

Kritische Textreihe zu  
kuenstlicher Intelligenz



capulcu. blackblogs. org

# Kritische Textreihe zu Künstlicher Intelligenz

November 2023

## Vorwort

Wir haben etwas gebraucht, um einzusehen, dass wir das Geld für die Stromrechnung nicht zusammenbekommen. Eine eigene KI (künstliche Intelligenz) steht uns also nicht zur Verfügung und wir müssen unsere natürliche Intelligenz bemühen, um eine Reihe von Texten zum Thema KI und dessen Folgen zu schreiben, diskutieren, verwerfen und zu veröffentlichen. Wir haben vor, in lockerer Abfolge auf einzelne Aspekte näher einzugehen und die gewogene Weltöffentlichkeit damit zu belästigen. Die Texte sind fokussiert und gehen eher in die Tiefe, als in die Breite – sie mögen Widerspruch oder Begeisterung auslösen. Wir verstehen sie als Diskussionsangebote im Prozess einer Entwicklung einer gemeinsamen Analyse dieses neuen Hypes – über kritische Anmerkungen würden wir uns freuen. In diesem Sinne:

*Viel Spaß beim Lesen.*

## Inhalt

- 1) DIE GOLDGRÄBER DER KÜNSTLICHEN INTELLIGENZ
- 2) KLIMA: DAS GRÜNE VEHIKEL FÜR DIE KI-OFFENSIVE
- 3) CHATGPT ALS HEGEMONIEVERSTÄRKER

keep the future unwritten



capulcu.blackblogs.org  
capulcu@nadir.org  
Fingerprint (pgp-key):  
AF52 0854 7EF1 711A F250 57CB D0D0 A3C5 DF30 9590

V.i.S.d.P. E. Schmidt | Am Zuckerberg 14 | 21984 Silikontal

# Die Goldgräber der Künstlichen Intelligenz



Ende Mai 2023 wird ein kurzes Statement<sup>1</sup> veröffentlicht, welches vor der Auslöschung der Menschheit durch Künstliche Intelligenz warnt. Unter das Statement haben Koryphäen der KI-Forschung, CEOs von KI-Unternehmen wie OpenAI und weitere prominente Figuren des Tech-Sektors ihre Unterschrift gesetzt. Diese apokalyptische Warnung reiht sich ein in eine ganze Serie gleichartiger Aussagen<sup>2</sup> von Personen und Institutionen aus dem genannten Dunstkreis. Eric Schmidt, Ex-CEO von Google und jetzt Regierungsberater, warnt vor Tausenden von Toten. Sam Altman, CEO von OpenAI, der Entwicklungsfirma von ChatGPT, fleht die US-Regierung an, Regularien für die Branche zu erlassen.<sup>3</sup>

Es ist schon etwas verwunderlich, dass ausgerechnet diejenigen vor einer Technologie warnen, die sie selbst mit Macht und viel Geld auf den heutigen Stand gebracht haben. Sie erscheinen wie Goethes Zauberlehrling, der die Geister, die er rief, nicht mehr unter Kontrolle hat. Nichts könnte falscher sein.

Spätestens mit der Veröffentlichung von ChatGPT hat KI den Mainstream erreicht und genießt die volle Aufmerksamkeit der Medien. Die Leistung dieser Software scheint eine Schwelle überschritten zu haben. KI wird jetzt nicht mehr als belächelnswerter netter Versuch wahrgenommen oder als beachtenswerte Leistung in einer nerdigen kleinen Nische, sondern ist in ihre Gegenteil gekippt; eine Technologie, die auf dem besten Wege sei, die Menschheit in Sachen Intelligenz zu überflügeln. Das geht soweit, dass ein an der Entwicklung beteiligter Ingenieur bei Google nicht von der Behauptung abzubringen war, dass

“seine” KI ein Bewusstsein entwickelt hat. Google war das augenscheinlich so unangenehm, dass die Kündigung folgte.

ChatGPT als textgenerierende KI und andere bildgenerierende KIs (z.B. Midjourney) sind Schaufensterprodukte der Branche. Ohne sie hätte die teils enthusiastische, teils beunruhigte Reaktion der Öffentlichkeit auf KI nicht stattgefunden. AlphaGo, eine KI, die den amtierenden Weltmeister im Spiel Go wiederholt geschlagen hat, hatte dazu noch nicht ausgereicht. Die apokalyptischen Warnungen aus der KI-Szene greifen genau diese Stimmung auf. Ihre Dystopie einer übermächtigen Technologie, die die Menschheit ausrottet, betont im Wesentlichen eines: die Mächtigkeit dieser Technologie. Das apokalyptische Flair dient als Ablenkung, um von den maßlosen Übertreibungen abzulenken, und obendrein als warnender Insider, die Reputation eines kritischen und reflektierten Bewusstseins mit in die Waagschale zu werfen – CEOs, die besorgt sind um das Wohl der Menschheit. Dabei geht es mitnichten um eine Warnung und schon gar nicht um das Wohl der Menschheit, sondern um eine spezifische Verkaufsargumentation: Sei dabei, bediene dich dieser übermenschlichen Macht, investiere jetzt oder schließe zumindest ein Premium-Abo ab!

Die Goldgräberstimmung ist mit Händen zu greifen. Der Ruf nach einer Regulierung dieser Technologie durch Regierungen mag verwirren, ist aber folgerichtig. Regulierungen sind nicht zwangsläufig schädlich für die Branche, im Gegenteil: Sie ebnen das Spielfeld, schaffen Übersichtlichkeit, Planbarkeit und Investitionssicherheit. Sie können benutzt werden, um dem Einstieg von Nachzüglern in den Markt (z.B. China) Barrieren

1 <https://www.safe.ai/statement-on-ai-risk>

2 <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> <https://www.cold-takes.com/ai-could-defeat-all-of-us-combined/>

3 <https://www.c-span.org/video/?528117-1/openai-ceo-testifies-artificial-intelligence>



in den Weg zu legen. Außerdem wären Regulierungen eh gekommen, aus Sicht der Branche ist es deshalb vorteilhaft, hier die Initiative zu übernehmen.

Es gibt noch einen weiteren Aspekt: den Regulierungsbehörden fehlt das nötige Fachwissen. Das gilt allerdings für praktische alle Technologien, deren Einsatz reguliert wird – das Fachwissen muss von außen hinzugezogen werden. Besonders im Falle von KI ist dieses Fachwissen allerdings stark konzentriert, die Entwicklung wird im Wesentlichen von den Forschungsabteilungen der großen Konzerne voran getrieben. Das Fachwissen für die Regulierung kommt also ausgerechnet aus der Branche, die reguliert werden soll, die Interessenskonflikte sind vorprogrammiert. Für die Tech-Konzerne beste Startbedingungen, um eine Quasi-Selbstregulierung im eigenen Sinne durchzudrücken. Dieses Muster ist übrigens nicht neu, sondern lässt sich in vielen vergleichbaren Vorgängen wieder finden – und das nicht nur in den USA. Was vielleicht nicht neu, aber diesmal besonders auffällig ist, ist die Dringlichkeit, mit der das Anliegen vorgebracht wird.

Ein Schlaglicht darauf wirft Sam Altmans Kritik an den KI-Regulierungen der EU. Nach intensiver Lobbyarbeit ist es OpenAI und Google gelungen, „Allzweck“-KI-Anwendungen wie etwa ChatGPT aus der Kategorie der Hochrisiko-Technologien heraus zu bekommen, die mit strengen Auflagen belegt ist. Statt dessen wurde für diese Fälle eine neue Kategorie der „Foundation Models“ mit aufgeweichten Auflagen geschaffen.<sup>4</sup> Regulierungen sind für Herrn Altman nur solange OK, wie sie nicht geschäftsschädigend sind.

Ein Seiteneffekt der suggerierten Dringlichkeit ist die Erzeugung eines Eindrucks, dass jetzt etwas Neues aufgetaucht sei. KI blickt aber auf eine Jahrzehnte alte Geschichte. AlphaGo wurde schon genannt. KI-basierte Gesichtserkennung zum Beispiel bei Zugangssystemen, aber auch in Überwachungskameras, wie etwa am Berliner Südkreuz mit notorischer Schwäche, Menschen mit nicht-weißer Hautfarbe zu erkennen. Auch Betrugserkennungssysteme basieren auf KI und haben dort für eine Reihe von desaströsen Skandalen gesorgt – erinnert sei hier beispielsweise an die Toeslagenaffaire<sup>5</sup> in den Niederlanden. Unternehmen wie Clearview AI oder PimEye haben mit Porträtfotos aus dem Internet Bilddatenbanken aufgebaut, die sich mit Hilfe von KI durchsuchen lassen – ein Schnappschuss einer Person kann schon ausreichen, um Name, Arbeitgeber oder Adresse herauszufinden<sup>6</sup> - Stalkerware, nicht nur für Repressionsbehörden. Eine ausführlichere Liste lässt sich in unserem Text „KI zur Programmatischen Ungleichbehandlung“ finden<sup>7</sup>. Ein solches Erbe im Gepäck verdüstert die Akzeptanz.

Das Gold, welches die Goldgräber der Tech-Branche zu finden hoffen, ist die Automatisierungsdividende. KI verspricht, Vor-

gänge automatisieren zu können, die sich bislang erfolgreich entzogen haben. In einer Studie<sup>8</sup> über die Auswirkungen der KI schätzt GoldmanSachs, dass 66% aller Arbeitsplätze in den USA betroffen sein werden. Dort könnten 25-50% der anfallenden Aufgaben von KI übernommen werden. Andere Studien<sup>9</sup> kommen zu ähnlichen Zahlen. Es ist die Aussicht auf diese Produktivitätssteigerung, die die oben genannten Aktivitäten anspricht.

