignern. Sie werden damit, um einen Begriff der Philosophin Elizabeth Anderson zu zweckentfremden, zu »privater Re-

14 Diese Wendung wurde auch von Musk als Grund für die Gründung seines eigenen KI-Unternehmens angeführt.

15 Hannes Bajohr, *Wer sind wir? Warum künstliche Intelligenz immer ideologisch ist*. In: *Republik* vom 6. April 2021 (www.republik.ch/2021/04/06/warum-kuenstliche-intelligenz-immer-ideologisch-ist); Leif Weatherby, *ChatGPT Is an Ideology Machine*. In: *Jacobin* vom 17. April 2023 (jacobin.com/2023/04/chatgpt-ai-language-models-ideology-media-production). Vgl. auch Paola Lopez, *Artificial Intelligence und die normative Kraft des Faktischen*. In: *Merkur*, Nr. 863, April 2021 (www.merkur-zeitschrift.de/artikel/artificial-intelligence-und-die-normative-kraft-des-faktischen-a-mr-75-4-42/).

gierung«. ¹⁶ Auf den ersten Blick mag dies nicht ganz so neu klingen. »Artifacts have politics«, wie Langdon Winner es formulierte, ¹⁷ und das gilt auch für digitale Artefakte: Allein durch die Art und Weise, wie es Informationen zugänglich macht, hat Google bereits einen übergroßen Einfluss darauf, was Nutzern als Realität erscheint. ¹⁸

Doch mit der neuentstehenden privaten Regierung, die maschinelles Lernen kapitalisiert, hat eine noch tiefere Vereinnahmung eingesetzt. Denn das Produkt der KI-Unternehmen ist exakt jene Ressource, die eine lebendige Demokratie ausmacht: Sprache. Es ist die Sprache, durch die wir politische Alternativen auf der einzigen Ebene aushandeln, auf der das möglich ist, jener der politischen Öffentlichkeit. Statt zu debattieren, in welcher Welt wir leben wollen, ist mit LLMs ein Teil dieser Entscheidung schon getroffen, bevor ein einziges Wort gesprochen ist, weil die Sprache selbst schon einer politischen Vorentscheidung unterstand. Je mehr die Sprache, die von solchen Modellen produ-

ziert wird, zukünftig die feinsten Verästelungen unseres Alltagslebens durchdringt, desto düsterer muss eine solche Perspektive erscheinen.

Da hilft es wenig, dass man ein Sprachmodell logischerweise auch nach rechts dirigieren kann, wie der Informatiker David Rozado gezeigt hat, als er kürzlich RightWingGPT schuf. ¹⁹ Dies verweist auf eine zweite Sorge aus demokratietheoretischer Sicht. Eine Zukunft, in der es neben einer progressiven auch eine konservative Sprach-KI gäbe, würde nicht zu einer Art Gleichgewicht oder einer Entwicklung hin zu immer nuancierteren Positionen führen. Ebenso wenig würde eine Vielzahl per LLMs vertretene Fraktionen eine gesunde »Vielfalt der Sekten« darstellen, um James Madisons republikanisches Allheilmittel aus den *Federalist Papers No. 10* (1787) gegen die Dominanz einer Gruppe über eine andere zu zitieren.

Denn es wären nicht länger Sekten oder Fraktionen, die miteinander sprechen, sondern stochastisch modellierte Rede selbst. Politische LLMs würden so die Diskussion zwischen gesellschaftlichen Gruppen eliminieren, deren Konflikte idealerweise zur Meinungsbildung einer informierten Öffentlichkeit beitragen. Statt Austausch gäbe es nur die Verstärkung bereits bestehender Meinungen; und anders als bei den vielbeschworenen Echokammern sozialer Medien, wären es hier nicht einmal mehr Menschen, die die Parameter dieser Diskussion festlegten. An ihre Stelle träte ein komplexes System

16 Elizabeth Anderson, *Private Government. How Employers Rule Our Lives (and Why We Don't Talk about It)*. Princeton University Press 2017. Ich sage »zweckentfremden«, weil Anderson auf die Macht der Unternehmen verwies, das Leben ihrer Mitarbeiter zu regulieren, wenn sie beispielsweise Normen für deren Rede auch in ihrer privaten Zeit festlegen. Bei LLMs muss man nicht für ein Unternehmen arbeiten, um betroffen zu sein. Nichtsdestotrotz haben die Begrenzung und Umlenkung von Sprache eine ähnlich quasi-regierungshafte Qualität.