Ende des 19ten, Anfang des 20ten Jahrhunderts entwickelte und popularisierte Frederick Taylor u.a. eine Methode, die später als „Wissenschaftliches Management“ oder besser „Taylorismus“ bekannt wurde. Erklärtes Ziel dieser Methode war es, die Arbeit zu effektivieren, mehr Leistung aus jede\*r Arbeiter\*in rauszupressen. Dazu wurden Arbeitsabläufe minutiös dokumentiert, analysiert und optimiert, um sie dann in neu zusammengesetzter Form in den Produktionsprozess zurückzubringen, mit der klaren Absicht, „ineffiziente“, tradierte Abläufe und Arbeitsformen zu ersetzen. Aus Facharbeiter\*innen wurden austauschbare Massenarbeiter\*innen, Anhängsel der Maschinen, die ab sofort den Arbeitstakt vorgaben.

Zentraler Baustein dieser Methode war ein Wissenstransfer von den Facharbeiter\*innen in die Ingenieursetage. So war es möglich, dieses Wissen einzusetzen, ohne von den Menschen, von denen das Wissen stammte, abhängig zu sein. Der Transfer war im Kern ein Transfer der Verfügungsgewalt über dieses Wissen. Die Folge war eine Entmachtung der Facharbeiter\*innen im Produktionsprozess, eine Dequalifizierung der Arbeit und damit eine Verschlechterung der Verhandlungsbedingungen, wenn es z.B. um Lohnforderungen oder Arbeitsschutz ging. Ein vergleichbarer Transfer findet beim Training einer KI statt.

Der Taylorismus treibt sein Unwesen seit mehr als hundert Jahren und Computer sind auch nicht erst gestern erfunden worden. Damit Abläufe aus der „analogen“ Welt in einem Computer repräsentiert und ausgeführt werden können, müssen sie formalisiert werden: in einen Satz von detaillierten Regeln bzw. Anweisungen übersetzt werden, ganz ähnlich wie auch im „wissenschaftlichen Management“. Das funktioniert je nach betrachtetem Ablauf bzw. Problemstellung unterschiedlich gut. In vielen Fällen bleibt ein „Rest“, der sich einer Formalisierung entzieht, das Ergebnis passt dann nur unvollkommen auf die Problemstellung. In anderen Fällen ist es schon schwierig zu benennen, wie überhaupt an die Formalisierung einer Problemstellung herangegangen werden kann.

Formalisierung lässt sich begreifen als eine Art notwendiger Übersetzungsschritt, der eine Aufgabenstellung „computergängig“ macht – ein Schritt, der von Menschen geleistet wird. Für die KIs der aktuellen Generation wird erst gar nicht versucht, eine Aufgabenstellung zu formalisieren. Statt dessen wird die

4 <https://time.com/6288245/openai-eu-lobbying-ai-act/>

5 <https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/>

6 <https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html>

7 <https://capulcu.blackblogs.org/wp-content/uploads/sites/54/2020/06/DIVERGE-small.pdf> – ab Seite 33

8 [https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst\\_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs\\_Kodnani.pdf](https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf)

9 <https://arxiv.org/abs/2303.10130>



KI in einem trial-and-error-Prozess unter enormen Aufwand an die Aufgabenstellung heran trainiert. Der Schritt, der die Aufgabenstellung computergängig macht, wird also vom Computer selbst ausgeführt. Die Trainingsdaten werden so häufig „durchgekaut“<sup>10</sup>, bis die KI zufriedenstellend plausibel die in den Trainingsdaten enthaltenen Eigenschaften nachahmen bzw. wiedererkennen kann.

Beim Training entsteht eine Art stochastisches Extrakt der Trainingsdaten, ein Tensor aus Billionen von Zahlen, der sich im Hauptspeicher der KI ausbildet. Welche Aspekte der Trainingsdaten extrahiert werden, hängt von der Gestaltung des Trainings, der Topologie der KI, der Aufbereitung der rohen Trainingsdaten und weiteren begleitenden Maßnahmen ab. Entscheidend ist, dass im Tensor die notwendigen Informationen für das plausible Nachahmen bzw. Wiedererkennen landen – im allerweitesten Sinne also das „Wissen“. Wie auch immer dieser Wissenstransfer in den Tensor beurteilt werden mag – schließlich ist Nachahmen etwas anderes als Verstehen –, er erlaubt eine „Reproduktion“, ohne auf die Menschen zurückgreifen zu müssen, von denen das Wissen stammt. Wie schon beim Taylorismus findet ein Transfer von Verfügungsgewalt statt. Dieser Transfer ist das Fundament der Automatisierungsdividende.

Auch die in klassischer Programmierung verwendete Formalisierung implementiert einen Transfer der Verfügungsgewalt, allerdings wird dieser von Menschen in Hand- oder besser Kopfarbeit gemacht und ist deshalb nur schlecht zu skalieren. KI verspricht, diesen Formalisierungsschritt zu überspringen und den Transfer selbst in einen automatisierbaren und damit skalierbaren Prozess zu verwandeln – und das ist der qualitative Sprung in der Enteignung von „Wissen“.

Das Umschiffen der Formalisierung zur Übertragung einer Aufgabenstellung auf Computer erlaubt es zwar, neue Anwendungsgebiete zu erschließen, kommt aber mit einigen Nachteilen daher. Formalisierung setzt voraus, dass ein Problem bis ins Detail verstanden wurde – dass dabei Fehler passieren und Missverständnisse ausdetailliert werden, ist dazu kein Widerspruch. Das Ergebnis lässt sich überprüfen, mit einigem Aufwand ist es sogar möglich, einen mathematikartigen Beweis zu führen.

Bei KI ersetzt das Training das Verständnis, letztendlich ist das Training ein Schuss ins Blaue. Die Begeisterung vieler Ingenieur\*innen von ChatGPT und anderen KIs reflektiert deren Überraschung, wie gut dieser Schuss gelungen scheint.

Dem Extrakt der Trainingsdaten ist nicht anzusehen, was genau extrahiert wurde – was genau die KI „gelernt“ hat. Dementsprechend sind die Ausgaben, die eine KI produziert, fehler-

behaftet. Das Einsatzgebiet für KIs zielt auf Anwendungen, bei denen entweder Fehler „tolerierbar“ sind oder sie in Konkurrenz zu menschlicher Arbeit treten, die ebenfalls fehlerbehaftet ist. Oder sie tritt in Konkurrenz zu im weitesten Sinne kreativen Tätigkeiten, die nicht binär richtig oder falsch, sondern besser oder schlechter sind. Letztendlich findet hier eine ökonomische Abwägung über Kosten und Nutzen statt, deren Ergebnis einzig vom (positiven) Einfluss auf das Geschäftsergebnis abhängen wird.

Das Ergebnis ist überaus zynisch: Wenn ein System zur Aufdeckung von Sozialhilfebetrug Fehler macht und die Falschen beschuldigt (und in Folge die Unterstützungszahlungen verweigert), dann trifft es Menschen, die sich nur schlecht wehren können. Selbst wenn KIs fachlich schlechte Ergebnisse liefern, bauen sie (oder genauer diejenigen, die die KI einsetzen) einen Konkurrenzdruck auf Arbeiter\*innen und Angestellte auf, der Folgen z.B. bei Tarifverhandlungen haben kann. Einen Eindruck davon liefert die Antwort von Netflix auf den Streik von Schauspieler\*innen und Autor\*innen in Hollywood, die u.a. eine „Zweitverwertung“ ihrer Leistungen durch KI-„generierte“ (besser: kopierte) Inhalte verhindern wollen: Netflix schreibt eine gut bezahlte Stelle für einen „KI Produkt Manager“ aus, für „alle Bereiche“, was genau diese Zweitverwertung beinhaltet.<sup>11</sup>

Wie schon der Taylorismus, wird die KI zu einer Verschiebung gesellschaftlicher Macht „nach oben“ führen, gefolgt und verstärkt durch eine entsprechende Reichtumsumverteilung in die gleiche Richtung. KI wirkt wie ein Verstärker gesellschaftlicher Ungleichheit. Der enorme Ressourcenaufwand, den die KI-Technologie verlangt – Trainingsdaten, Energie, Wasser und leistungsfähige Hardware – lässt an einer „Demokratisierung“ dieser Technologie zweifeln. Einen eigenen Web- oder Mailserver im Internet zu betreiben ist vielleicht nicht trivial, aber durchaus von normalsterblichen Individuen leistbar. Für KI gilt das auf absehbare Zeit nicht, sie wird Werkzeug der Mächtigen bleiben.

Selbst, wenn sich in einem utopischen Szenario ein gesellschaftlich sinnvoller Einsatz denken ließe – die jetzige gesellschaftliche Realität besteht praktisch vollständig aus Anwendungen zu Lasten der großen Mehrheit der Menschen und reflektiert damit die aktuelle gesellschaftliche Machtverteilung.

Es muss also – in bester luddistischer Tradition – gefragt werden, wer KI für welchen Zweck einsetzt und ob die Resultate gesellschaftlich und ökologisch erstrebenswert sind. Diese Frage kann klar verneint werden.

10 Eine lesbare Einführung in die Funktionsweise von KIs vom Typ ChatGPT und was genau mit „durchkauen“ gemeint ist: <https://arstechnica.com/science/2023/07/a-jargon-free-explanation-of-how-ai-large-language-models-work/>

11 <https://theintercept.com/2023/07/25/strike-hollywood-ai-disney-netflix/>



# Klima: Das grüne Vehikel für die KI-Offensive



/imagine Visiting Speicherstadt in Hamburg in 2072 ::considering the impact of climate change:: --ar 3:2

„Die KI versteht. Warum wir nicht?“ - Die Mobilisierung von Fridays for Future Hamburg zum Klimastreik am 15.09.2023 unterstellt dem KI-basierten Bildgenerator den Klimawandel besser zu durchschauen als die Menschen.

Als Microsoft-CEO Satya Nadella im Februar die neue KI-gestützte Version der Suchmaschine Bing vorstellte, beschwor er mit dem „dreiköpfigen Monster aus Inflation, Rezession und Energiekrise“ das ganz große Bedrohungsszenario herauf. Natürlich nur, um direkt im Anschluss zu verkünden, dass die technische Lösung der genannten Probleme greifbar nahe sei. Es überrascht nicht, dass ein Unternehmen wie Microsoft danach „strebt, Technologie einzusetzen, um die großen Herausforderungen, die sich Menschen, Organisationen und Staaten stellen, zu bewältigen,“ wie Nadella die Ziele des Konzerns formulierte.<sup>1</sup> Schließlich geht es dem Konzern bei der öffentlichkeitswirksamen KI-Offensive u.a. darum, endlich wieder Marktanteile im Google-dominierten Suchmaschinenmarkt zu gewinnen.

Erstaunlicher als die PR-Darstellung Microsofts ist, dass auch weite Teile der Grünen und der (professionalisierten) Klimabewegung diese Vision übernommen haben. Der massive Ausbau des Einsatzes von Technologie und insbesondere von sogenannter Künstlicher Intelligenz wird als große Chance, wenn nicht gar Notwendigkeit, im Kampf gegen die menschengemachte Klimakrise wahrgenommen. Kein Tag vergeht, ohne dass nicht eine neue Technologie als „die“ Lösung für das Klimaproblem präsentiert wird. Sollte die Vision des umfassenden Einsatzes von KI zur Lösung politisch-sozialer Probleme Realität werden, käme es zu tiefgreifenden sozialen Umwälzungen

im Sinne der leistungsfähigen, progressiven bürgerlichen und inzwischen oft grünen Eliten. Im diesem Artikel wollen wir uns genauer mit der grünen Technologie-Gläubigkeit befassen. Leitend für unsere Überlegungen ist die Frage, welche Folgen für individuelle und kollektive Selbstbestimmung von der Bekämpfung des Klimawandels mittel Künstlicher Intelligenz zu erwarten sind.

## DIE HOFFUNG AUF TECHNISCHE LÖSUNGEN

Die Initiative, einen informationstechnologischen Vorstoß in die Weiterentwicklung des Kapitalismus einzubringen, reicht weit zurück. Die Idee entstand jedoch – anders als man meinen könnte – zunächst nicht in Konzernzentralen wie der von Microsoft, sondern in der US-amerikanischen Gegenkultur der 60er Jahre. Die aktuellen grünen Vorstellungen eines sich selbst regulierenden Klimaschutzes und einer mittels Rückkopplungsschleifen von Informationsströmen optimierten Gesellschaft knüpfen an die Utopien der damaligen Alternativbewegung an.<sup>2</sup> Diese Vorstellungen münden heute in dem Vorstoß, Klimapolitik und KI in einem technopolitischen Komplex neuer Art zum Kern eines totalisierenden Durchbruchs in eine neue Ära des Kapitalismus zu machen.

1 <https://news.microsoft.com/wp-content/uploads/prod/2023/02/Reinventing-search-with-a-new-AI-powered-Bing-and-Edge-1.pdf>

2 Vgl. <https://jacobin.de/artikel/techno-okologie-astrid-zimmermann-klimawandel-blockchain-terra0-web3-whole-earth-catalogue-buckminster-fuller-solutionismus-design-thinking/>



Bei der Heinrich-Böll-Stiftung läuft dieser Ansatz unter der Reflexion und Offenheit betonenden Überschrift: „KI & Klimawandel – Hype oder Chance?“ Die Heinrich-Böll-Stiftung ist, mehr als jede andere politische Stiftung anderer Parteien, das Strategie- und Diskursorgan (zur Meinungsbildung und Stimmungstest) der Grünen. Klimapolitik erscheint dabei sowohl als Vehikel des technologischen Durchbruchs, als auch als Vehikel zur Verdrängung alternativer politischer Ansätze (des radikalen Klimaschutzes und der gesellschaftlichen Veränderung) und weist den IT-Technologien ein Monopol in der Klimapolitik zu. Auch Ralf Fücks, einer der Gründer:innen des grünen Thinktanks Zentrum Liberale Moderne und ehemaliger Leiter der Heinrich-Böll-Stiftung, schwärmt von einer leuchtenden Zukunft dank technologischer Innovation:

*„Aus dem Wettlauf gegen den Klimawandel [...] kann eine neue ökonomische Dynamik entstehen, eine lange Welle umweltfreundlichen Wachstums. Ihre Treiber sind Künstliche Intelligenz und die kybernetische Steuerung von Produktion und Logistik, Wasserstoff und synthetische Kraftstoffe, E-Mobilität und Batterietechnik, nachwachsende Werkstoffe, Bionik und das weite Feld der Biotechnologie mit ertragreicheren, robusteren Nutzpflanzen und Lebensmitteln aus Zellkulturen.“*

Von dem Literaturwissenschaftler Roberto Simanowski wird diese Haltung so zusammengefasst:

*„Die Hoffnung, dass die Technik uns rechtzeitig aus dem Desaster, auf das wir zusteuern, errettet, indem sie, gepaart mit ‚nachhaltigem‘ Konsum, ein ‚grünes‘ Wirtschaftswachstum ermöglicht, ist [...] nur Ausrede dafür, am Status quo nichts Wesentliches ändern zu müssen. KI in ihrer schwachen Form ist erklärter Ausdruck dieser Hoffnung, sich mit Technik, statt einer Kehrtwende, vor den Folgen der bisher entwickelten Technik schützen zu können; durch den effizienten Einsatz von Wärme mittels intelligenter Thermostate, die Optimierung von Verkehrslenkung in der intelligenten Stadt oder die CO<sub>2</sub>-Rückbindung aus der Atmosphäre in den Boden.“<sup>3</sup>*

Die Hoffnung auf Technik ist demnach ein solutionistischer Ansatz, d.h., das sozial erzeugte Problem des Klimawandels wird in technische Ersatzprobleme, z.B. den enormen Verbrauch von fossiler Energie übersetzt, die dann mittels KI gelöst bzw. weg-optimiert werden können.<sup>4</sup>

## TECHNOZÄN STATT ANTHROPOZÄN?

Technokrat:innen und Solutionist:innen sehen in der (derzeitigen) politischen Unfähigkeit, einen klimagerechten Richtungswechsel auch nur einzuleiten, die Bestätigung dafür, dass ‚der Mensch‘ nicht in der Lage sei, a) über seine eigenen Bedürfnisse hinaus und b) über das unmittelbare Hier und Jetzt hinaus rationale Entscheidungen im Sinne eines (globalen) Gemeinwohls zu treffen. Als quasi naturgesetzlich soll diese Einsicht in die ‚menschliche Unfähigkeit‘ den Weg für eine Künstliche Intelligenz ebnen. Sie könne die Klimakrise weit besser als der

Mensch lösen, weil sie viel besser Daten prozessieren und komplexe, klimarelevante Zusammenhänge detektieren kann.

Mensch könnte zynisch anmerken, eine KI (ausgestattet mit weitgehender Entscheidungsbefugnis) könne kaum eine noch schlechtere Klimapolitik machen als die derzeitige Politik. Doch anders, als es *James Lovelock*, eine der intellektuellen Referenzen von Teilen der Ökologiebewegung in seinem Buch *No-vo-zän - Das kommende Zeitalter der Hyperintelligenz* analysiert, stehen wir schon lange nicht mehr vor einem Wissens- sondern vor einem Willensproblem. Der politische Wechsel vom Fokus des *Individuums* in seiner Jetzt-Zeitigkeit hin zu einer Gesellschaft, welche die zukunftsfähige *Gemeinschaft* radikal in den Mittelpunkt stellt, lässt sich nicht per Entscheidungs-Überantwortung an eine künstliche Intelligenz abkürzen. Der Grund dafür liegt weniger im technischen Problem einer niemals ausgewogenen Datenbasis, die zum Training selbstlernender Algorithmen der KI herangezogen wird, und damit zu einer unbrauchbaren Verstärkung dieser Datenvorurteile durch die KI führt. Der Grund liegt vielmehr in der konzeptionellen Unzulänglichkeit des Maschinenlernens, eine brauchbare Notion von *Gemeinwohl* abzubilden.

Eine semantisch ahnungslose KI, die lediglich Mustererkennung und -optimierung statistischer Gewichte betreibt, hat keine Idee von dem, was ein Gemeinwohl sein könnte und wie es sich sinnvoll dynamisch weiterentwickeln lässt, egal wie beeindruckend ‚menschenähnlich‘ selbstlernende Sprachmodelle à la ChatGPT Problemlösungsstrategien bereits jetzt imitieren. Schlimmer noch - das Konzept der Optimierung auf der Basis bestehender Daten schreibt unweigerlich Vergangenes (herrschaftsstabilisierend) in die Zukunft fort. Damit entpuppt sich eine entmündigende KI als Empfehlungs- und Entscheidungsassistent zur Lösung der ökologischen Krise als sozio-technologische Sackgasse - sie ist als (techno-)revolutionäres Hilfsinstrument untauglich. Ihre Attraktivität speist sich lediglich aus einer doppelten Verantwortungsabgabe erstens für die Mehrheit der Menschen, die sich nicht mehr mit dem von ihnen verursachten Klimawandel befassen müssen, da eine KI ohnehin bessere Lösungen findet als sie selbst, und zweitens für die Entscheidungsträger:innen, die den politischen Charakter wegweisender gesellschaftlicher Entscheidungen verschleiern und ihre eigene Verantwortung der Öffentlichkeit gegenüber auf die KI abwälzen können, um so ggf. auch unpopuläre Maßnahmen durchsetzen zu können.