17 Langdon Winner, *Do Artifacts Have Politics?* In: *Daedalus*, Nr. 109/1, Winter 1980.

18 Safiya Umoja Noble, *Algorithms of Oppression. How Search Engines Reinforce Racism*. New York University Press 2018.

19 David Rozado, *RightWingGPT – An AI Manifesting the Opposite Political Biases of ChatGPT*. In: *Rozado's Visual Analytics* vom 16. Februar 2023 (davidrozado.substack.com/p/rightwinggpt).

aus maschineller Sprachverarbeitung und profitorientierten Privatunternehmen.

Die nachteiligen Auswirkungen speziell auf demokratische Politik lassen sich anhand eines Arguments der politischen Philosophin Judith Shklar veranschaulichen. Sie hielt es für die wichtigste Aufgabe liberaler Demokratien, ihre öffentlichen Institutionen und Foren so zu gestalten, dass die Stimmen der Marginalisierten gehört werden können. Dies sei nicht nur ein moralisches Gebot, sondern ein demokratisches: Da es sich bei den hegemonialen Vorstellungen von Gerechtigkeit, Gleichheit, Fairness usw. um positive Konzepte handele, die sich auf Fälle beschränken, die von diesen Konzepten ausdrücklich berücksichtigt werden – und so auf Fälle, die sich die Mehrheit als relevant vorstellen kann –, ist es unerlässlich, auch negativen Erkenntnissen Gehör zu schenken, also Fällen, die durch die Riten der offiziellen Begriffsmatrix fallen. Manchmal wird etwas gar nicht als Ungerechtigkeit wahrgenommen, bis es von den Betroffenen als solche benannt wird.

Der »Sinn für Ungerechtigkeit«,²⁰ den die Marginalisierten artikulieren – das unmittelbare Gefühl der Verletzung, das allen explizit formulierten Konzepten von Gerechtigkeit vorausgeht –, ist nicht nur ein Ruf nach konkreter Linderung; er bringt auch das demokratische Projekt voran, indem es den Begriff dessen ausweitet, was als mögliche Ungerechtigkeit erkannt werden kann. Die Marginalisierten anzuhören bedeutet, sie als Bürger ein-

zubeziehen und somit mehr Menschen im Gemeinwesen zu vertreten.

Miranda Fricker hat »epistemische Ungerechtigkeit« als das Phänomen bezeichnet,²¹ bei dem Bürger den Rahmen dessen, was Gerechtigkeit in einer gegebenen Demokratie ausmacht, nicht verändern können, weil sie ungehört bleiben und nicht berücksichtigt werden. Die hegemoniale Sichtweise, die LLMs formulieren und kodieren, lässt diese epistemische Korrektur, die sich aus dem Hören auf den Sinn für Ungerechtigkeit ergibt, nicht zu. LLMs sind damit bereits ihrer Struktur nach Maschinen epistemischer Ungerechtigkeit.

Dies wird zu einem ernststen Problem, wenn LLMs langfristig selbst zu einem Surrogat der Öffentlichkeit, also einer synthetischen Öffentlichkeit werden. Werden immer mehr Texte durch KI-Systeme generiert – und davon ist auszugehen –, sinkt der Anteil an menschenproduzierten Diskursbeiträgen stetig. Da Sprachmodelle, einmal trainiert, schwer zu verändern sind und zudem aus Fakten der Vergangenheit Normen für die Zukunft ableiten,²² besteht die Gefahr eines *value lock*, wie Emily Bender es nennt. Meinungen, Werte, Normen und Tendenzen, die ansonsten durch Diskurs oder ideologische Auseinandersetzung verändert werden können – einschließlich

20 Judith N. Shklar, *Über Ungerechtigkeit. Erkundungen zu einem moralischen Gefühl*. Übersetzt und herausgegeben von Hannes Bajohr. Berlin: Matthes & Seitz 2021.