## GRÜNE KI ALS AKZEPTANZBESCHAFFUNG FÜR SOZIALE UMWÄLZUNGEN

Es ist festzustellen, dass die Hoffnung auf Technik in der grünen Vision nicht lediglich als bewahrend im Sinne der ‚Verwaltung einer bereits etablierten Ordnung‘ zu charakterisieren ist, sondern die Verheißung der KI gegenüber denjenigen, die über ihren Einsatz bestimmen, gerade darin besteht, die geltenden Regeln bis in die letzten Verästelungen des sozialen Gefüges durchsetzen zu können. Eine solche Optimierung der Durch-

<sup>3</sup> Roberto Simanowski. 2020. Der Todesalgorithmus. Das Dilemma der künstlichen Intelligenz. S. 112.

<sup>4</sup> Zur Diskussion des Solutionismus vgl. Redaktionskollektiv Capulcu. KI zur programmatischen Ungleichbehandlung, In: Redaktionskollektiv Capulcu. 2020. Diverge – Abweichendes vom rückschrittlichen „Fortschritt“.



setzung sozialer Ordnung mittels KI-getriebener Automatisierung – auch wenn diese den progressiven Zielen im Geiste des Klimaschutzes folgt – ist mehr als nur eine Erfassung immer weiterer Bereiche menschlichen Lebens im Sinne quantitativer Expansion. Sie bildet die Grundlage für eine tiefgreifende qualitative Transformation sozialer Beziehungen. Die Politik eines kombinierten klimapolitisch/technologischen Durchbruchs kann daher als total, oder besser „totalisierend“ bezeichnet werden in dem Sinne, wie es vor hundert Jahren die Kombination von Taylorismus/Fordismus und Durchsetzung von Einsparnissen und durchaus, wie in den USA, ökologischen Zielsetzungen gewesen ist. Totalisierend ist die grüne KI-Offensive auch deswegen, weil sie im Sinne einer gesamtgesellschaftlichen Veränderung bis in die Mentalitäten der Menschen hinein wirkt und wirken soll.

Das lässt sich aus der Schrift der Heinrich-Böll-Stiftung „Smarte Technologie gegen den Klimawandel, 15 Fakten über künstliche Intelligenz“ bis ins Einzelne ablesen.<sup>5</sup> Sie stellt den Aufriss eines umfassenden gesellschaftspolitischen Projekts dar. Hier wird der klimapolitische/technologische Einsatz auf den Gebieten von Ressourcenverbrauch, Industrie 4.0, Verkehr und Mobilität, Landwirtschaft, Waldwirtschaft, Artenverwaltung durchgespielt. Eindeutig mit der Tendenz der Ausweitung auf weitere gesellschaftliche Bereiche. Herausheben möchten wir hier Folgendes: Eine Politik unbedingter Vermeidung der Katastrophe wird gar nicht erst verfolgt, es geht um „gesellschaftliche Anpassung“ (S.8). Kapitalismuskritische Gesichtspunkte tauchen ebenfalls nicht mehr auf, wenn es heißt „Energemarkt nachvollziehbar machen“ (S. 14). Die Propaganda einer „Präzisionslandwirtschaft“ unter Einsatz von KI steht in frappanter Analogie zu der stalinistischen Strategie eines totalisierenden Zugriffs, die gesamte Landwirtschaft bis ins Letzte in eine fordistische/tayloristische Maschine zu transformieren, mit den bekannten katastrophalen Ergebnissen, wie Josephson sie als „Brute-Force-Technologie“ beschrieben hat. Die totalisierende Tendenz zu einer neuen klimapolitisch-technologischen Expertokratie ist derart hermetisch, dass alternative politische Formen der Klimapolitik und des Verhältnisses zu den neuen Technologien gar nicht mehr auftauchen. Klima- und KI-Politik nehmen die Form eines geschlossenen Systems an, das Spielräume für autonome Prozesse nicht mehr zulässt. Kritik? „Um Vertrauen in die KI zu schaffen, müssen wir uns auch mit ihren möglichen negativen Folgen auseinandersetzen“ (S.32).<sup>6</sup>

Mit ihrer betont reflektierten Herangehensweise an das Thema sind die Grünen nicht alleine. Immerhin hat das Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU) schon 2021, also noch vor der grünen Regierungsbeteiligung, auf Bundesebene ein „Fünf-Punkte-Programm

„Künstliche Intelligenz für Umwelt und Klima“ ins Leben gerufen. Darin geht es vor allem um die Schaffung sogenannter KI-Leuchttürme, „Projekten also mit Strahlkraft für den Umweltschutz.“<sup>7</sup> Und auch im BMU gibt man sich gern kritisch: „Denn es gibt ökologische Schattenseiten, die wir in den Blick nehmen müssen: Die Milliarden von Berechnungen auf Hochleistungsprozessoren, die den KI-Systemen ihre beeindruckenden Fähigkeiten verleihen, verschlingen viel Energie,“ heißt es im zugehörigen Factsheet, nur um einen Absatz weiter davon zu träumen, „eine starke Marke ‚Sustainable AI made in Europe‘“ zum eigenen Wettbewerbsvorteil zu machen.

Zusammen mit den Bundesministerien für Arbeit und Soziales sowie für Familie, Frauen, Senioren und Jugend hat das BMU auch das Forschungsprojekt „Civic Coding - Innovationsnetz KI für das Gemeinwohl“ gefördert. Darin legt die Bundesregierung einen international einzigartigen Fokus auf „gemeinwohlorientierte“ KI-Entwicklung, wobei einer der Schwerpunkte mit der „KI-Ideenwerkstatt“ auch hier auf Umweltzielen liegt.<sup>8</sup> Die faktisch genannten Projekte sind dann allerdings meist eher mickrige Leuchttürme. Da werden hier mal ein paar Bienenstöcke mit Sensoren ausgestattet, um dem Bienensterben auf den Grund zu gehen, und dort mal der Bewässerungsbedarf von Stadtbäumen prognostiziert. Es drängt sich der wenig überraschende Eindruck auf, dass komplexere Probleme von einer automatisierten Lösung durch KI weit entfernt sind und es eher darum geht, den Diskurs über diese Technologien von Seiten des Ministeriums aktiv mitzugestalten.

## KLIMAKILLER KI

Die Bundesregierung gibt den KI-Leuchttürmen nicht ohne Grund einen kritisch-reflektierenden Anstrich. Denn es ist keineswegs ausgemacht, dass KI Teil der Lösung und nicht Teil des Problems ist. Schließlich ist der Ressourcen- und Energieverbrauch des maschinellen Lernens enorm. So schätzt etwa Facebook die beim Training des eigenen Sprachmodells LLaMA verbrauchte Energie auf 2638 MWh.<sup>9</sup> Zur Einordnung: Ein modernes Windrad, das etwa 3500 Haushalte mit Energie versorgen kann, muss über drei Monate laufen, um eine solche Menge an Energie zu produzieren.<sup>10</sup> Während Facebook in Aussicht stellt, dass nach dem einmaligen Trainingsprozess der Energieverbrauch relativ gering ausfalle und u.U. schon eine einzelne GPU ausreichen könne, um die trainierte KI zu betreiben, muss hier festgestellt werden, dass zumindest durch den geplanten massenhaften Einsatz der KI ein enormer Energiebedarf zu erwarten ist – von anderen versteckten Energiekosten wie der Herstellung der nötigen Hardware einmal gänzlich abgesehen.

5 [https://www.boell.de/sites/default/files/2022-04/BoellFakten\\_Smarte\\_Technologie\\_gegen\\_den\\_Klimawandel\\_15\\_Fakten\\_ueber\\_Kuenstliche\\_Intelligenz.pdf](https://www.boell.de/sites/default/files/2022-04/BoellFakten_Smarte_Technologie_gegen_den_Klimawandel_15_Fakten_ueber_Kuenstliche_Intelligenz.pdf)

6 Vgl. Redaktionskollektiv Capulcu. IT – Der technologische Angriff des 21. Jahrhunderts. In: Redaktionskollektiv Capulcu. 2017s. Disrupt – Widerstand gegen den technologischen Angriff.

7 [https://www.bmu.de/fileadmin/Daten\\_BMU/Download\\_PDF/Digitalisierung/factsheet\\_ki\\_bf.pdf](https://www.bmu.de/fileadmin/Daten_BMU/Download_PDF/Digitalisierung/factsheet_ki_bf.pdf)

8 <https://www.civic-coding.de/angebote/publikationen>

9 Hugo Touvron et al. 2023. LLaMA: Open and Efficient Foundation Language Models.

10 <https://www.ndr.de/nachrichten/info/Watt-Das-leisten-die-Anlagen-im-Vergleich,watt250.html>





Schon heute wird geschätzt, dass etwa 12 Prozent des weltweiten Strombedarfs in digitale Geräte fließen. Tendenz steigend.<sup>11</sup> Eine einzelne Google-Anfrage – wohlgermerkt ohne Verwendung KI-basierter Chat-Assistenten – verbraucht etwa 0,3 Wh – soviel Energie, wie benötigt wird, um eine energiesparende LED für drei Minuten zu betreiben. Eine Chat-GPT3-Anfrage, die mittels bereits trainierter KI beantwortet wird, verbraucht dagegen schon 1,3 Wh, also mehr als das 4-fache – wobei das besonders energieintensive Training hier nicht beachtet wird.<sup>12</sup> Diese Zahlen sind natürlich lediglich Schätzungen und der reale Energieverbrauch dürfte um ein Vielfaches höher liegen.<sup>13</sup> Schließlich wird nicht nur der Strom zum Betreiben der Rechenzentren verbraucht, sondern auch Produktion und Transport der Hardware sowie Entwickler:innen und ihre Ausstattung verbrauchen Energie, die in den genannten Schätzungen oftmals nicht einbezogen wird.

Angesichts des enormen Energieverbrauchs digitaler Technologien ist Skepsis angebracht, wenn diese als Mittel zur Senkung des Energieverbrauchs angepriesen werden. Die Bundesregierung scheint sich jedoch das Ziel gesetzt zu haben, diese Zweifel zu zerstreuen. Daher befinden sich unter den erwähnten KI-Leuchttürmen auch einige Projekte, die das Ziel haben, den Ressourcenverbrauch der KI selbst zu optimieren bzw. ihn transparenter zu machen. Ein solcher Leuchtturm ist das Projekt NADIKI an der Universität Stuttgart, das sich zum Ziel setzt, den realen Energie- und Ressourcenverbrauch von KI über eine Software-Schnittstelle bereitzustellen. In der Pressemitteilung zum Förderbescheid heißt es:

*„Für eine nachhaltige KI-Nutzung ist es daher wichtig, vorhandene Infrastruktur bestmöglich zu nutzen, um den Bau neuer Rechenzentren, Server oder Netzwerkequipment zu reduzieren oder zu vermeiden. Gleichzeitig sollten KI-Systeme optimal ausgelastet sowie der Ressourcenverbrauch erfasst und offengelegt werden.“<sup>14</sup>*

Damit ist der Rahmen der kritischen Auseinandersetzung über die ökologischen Folgen abgesteckt. Eine ergebnisoffene Hinterfragung, ob KI – und sei es lediglich aus ökologischer Sicht – überhaupt eingesetzt werden sollte, steht nicht im Fokus. Zu klären ist lediglich noch, wie ihr Einsatz „nachhaltig“ gestaltet werden kann – oder anders ausgedrückt: Der Einsatz von KI wird zu einem Optimierungsproblem zweiter Ordnung. Eines ist auch ohne die Ergebnisse der Stuttgarter Forscher:innen sicher – zunächst steigt der Energiebedarf durch den Einsatz maschinellen Lernens. Die hohen Fixkosten, die im Training von KI-Modellen stecken, führen dazu, dass ein effizienter Einsatz nur dann denkbar ist, wenn das Modell anschließend im großen Stil angewendet wird. KI-Modelle kommen daher zur

Lösung spezialisierter (Klima-)Probleme, bei denen kein groß angelegter Einsatz zu erwarten ist, kaum in Frage.

Es gibt noch einen weiteren Grund, den Energiespar-Versprechungen grüner KI gegenüber skeptisch zu sein: den sogenannten Rebound-Effekt. Dieser besagt, dass Effizienzsteigerungen beim Energieverbrauch nicht dazu führen, dass der Verbrauch insgesamt sinkt, sondern lediglich dazu, dass die Kosten sinken und die überschüssige, nicht mehr benötigte Energie stattdessen an anderer Stelle verbraucht wird.<sup>15</sup> Ein einfaches Beispiel: die Senkung des Kraftstoffverbrauchs moderner PKW hat nicht dazu geführt, dass weniger Kraftstoffe verbraucht werden, sondern dazu, dass erstens mehr Auto gefahren wird, weil es sich mehr Menschen leisten können, und zweitens größere Autos wie SUVs produziert werden, die wiederum einen sehr hohen Kraftstoffverbrauch haben und ohne die Effizienzsteigerungen gar nicht denkbar gewesen wären. Eine Optimierung der bestehenden Wirtschaftsbereiche wird nicht zu einer realen Senkung des Energieverbrauchs führen. Unter diesen Voraussetzungen wiederum erscheinen auch die Effizienzsteigerungen durch grüne KI wenig geeignet, einen wesentlichen Beitrag für die Bekämpfung des Klimawandels zu leisten. Dennoch setzen viele Liberale, wie der Wirtschaftshistoriker Adam Tooze, angesichts des Widerstands der politischen und wirtschaftlichen Eliten gegen grundlegende soziale Veränderungen – mal mehr, mal weniger zähneknirschend – vor allem auf technische Lösungen.<sup>16</sup> Das wirft die Frage auf: Wie kann es gelingen, eine Debatte über ein Ansetzen an den polit-ökonomischen Ursachen des Klimawandels zu führen, die nicht durch den Verweis auf bevorstehende technische Lösungen schon im Ansatz abgewürgt wird?

## PROGRESSIV, ABER NICHT EMANZIPATORISCH

In ihrer taz-Kolumne kritisiert Charlotte Wiedemann die Grünen scharf für das „nicht ergründete Ausmaß europäischer Gewaltgläubigkeit im Ukrainekrieg,“ das die Grünen in Deutschland mit ihrer „feministischen Außenpolitik“ wie keine andere politische Kraft vorantreiben.<sup>17</sup> Sie präzisiert:

*„Heute sind die Grünen indes eine Kraft der Disziplinierung, der Einhegung geworden, der Betäubung und Verbravung des Denkens. Während sich andere verzweifelt ans Pflaster kleben, sind die Grünen mit den herrschenden Verhältnissen verleimt.“*

Und tatsächlich – nicht nur die grüne Position im Ukrainekrieg, sondern auch der Einsatz von KI zur Bekämpfung des Klimawandels sind Symptom einer fatalen Idee, wie sie typisch

11 <https://www.deutschlandfunk.de/stromverbrauch-digitalisierung-internet-bitcoin-rechenzentren-abwaerme-100.html>

12 Vgl. <https://medium.com/@zodhyatech/how-much-energy-does-chatgpt-consume-4cba1a7aef85>

13 Für einen Überblick über die Problematik der korrekten Erfassung des Energieverbrauchs von KI und den aktuellen Stand der Forschung s. <https://www.theguardian.com/technology/2023/aug/01/techscape-environment-cost-ai-artificial-intelligence>

14 <https://www.uni-stuttgart.de/universitaet/aktuelles/meldungen/Foerderbescheid-fuer-KI-Leuchtturmprojekt-NADIKI/>

15 Vgl. <https://www.umweltbundesamt.de/themen/abfall-ressourcen/oekonomische-rechtliche-aspekte-der/rebound-effekte>

16 Vgl. <https://nymag.com/intelligencer/2021/05/adam-tooze-on-climate-politics-after-covid.html> und <https://www.youtube.com/watch?v=w4Y9SomH9Nc>

17 <https://taz.de/Die-Entwicklung-der-Gruenen!/5940274/>



ist für postdemokratische Gesellschaften: die Durchsetzung progressiver Politik bei gleichzeitiger Aufgabe emanzipatorischer Ansprüche. Sozial bzw. ökologisch progressiv ist diese Politik deswegen, weil der Klimawandel gestoppt werden muss, um die katastrophalen Folgen insbesondere für arme Menschen und die Natur zu verhindern. Die Grünen möchten – in Teilen – die Bekämpfung des Klimawandels auf die Umsetzung algorithmisch errechneter Maßnahmen reduzieren. Eine gesellschaftliche Auseinandersetzung über die konkreten Ziele und daraus folgenden Maßnahmen ist da lediglich Beiwerk. Der Verweis auf technologische Lösungen hat somit gerade den Zweck, eine solche Debatte über tiefgreifende Änderungen an den ursächlichen gesellschaftlichen Machtverhältnissen von vornherein abzuschmettern. Anders als der oben zitierte Adam Tooze bedauern es die meisten Grünen daher nicht, mangels politischer Mehrheiten fast ausschließlich auf technische Lösungen zu setzen. Im Gegenteil: Die kybernetische Gesellschaft soll uns als grüne Utopie verkauft werden, obwohl sie doch eigentlich vor allem mehr vom Bestehenden enthält. Immerhin erkennen weite Teile der Klimabewegung, dass diese grüne Haltung einen Frontalangriff auf all jene bedeutet, die an der zapatistischen Parole „Eine andere Welt ist möglich“ festhalten wollen. Am deutlichsten wurde dieser Dissens zwischen grüner Parteispitze und Klimabewegung Anfang des Jahres am breiten Widerstand der Bewegung gegen die von den Grünen ausgehandelte Räumung des Dorfes Lützerath.<sup>18</sup>

Auch wenn die Grünen nicht – wie es im Sinne emanzipatorischer Politik wäre – die gemeinsame und offene Aushandlung von Maßnahmen gegen den Klimawandel betreiben, so bleiben sie auf einem liberalen Weg und unterscheiden sich (noch) von einem autoritären Weg, wie er etwa in China gegangen wird. Für diese spezifische Haltung, die zwar formal die liberalen Grundfreiheiten hochhält, gleichzeitig den Fokus aber gänzlich auf die Effektivität der Umsetzung politischer Ziele verschiebt, hat Colin Crouch den Begriff der Postdemokratie geprägt.<sup>19</sup> Der schon erwähnte Ralf Fücks vom Zentrum Liberale Moderne grenzt sich dementsprechend scharf von einem dezidiert autoritären Projekt ab, indem er typisch neoliberale Argumentationsmuster mit der Idee nationaler Wettbewerbsfähigkeit kombiniert:

*„Wer Freiheit und Ökologie in Einklang bringen will, muss vor allem auf Innovation setzen und den Wettbewerb um die besten Lösungen fördern. Dafür braucht es einen ökologischen Ordnungsrahmen, der die Dynamik der Marktwirtschaft in eine ökologische Richtung lenkt. Auch eine marktwirtschaftliche Klimapolitik kommt nicht ohne Gebote und Verbote aus. Sie sind aber nicht der Königsweg für die Bewältigung der ökologischen Krise. Eine Top-Down-Steuerung durch engmaschige staatliche Vorgaben kann niemals die Innovationskraft der Marktwirtschaft ersetzen, die das Wissen und die Eigeninitiative von Abermillionen Produzenten und Konsumenten bündelt.“<sup>20</sup>*

## UND DIE KLIMABEWEGUNG?

Progressiv, aber nicht emanzipatorisch – mit einer solchen Haltung ist die grüne Parteispitze nicht allein. Vielmehr spiegelt sich hier eine größere gesellschaftliche Entwicklung wider. Nicht von ungefähr gilt mit der Letzten Generation eine Kraft als die medial meist beachtete politische Bewegung in Deutschland 2023, die ihr Desinteresse an den Werten von Aufklärung und der Tradition (linker) Befreiungskämpfe offen kundtut. Carla Rochel, ein Mitglied des sogenannten Strategieteams der Letzten Generation, macht emanzipatorischer Politik eine unmissverständliche Absage zugunsten der vermeintlichen real-politischen Durchsetzung der eigenen Ziele: „Wir tun alles für eine gute Feedback-Kultur, aber wir haben leider bei anderen Organisationen gesehen, dass Basisdemokratie zu viel Zeit braucht, die wir nicht haben.“<sup>21</sup> In der Praxis bedeutet dies, dass das Strategieteam, also eine handvoll Leute, plant und die sogenannten „Bienen“ bloß auf ihren Einsatzbefehl warten, der ihnen mitteilt, wann und wo sie sich auf die Straße zu kleben haben. Nicht unwahrscheinlich, dass eine solche Haltung der Gruppe noch auf die Füße fallen wird. Wenn sich nämlich herausstellen sollte, dass das Aktions-Know-how nicht breit genug verteilt ist, um trotz staatlicher Repression gegen die Gruppe auf Dauer weiterzumachen und flexibel auf politische Veränderungen zu reagieren, könnten auch die Vorteile dezentraler Organisationsformen wieder in Erinnerung geraten.