21 Miranda Fricker, *Epistemische Ungerechtigkeit. Macht und die Ethik des Wissens*. München: Beck 2023.

22 Cathy O'Neil, *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. New York: Crown 2016; Virginia Eubanks, *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St Martin's Press 2018.

des artikulierten Unrechtsempfindens von Minderheiten –, werden festgeschrieben aufgrund der Unfähigkeit des Systems, sich anzupassen.

Keine noch so große Diskussion und kein noch so großer Hegemoniekampf können diese eingebrannten Werte verändern; das Ergebnis ist eine technologisch bedingte politische Stagnation, die einen immer engeren Erkenntnishorizont einschließt. Der Inhalt der Beiträge zu einer solchen synthetischen Öffentlichkeit ist also nicht nur durch technologische Systeme und Kapitalinteressen vorbestimmt, sondern er bleibt auch zwangsläufig derselbe, da er parallel zu und unbeeinflusst von nicht-LLM-generiertem öffentlichem Diskurs verläuft.

Für diese Überlegungen spielt es keine Rolle, ob man einem deliberativen oder einem agonistischen Ansatz der Demokratietheorie folgt. Aus der Sicht deliberativer Theorie – die Politik als Prozess kollektiver Argumentation unter Bürgern versteht – hängt das »Erfordernis der freien Deliberation« von der »diskursiven Qualität der Beiträge« in der Öffentlichkeit und der möglichen Einbeziehung aller Bürger ab.²³ In einer synthetischen Öffentlichkeit, die durch ideologisch grundierte LLMs im *value lock* erzeugt wird, stünde beides auf dem Spiel. Vom Standpunkt einer agonistischen politischen Theorie aus gesehen – die Politik als einen Bereich des Kampfes und der Anfechtungen betrachtet, der nicht unbedingt zu einem eindeutigen Konsens führt – würde eine solche synthetische Öffentlichkeit »für widerstreitende

Stimmen legitime [...] Artikulationsmöglichkeiten« eliminieren, die den »Kampf zwischen unvereinbaren hegemonialen Projekten« in die Kanäle der demokratischen Agonistik übertragen sollen.²⁴

Auch dies muss nicht unbedingt das Ergebnis böswilliger Absicht sein. Ein gutes Beispiel dafür, wie der technische Geist des »fixing things with technology« die demokratische Debatte, egal, ob deliberativ oder agonistisch, abwürgen kann, bietet das Paper *AI Chat Assistants Can Improve Conversations About Divisive Topics*.²⁵ Die Ingenieure versuchten, »die Qualität von kontroversen Gesprächen zu verbessern«, indem sie ein KI-System zwischen den Dialog zweier debattierender Parteien schalteten, das ihre Positionen in einer »neutraleren« Sprache umformulierte. Die Autoren sind der Meinung, dass »die Verbesserung der Qualität des politischen Diskurses« – durch einen erzwungenen zivilen Ton – »einen größeren Nutzen für den sozialen Zusammenhalt und die Demokratie haben« werde.

Diese Lösung ist jedoch nicht nur inhärent paternalistisch, da sie die bewussten kommunikativen Entscheidungen der Debatteilnehmer nicht respektiert; sie ist auch, auf einen größeren Maßstab projiziert – zum Beispiel in Messengern oder Diskussionsforen integriert – ganz und gar nicht neutral, sondern wiederum das Ergebnis einer vorherigen Entscheidung

23 Jürgen Habermas, *Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik*. Berlin: Suhrkamp 2022.

24 Chantal Mouffe, *Über das Politische. Wider die kosmopolitische Illusion*. Aus dem Englischen von Niels Neumeier. Frankfurt: Suhrkamp 2007.