Es gibt jedoch auch andere Ansätze innerhalb der Klimabewegung, die mehr Hoffnung machen, weil sie sich der grünen Einhegung des Denkens nicht kampflös fügen wollen. Zu nennen sind da neben dem bereits erwähnten Widerstand gegen den Kohletagebau (in Lützerath) auch die Proteste in Sainte Soline, Frankreich, gegen die sogenannten Mega-Bassins, riesige künstliche Seen, mit dem Zweck, die industrielle Landwirtschaft in Zeiten sich häufender Dürren mit Wasser zu versorgen.<sup>22</sup> Oder auch viele kleinere Aktionen, wie sie auf dem Blog <https://switchoff.noblogs.org/> dokumentiert werden. In dem dort veröffentlichten Aktionsaufruf heißt es explizit:

*„Wenn uns die Illusion verkauft wird, der Klimawandel wäre technologisch zu stoppen, dann liegt dem das Vertrauen zugrunde, die Machthabenden müssten nur die richtigen Schritte unternehmen, die richtigen Maßnahmen ergreifen, um diese Welt zu retten. Zum einen haben sie überhaupt kein Interesse an einem Ende des Expansionskapitalismus, der ihre Machtposition sichert. Und zum anderen ist die technologische Reform mit den neuen Abhängigkeiten, die sie produziert, ebenfalls zum Scheitern verurteilt.“*

Fragt sich, ob eine solche Haltung in Klima- und linker Bewegung noch mehrheitsfähig ist, oder ob der technologische Angriff auf die Aufgabe emanzipatorischer Politik schon zu weit vorangeschritten ist.

18 Vgl. <https://www.tagesschau.de/inland/innenpolitik/gruene-luetzerath-107.html>

19 Colin Crouch. 2008. Postdemokratie.

20 <https://libmod.de/aufbruch-statt-abbruch-mit-gruenem-wachstum-aus-der-klimakrise/>

21 <https://taz.de/Wer-ist-die-Letzte-Generation!/5898641/>

22 <https://tumulte.org/2023/03/articles/berichte-aus-sant-soline/>



# ChatGPT als Hegemonieverstärker



*Eine Gesellschaft mit unechten Menschen, die wir nicht von echten unterscheiden können, wird bald gar keine Gesellschaft mehr sein.<sup>1</sup> (Emily Bender, Computerlinguistin)*

Die Künstliche Intelligenz (KI) erlebt aktuell ihren *iPhone*-Moment. ChatGPT hat einen beispiellosen Hype um künstliche Intelligenz ausgelöst. Innerhalb von zwei Monaten haben mehr als 100 Millionen Menschen weltweit die neue Technik ausprobiert.

## SPRACHMODELLE – KEINE WISSENSMODELLE

Der *Chatbot*<sup>2</sup> ChatGPT basiert auf einem sogenannten *großen Sprachmodell*, das wir uns wie einen sehr großen Schaltkreis mit (im aktuellen Fall von GPT-4) einer Billion justierbarer Parameter vorstellen können. Ein Sprachmodell beginnt als unbeschriebenes Blatt und wird mit mehreren Billionen Wörtern Text trainiert. Die Funktionsweise eines solchen Modells ist, das nächste Wort in einer Folge von Wörtern aus dem ‚Erlernen‘ zu erraten. Die Bedeutung von Worten ist für ein Sprachmodell lediglich die statistische Erfassung des Kontexts, in dem sie auftauchen.

Dieses Imitieren von Text-‚Verständnis‘ bzw. ‚Wissen‘ über die Berechnung von Wahrscheinlichkeiten für das Auftauchen einzelner Wörter innerhalb von komplexen Wortmustern klappt teilweise verblüffend gut. Das Generieren von Inhalt ohne jegliches semantisches Verständnis hat natürlich den Nachteil, dass auch sehr viel Unsinn (im engeren Sinn) produziert wird. ChatGPT erzeugt mit dieser Taktik der Nachahmung von Trainingstexten beispielsweise wissenschaftlich anmutende Abhandlungen, inklusive ‚frei erfundener‘ Referenzen, die strukturell stimmig aussehen, aber nicht existieren. ChatGPT ‚erfindet‘ Dinge und produziert dadurch massenweise Fake-Inhalte – das liegt daran, dass es sich um ein statistisches Sprachmodell und nicht um ein wissensbasiertes Modell handelt.

Es ist daher für ein Restmaß an ‚Faktizität‘ im Internet wenig förderlich, dass Google und Microsoft die neuesten Versionen ihrer Suchmaschinen mit den jeweiligen Sprachmodellen ChatGPT bzw. *Bard* koppeln. Denn eines kann Künstliche Intelligenz in Form von Sprachmodellen noch weniger als jede aggregierte themenbasierte Internetsuche: Fakten prüfen. Da Sprachmodelle lediglich Wahrscheinlichkeiten von für sie bedeutungslosen Sprachformen berechnen, ist ein Faktencheck neuen ‚Wissens‘ (über die Trainingsdaten hinaus) ein blinder Fleck: Sprachmodelle leiden unter einem Phänomen, das Programmierer:innen „Halluzinieren“ nennen<sup>3</sup>. Sie sind darauf programmiert, (fast) immer eine Antwort zu geben, die auf der Ebene von ‚sich nahe stehenden‘ Wortgruppen eine hin-

<sup>1</sup> <https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html>

<sup>2</sup> Ein Computerprogramm, welches möglichst menschenähnlich kommuniziert.

<sup>3</sup> Die Psychologie spricht genauer von „Konfabulationen“.



reichend hohe Wahrscheinlichkeit haben, um für die Nutzer:in (nachträglich) Sinn zu ergeben. ChatGPT ist daher konzeptionell eine Fake-Maschine zur Produktion von plausibel erscheinenden, aber nicht notwendigerweise faktenbasierten Inhalten und damit hervorragend geeignet für die Verbreitung von Misoder gar Desinformation.

Damit verstärkt sich ein Effekt, der bereits durch das algorithmische Ranking bei den sozialen Medien erkennbar wurde. Nicht-faktengebundene Inhalte können so weit selbstverstärkend im individuellen Nachrichtenstrom ‚nach oben‘ gespült werden, dass Meinungsbilder verzerrt werden. Und damit ist die zentrale These dieses Textes:

ChatGPT ermöglicht das (automatisierte und voraussetzungslose) Produzieren von post-faktischen Inhalten, die im Wechselspiel mit der algorithmischen Reichweitensteuerung sozialer Medien und den Ranking-Algorithmen der Suchmaschinen statistisches Gewicht erlangen. Die Rückkopplung der so generierten Inhalte sozialer Medien in den Trainingsdatensatz der nächsten Generation von Sprachmodellen ermöglicht sogar eine Dominanz synthetischer Inhalte im Netz.

## ‚KANNIBALISMUS‘ UND ZENSUR BEI WACHSENDEM ANTEIL KI-GENERierter INHALTE

Eine derartige Dominanz hat messbare Konsequenzen. Die Größe von Sprachmodellen nimmt zu und damit auch der Bedarf an Trainingsmaterial für das maschinelle Lernen. Immer mehr synthetische Inhalte werden zum Training herangezogen, denn je mehr Inhalte KIs wie ChatGPT oder Google Bard produzieren, desto häufiger werden sie ihre eigenen Inhalte in ihren Datensatz aufnehmen. Das geschieht beim sogenannten Datenschrufen, bei dem automatisierte Programme nahezu alles an Daten aufsaugen, was frei im Internet verfügbar ist. Google bedient sich zudem bei den eigenen Anwendungen wie Gmail, in Speicherdiensten wie Google Drive oder Google Docs.

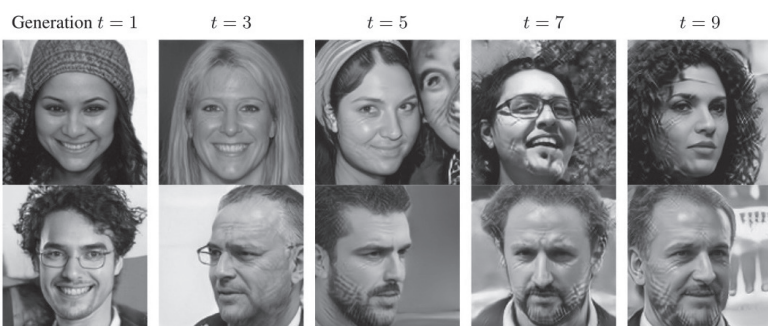
Durch das Wiederverdauen selbst generierter Inhalte entsteht eine „selbstverzehrende“ Rückkopplungsschleife, die einer nachweisbaren Störung unterliegt, der sogenannten *Model Autophagy Disorder* (MAD)<sup>4</sup>: Die Fehler zum Beispiel von

Bild-Generatoren verstärken sich rekursiv zu regelrechten Artefakten und sorgen für eine abnehmende Datenqualität. Siehe dazu die Abbildung künstlich erzeugter Gesichter bei deren Wiederverwendung als Trainingsmaterial in der nächsten Generation ( $t=3$ ) bzw. in übernächster Generation ( $t=5$ ), usw. Noch wesentlicher ist eine massiv schrumpfende Diversität der Inhalte im Netz bei zu geringer Beimischung neuer, nicht-synthetischer Inhalte. Ähnliches lässt sich bei der Text-Erzeugung durch ChatGPT beobachten.

Bereits im April 2023 (kurz nach Freischaltung des kostenfreien Schnittstelle zur Nutzung von ChatGPT) hat *NewsGuard* rund 50 Nachrichten- und Informationsseiten in sieben Sprachen identifiziert, die fast vollständig von KI-Sprachmodellen generiert werden.<sup>5</sup> Diese Webseiten produzieren eine Vielzahl von rein synthetischen Artikeln zu verschiedenen Themen, darunter Politik, Gesundheit, Unterhaltung, Finanzen und Technologie. Damit scheinen sich die Befürchtungen von Medienwissenschaftler:innen zu bestätigen: Zur Erzeugung von Werbeeinnahmen und / oder zur Debattenbeeinflussung verbreiten algorithmisch generierte Nachrichtenseiten KI-generierte Inhalte von fiktiven Verfasser:innen. Den meisten Leser:innen stehen keine Möglichkeiten zur Verfügung, diese Artikel als synthetisch zu identifizieren.<sup>6</sup>

Von Produktrezensionen über Rezeptsammlungen bis hin zu Blogbeiträgen, Pressemitteilungen, Bildern und Videos – die menschliche Urheberschaft von Online-Texten ist auf dem besten Weg, von der Norm zur Ausnahme zu werden. Pessimistische Prognosen sagen bis zum Ende dieses Jahrzehnts einen Anteil von bis zu 90% KI-generierter Inhalte im Internet voraus.<sup>7</sup> Schon jetzt tauchen diese KI-generierten Texte in den Ergebnislisten der Suchmaschinen auf. Eingreifen will Google erst bei „Inhalten mit dem Hauptzweck, das Ranking in den Suchergebnissen zu manipulieren“.<sup>8</sup>

Wie sollen wir mit der Datenexplosion umgehen, die diese KIs nun verursachen werden? Wie verändert sich eine Öffentlichkeit, die so unkompliziert mit Mis- und Desinformation geflutet werden kann? Bei steigendem Anteil können derartige synthetische Inhalte den ‚Nutzen‘ des Internet drastisch reduzieren: Wer kämpft sich durch einen (noch viel) größeren Berg an quasi-sinnloser Information – ohne Bezug zur Lebensrealität menschlicher Autor:innen? Lässt sich feststellen, ob ein Text, ein Bild, eine Audio- oder eine Videosequenz durch eine KI generiert bzw. gefälscht wurde? Schon bieten Software-Hersteller Werkzeuge zur Detektion von KI-generierten Inhalten an – selbstverständlich ebenfalls auf der Basis einer künstlich-intelligenten Mustererkennung. Menschlich verfasste Texte sollen sich über statistische Abweichungen von den Wahrscheinlichkeitsmustern der verwendeten Wortgruppierungen der KI-Sprachmodelle unterscheiden lassen. Dies sind jedoch statistische Differenzen, deren Erkennung im *Einzelfall* damit hochgradig fehleranfällig ist.



4 Alemohammad et al., Self-Consuming Generative Models Go MAD, 2023, <https://arxiv.org/abs/2307.01850>

5 <https://www.newsguardtech.com/de/special-reports/newsbots-vermehrt-ai-generierte-nachrichten-webseiten/>

6 <https://www.theguardian.com/books/2023/sep/20/amazon-restricts-authors-from-self-publishing-more-than-three-books-a-day-after-ai-concerns>

7 <https://www.youtube.com/watch?v=DgYCcdwGwrE>

8 <https://developers.google.com/search/blog/2023/02/google-search-and-ai-content?hl=de>



Im Falle einer Dominanz von synthetischen Inhalten wird die Mehrheit der Nutzer:innen von Kommunikationsplattformen nach automatisierter Löschung rufen, da ein ‚unbereinigter‘ Nachrichtenstrom für sie zu viel und zu schwer erkennbaren ‚Unsinn‘ enthält. Damit ergibt sich eine Lizenz zum (immanent politischen) Löschen bzw. zur Unsichtbarmachung von Inhalten im Netz. Den Architekt:innen der nun anzupassenden Social Media-Algorithmen und den Datenaufbereiter:innen für Training und Output der großen Sprachmodelle kommt dann eine nicht hinnehmbare Macht innerhalb der politischen Öffentlichkeit zu:

*Eine KI-basierte Bewältigung des Problems synthetischer Inhalte im Netz ist ein politisches Desaster für die historische Entwicklung des Internet, welches vorgab, die Demokratisierung der Wissenszugänge und des Informationsaustauschs voranzutreiben.*

Die Machtkonzentration auf ein kleines Oligopol ist umso größer, als die Privatisierung von Sprachtechnologien massiv voranschreitet. Als die Chef-Entwicklerin von ChatGPT *Mira Murati* 2018 bei OpenAI startete, war das Unternehmen noch als gemeinnütziges Forschungsinstitut konzipiert: Es ging darum, „sicherzustellen, dass künstliche allgemeine Intelligenz der ganzen Menschheit zugutekommt“. 2019 folgte, wie gewöhnlich bei angehenden *Einhörnern*, die als offene Entwickler:innen-Projekte gestartet sind, die Abkehr vom Non-Profit-Modell. Die mächtigsten KI-Unternehmen halten ihre Forschung unter Verschluss. Das soll verhindern, dass die Konkurrenz von der eigenen Arbeit profitiert. Der Wettlauf um immer umfangreichere Modelle hat schon jetzt dazu geführt, dass nur noch wenige Firmen im Rennen verbleiben werden – neben dem GPT-Entwickler *Open AI* und seiner Microsoft-Nähe sind das *Google*, *Facebook*, *xAI* (neue Firma von Elon Musk), *Amazon* und mit Einschränkung<sup>9</sup> chinesische Anbieter wie *Baidu*. Kleinere, nichtkommerzielle Unternehmen und Universitäten spielen dann so gut wie keine Rolle mehr. Der ökonomische Hintergrund dieser drastisch ausgedünnten Forschungslandschaft: Das Training der Sprachmodelle ist eine Ressourcen-intensive Angelegenheit, welches eine massive Rechenleistung und damit einen beträchtlichen Energieaufwand erfordert. Ein einziger Trainingslauf für das derzeit größte Sprachmodell *GPT-4* kostet aktuell 63 Millionen Dollar.<sup>10</sup>

## AUF DER ÜBERHOLSPUR INS ZEITALTER VON DEEPFAKES

Analog zur (Text-zu-)Texterzeugung per ChatGPT nutzen Programm wie *Midjourney* oder *Stablediffusion* einen ebenfalls auf maschinellem Lernen basierenden (Text-zu-)Bildgenera-

tor, um aus einer textförmigen Bildbeschreibung synthetische Bilder zu erzeugen. Die so erstellten Fake-Bilder einer fiktiven Festnahme von Donald Trump und eines im Rapper-Style verfremdeten Papstes galten dem Feuilleton zu Anfang des Jahres weltweit als ikonische Zeugnisse einer ‚neuen Fake-Ära‘ des Internet. Dabei waren beide lediglich gut gemachte, aber harmlose Bildfälschungen. Andere Formen der sprachmodellbasierten Mis- und Desinformation sind von weit größerer Tragweite.

Auf der *Code Conference 2016* äußerte sich Elon Musk wie folgt zu den Fähigkeiten seines Tesla-Autopiloten: „Ein Model S und Model X können zum jetzigen Zeitpunkt mit größerer Sicherheit autonom fahren als ein Mensch. Und das bereits jetzt.“<sup>11</sup> Elon Musks Anwält:innen behaupteten nun im April 2023 zur Abwehr einer Schadensersatzklage vor Gericht, das Video des Konferenzbeitrags, in dem Musk diese juristische folgenreiche Behauptung aufstellte, sei ein *Deepfake*.<sup>12</sup>

Bereits ein Jahr zuvor argumentierten zwei Angeklagte, die wegen der Kapitolerstürmung im Januar 2021 vor Gericht standen, das Video, welches sie im Kapitol zeige, könne von einer Künstlichen Intelligenz erstellt oder manipuliert worden sein. Täuschung und vorgetäuschte Täuschung gab es schon immer. Diese Debatte hatten wir bereits bei der Popularisierung der Bildbearbeitungssoftware *Photoshop*. Neu ist, dass es keiner handwerklichen Fertigkeiten bedarf und die für alle zugängliche, quasi-instantane Manipulierbarkeit auch Video- und Audio-Sequenzen betrifft.

„Das Hauptproblem ist, dass wir nicht mehr wissen, was Wahrheit ist“ (Margaret Mitchell, ehemalige Google-Mitarbeiterin und jetzige Chefethikerin des KI-Startups *Hugging Face*).

„Das ist genau das, worüber wir uns Sorgen gemacht haben: Wenn wir in das Zeitalter der Deepfakes eintreten, kann jeder die Realität leugnen“, so Hany Farid, ein Experte für digitale Forensik und Professor an der University of California, Berkeley. „Das ist die klassische Lügendividende<sup>13</sup>.“ Eine skeptische Öffentlichkeit wird dazu gebracht, die Echtheit von *echten* Text-, Audio-, und Videodokumenten anzuzweifeln.

Angesichts der beachtlichen Geschwindigkeit, mit der ChatGPT neue Nutzer:innen gewinnt, bedeutet dies einen enormen zukünftigen Schub für das Postfaktische, dessen Hauptwirkungsweise nicht darin besteht, dass selbstkonsistente Parallelwelten von Falscherzählungen für sich ‚Wahrheit‘ im Sinne einer Faktizität reklamieren, sondern dass sie die Frage „Was ist wahr und was ist falsch?“ (zumindest in Teilen des öffentlichen Diskursraums) für unwichtig erklären.