25 Lisa P. Argyle/Ethan Busby u. a., *AI Chat Assistants Can Improve Conversations About Divisive Topics*. 2023 (arxiv.org/abs/2302.07268).

darüber, was Neutralität und Höflichkeit sind, wie also die Grenzen und Bedingungen des Diskurses aussehen sollen.²⁶ Und auch dies ist eine zutiefst politische Entscheidung.

Die Regulierung von KI – die Big Tech vorgeblich fordert, aber nur als freiwillige Selbstkontrolle der Industrie, als »regulatory capture« akzeptieren will²⁷ – kann

26 Das ist keine hypothetische Sorge. Vgl. Maurice Jakesch / Advait Bhat u. a., *Co-Writing with Opinionated Language Models Affects Users' Views*. CHI Conference on Human Factors in Computing Systems, Hamburg 2023.

27 Der ehemalige Google-CEO Eric Schmidt hat dies in einem Interview mit NBC im Vorfeld der KI-Anhörungen des US-Senats (in dem der CEO von OpenAI, Sam Altman, den Senatoren einen wirtschaftsfreundlichen Regulierungsrahmen vorschlug) unmissverständlich zugegeben. Schmidt plädierte nicht nur für eine Selbstregulierung durch eine Business-to-Law-Pipeline – also ein rechtliches Privileg –, sondern beanspruchte für seine Branche auch die alleinige Kompetenz, die Feinheiten dessen, was reguliert werden soll, zu erfassen: also ein epistemisches Privileg. Praktischerweise deckt sich die Herrschaft der Experten mit der Herrschaft von Big Tech: »Eric Schmidt: Wenn diese Technologie breiter verfügbar wird, was sehr schnell der Fall sein wird, werden die Probleme noch viel größer werden. Mir wäre es viel lieber, wenn die Unternehmen vernünftige Grenzen festlegen würden. – Reporter: Es soll kein regulatorischer Rahmen sein, es soll vielleicht nicht einmal eine Art demokratische Abstimmung sein, es soll das Fachwissen innerhalb der Branche sein, das hilft, das zu klären? – Schmidt: Das wird zuerst die Industrie tun, denn es gibt keine Möglichkeit für Nicht-Insider zu verstehen, was möglich ist. Es ist einfach zu neu, zu schwierig, es gibt nicht das nötige Fachwissen. Es gibt niemanden in der Regierung, der es richtig machen kann, aber die Industrie

sich nicht mit bloßen ethischen Richtlinien begnügen.²⁸ Wie die Entlassung von Timnit Gebru und Margaret Mitchell bei Google Ende 2020 gezeigt hat, sind sogenannte Ethikabteilungen bestenfalls ein Feigenblatt der Verantwortlichkeit, das sich Unternehmen nach Belieben anheften und stets wieder ablegen können.

Es ist unbestritten notwendig, gesetzliche Regelungen zu schaffen, angefangen mit dem Verbot der Nutzung von KI für betrügerische Zwecke und von LLM-Training ohne Zustimmung der Dateneigentümer.²⁹ Außerdem könnte das Kartellrecht die Zerschlagung großer Unternehmen ermöglichen,³⁰ beispielsweise mag es wünschenswert sein, das Sammeln von Datensätzen und den Trainingsprozess in zwei getrennten rechtlichen Einheiten zu halten.

Der Digital Markets and Services Act der Europäischen Union ist in dieser Hinsicht besser dazu in der Lage als derzeitige Bemühungen der USA. Auch das oft gehörte Argument, jegliche Regulierung behindere Innovation, ist unangebracht: Lokale Regulierungsbemühungen innerhalb starker Wirtschaftsblöcke haben globale Auswirkungen, und die aggressiven Wettbewerbsgesetze der EU haben ihr oft

kann es in etwa richtig machen, und dann kann die Regierung eine Regulierungsstruktur um das herum aufbauen.« NBC News, »Life and death decisions are being made« by artificial intelligence vom 14. Mai 2023.

28 Luke Stark, *Breaking Up (with) AI Ethics*. In: *American Literature*, Nr. 95/2, Juni 2023.

29 *Antitrust and Competition: It's Time for Structural Reforms to Big Tech*. In: *AI Now* vom 11. April 2023 (ainowinstitute.org/publication/antitrust-and-competition).

30 Nick Srnicek, *Platform Capitalism*. London: Polity Press 2017.

einen Vorteil gegenüber den Bemühungen der USA verschafft und dazu geführt, dass sie im Ausland nachgeahmt wurden.³¹ Wie ein Bericht des *Time Magazine* allerdings gezeigt hat, ist es OpenAI-Lobbyisten bereits gelungen, die Gesetzgebung des europäischen AI Acts in ihrem Sinn zu beeinflussen.³²

Aus diesem Grund es ist notwendig, auch hier größer zu denken. Wenn diese traditionelleren Regulierungsmaßnahmen unwirksam sind und KI-Systeme zum Ort von Gesellschaftsentwürfen, zu dominanten Faktoren der Öffentlichkeit oder gar zu einer politischen Infrastruktur selbst werden, spricht viel dafür, sie tatsächlich auch öffentlicher Kontrolle zu unterstellen. Sollten bloße Leitplanken sich als zu schwach dafür erweisen, weil etwa der Einsatzbereich von Allzweck-KIs gar nicht mehr klar abzustecken ist, bleibt als letztes Mittel die Vergesellschaftung.

Akzeptiert man die Vorstellung, dass die Infrastrukturen technologisch vermittelter Kommunikation Infrastrukturen wie jede andere sind – Wasser, Strom, Straßen – und dass ihr Aufbau in erheblichem Maß von direkter oder indirekter öffentlicher Finanzierung sowie von »data as labor« unbekannter Nutzer abhängt,³³

31 Anu Bradford/Adam Chilton u.a., *The Global Dominance of European Competition Law Over American Antitrust Law*. In: *Journal of Empirical Legal Studies*, Nr. 16/4, Dezember 2019.

32 Billy Perrigo, *Exclusive: OpenAI Lobbied the E.U. to Water Down AI Regulation*. In: *Time* vom 20. Juni 2023 (time.com/6288245/openai-eu-lobbying-ai-act/).

33 Eric A. Posner/E. Glen Weyl, *Radical Markets. Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press 2018.

www.klett-cotta.de



Monica Ali
Liebesheirat
Roman

Aus dem Englischen von Dorothee Merkel
592 Seiten, broschiert
ISBN 978-3-608-98746-1
€ 15,- (D) / € 15,50 (A)

**»Ein ebenso rasanter
wie kluger Gesellschafts-
roman«** Mareike Ilsemann, WDR 5

Yasmin Ghorami ist sich sicher:
Mit Joe Sangster hat sie den Mann
fürs Leben gefunden. Doch als der
Hochzeitstermin näherrückt und
Yasmins Eltern auf die unkonventionelle,
feministische Mutter von Joe treffen,
wirbeln ungeahnte Geheimnisse
Yasmins strukturiertes Leben durcheinander.



Klett-Cotta

dann erscheint die Vergesellschaftung bestimmter Technologien weniger als schockierende Übergriffigkeit des Staates denn vielmehr als die zwingende Folge der Einsicht, dass öffentliche Güter und Dienstleistungen in den Händen einer selbstverwalteten Öffentlichkeit liegen sollten.

So erscheint auch der offene Brief in einem anderen Licht. Nicht bloß als technologischer Katastrophismus einer Riege von Longtermists, sondern als Versuch, von den politischen Konsequenzen dieser

Technik abzulenken. Denn die sind sehr viel konkreter als die Diktatur der Maschinen, selbst in der krassen und heuristisch überspitzten Art und Weise, in der ich sie hier dargestellt habe. Die Überwachung dieser Technologien stellt damit verglichen eine viel gefährlichere Bedrohung für die Unternehmen und Einzelpersonen dar, die am meisten von dem Hype um die KI profitieren – aus demokratietheoretischer Sicht aber ist sie unabdingbar.