9 Die weitreichende Zensur von Trainingsdaten und Output chinesischer Sprachmodelle stellen wegen der damit verengten Datenbasis ein großen Wettbewerbsnachteil dar. Eine weitere Hürde ist die Hardware. US-Regulierungen verhindern den Export der neuesten KI-Chips von Nvidia u.a. nach China. Diese Chips sind (derzeit) entscheidend für die Entwicklung und Verbesserung von KI-Modellen.

10 <https://the-decoder.de/leaks-zeigen-gpt-4-architektur-datensätze-kosten-und-mehr/>

11 <https://www.vox.com/2016/6/6/11840936/elon-musk-tesla-spacex-mars-full-video-code>

12 <https://www.theguardian.com/technology/2023/apr/27/elon-musks-statements-could-be-deepfakes-tesla-defence-lawyers-tell-court>

13 Die „Lügendividende“ ist ein Begriff, den Robert Chesney und Danielle Citron 2018 in einer Veröffentlichung *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, prägten. Darin beschrieben sie die Herausforderungen, die Deepfakes für die Privatsphäre, die Demokratie und die nationale Sicherheit darstellen. Der zentrale Gedanke darin ist, dass die Allgemeinheit sich bewusst wird, wie einfach es ist, Audio- und Videomaterial zu fälschen, und dass diese Skepsis als Waffe einsetzbar ist: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3213954](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3213954)



Große Sprachmodelle sind geradezu das Ideal des *Bullshitters*, wie der Philosoph Harry Frankfurt, Autor von *On Bullshit*, den Begriff definierte. Bullshitter, so Frankfurt, sind schlimmer als Lügner. Ihnen ist es egal, ob etwas wahr oder falsch ist. Sie interessieren sich nur für die rhetorische Kraft einer Erzählung. Beide Aspekte, das Ignorieren der Frage nach wahr oder falsch, als auch deren aktive Dekonstruktion haben das Potential, Gewissheiten über das Funktionieren von Gesellschaft zu zerlegen. Selbstorganisiertes politisches Engagement von unten droht zu einem Blindflug entlang falscher Annahmen zu werden. Die darauf folgende Ernüchterung befördert den Rückzug ins Private – ein durchaus gewünschter und geförderter Aspekt<sup>14</sup>. Politisch profitieren können von einem hohen Anteil an Misinformation rechte Kräfte, denen an einer gesellschaftlichen Destabilisierung durch wachsende Polarisierung gelegen ist. Steve Bannon (ehemalige Berater von Donald Trump), bezeichnete die Medien immer wieder als Feind, den es niederzuringen gelte. Dazu müsse man „das [mediale] Feld mit Scheiße fluten“.

Je stärker die Akzeptanz verbreiteter Information von ihrem Wahrheitsgehalt entkoppelt ist, desto leichter lässt sich dann auch manipulative Desinformation verbreiten. Falschnachrichten sind meist überraschend und erzeugen deutlich mehr Aufmerksamkeit. Bewusst adressierte Affekte wie *Empörung*, *Angst und Hass* erzeugen bei der Leser:in nachweislich mehr Aktivität und halten die Nutzer:innen damit länger in sozialen Netzwerken als etwa *Freude*, *Zuversicht und Zuneigung*. Dieses Muster wird von der algorithmischen Reichweitensteuerung sozialer Medien erkannt und rückkoppelnd als Trend verstärkt. Über diese statistische Gewichtsverzerrung bevorzugt rechter Beiträge innerhalb politischer Debatten ist z.B. auf Twitter eine deutliche Rechtsverschiebung zu verzeichnen – und das bereits weit vor der Übernahme durch Elon Musk und dessen Neuausrichtung des Algorithmus.<sup>15</sup> Der Siegeszug des Trumpismus nach 2016 ist ein gut untersuchtes Beispiel derartig kontaminierter Diskursräume.

## BEDENKLICHER REDUKTIONISMUS

Suchmaschinen wie *Bing* oder *Google* haben begonnen, ihre KI-Sprachmodelle *GPT-4* bzw. *PaLM* zur zusammenfassenden Weiterverarbeitung gefundener Suchergebnisse zu implementieren. Damit wird die (per Ranking-Algorithmus vorsortierte, aber immerhin noch vorhandene) bisherige Auswahl von Suchergebnissen reduziert auf *ein* leicht konsumierbares Ergebnis auswählbaren Umfangs. Eine enorme Vereinfachung zugunsten einer beträchtlichen Zeitersparnis bei der Internet-suche, aber zulasten einer Vielfalt möglicher (kontroverser) Ergebnisse.

Wer erste Nutzungserfahrungen mit ChatGPT gesammelt hat, wird bei vielen Text-Antworten auf Wissensfragen zu kontro-

versen Themen eine *vermeintliche Ausgewogenheit* feststellen. Einer detailliert dargestellten Mehrheitsmeinung wird ein Zusatz angehängt, dass es dazu durchaus anderslautende Interpretationen gibt. Politische Widersprüche, die in den (sich widersprechenden) Suchergebnissen noch bestanden, werden nun mit einer durch das Sprachmodell vordefinierten Diversitätstiefe aufgelöst. Dadurch ergibt sich ein politisch bedenklicher Reduktionismus, der wohlgerichtet auf einem Sprachmodell(!) basiert – also nicht wissensbasiert ist, sondern mangels Verständnis von Begriffsbedeutungen rein statistisch bestimmt ist.

Diese ‚kritischen‘ Anmerkungen werden zukünftig zur sogenannten Medienkompetenz gezählt werden und bedeutungslos (wie alles in der Welt der Sprachmodelle) verhallen. Wer klickt noch in schier endlosen Suchergebnislisten herum, wenn die Suche bei Google oder Bing das ‚Wichtigste‘ für uns zusammenfasst?<sup>16</sup>

## VERGANGENHEIT IN DIE ZUKUNFT PROJIZIERT

ChatGPT ist ein *stochastischer Papagei*, der (willkürlich) Sequenzen sprachlicher Formen zusammenfügt, die er in seinen umfangreichen Trainingsdaten beobachtet hat, und zwar auf der Grundlage probabilistischer Informationen darüber, wie sie kombiniert werden, aber ohne jeglichen Bezug zu deren Bedeutung. Ein solcher Papagei reproduziert und verstärkt dabei nicht nur den *Bias* von verzerrten Trainingsdaten, sondern auch hegemoniale Weltanschauungen dieser Trainingsdaten. Gesellschaftliche Verhältnisse aus der Vergangenheit der Trainingsdaten werden in die Zukunft verstetigt. Die den Sprachmodellen immanente Rekombination statistisch dominanter Wissensbeiträge der Trainingsdaten wirkt die Verhältnisse konservierend und stabilisierend – ein sogenannter *value lock*, das Einrasten von Werten im Sinne einer politischen Stagnation droht.<sup>17</sup>

Die Bedingungen einer solchen Hegemonieverstärkung werden leider nur marginal gesellschaftlich (mit-)bestimmt. Das komplexe System aus Trainingsdatenaufbereitung, Parameterjustierung des Sprachmodells und nachträglicher Zensur des Outputs (allesamt unter der Kontrolle profitorientierter Privatunternehmen) bestimmen das Gewicht von *neuen* Wissensbeiträgen. Damit liegt die hohe Hürde einer ausreichenden statistischen Relevanz emanzipatorischer Debattenbeiträge außerhalb einer demokratisch verfassten, gesellschaftlichen Mitbestimmung. Angesichts eines deutlichen politischen Drifts nach rechts maßgeblicher Technokrat:innen des KI-Geschäftsmodells (wie Sam Altman, Elon Musk, Peter Thiel, ...) sind das keine hinnehmbaren Voraussetzungen für eine gesellschaftlich progressive Entwicklung.

14 Beispielhaft steht hierfür die Politik von Vladislav Surkov, Spindoktor Putins. <https://www.nytimes.com/2014/12/12/opinion/russias-ideology-there-is-no-truth.html>

15 Aral, S. (2018): The spread of true and false news, *Science* 359, 1146-1151(2018), <https://www.science.org/doi/10.1126/science.aap9559>

16 Amazons Spracherkennungs- und -steuerungssoftware *Alexa* befördert ebenfalls diesen Reduktionismus, da sich niemand von *Alexa* eine längere Liste von Sucheinträgen *vorlesen* lassen möchte. Wegen der oft wenig hilfreichen Sprachausgabe *eines* weit oben gelisteten Treffers ist die Googlesuche über *Alexa* allerdings weit weniger beliebt.

17 Bender et al: *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* (2021) <https://dl.acm.org/doi/pdf/10.1145/3442188.3445922>



## DIVERSITÄTSVERLUST UND RECHTSDRIFT

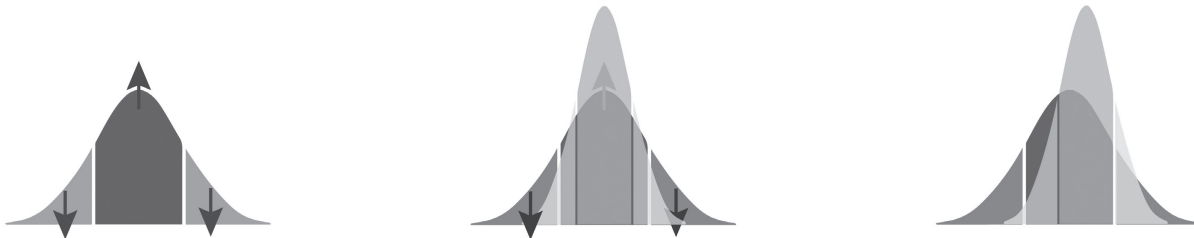
Die intrinsische Hegemonieverstärkung großer Sprachmodelle über ein selbstverstärkendes Wiederverdauen des eigenen Outputs als Input für das nächste Training des Modells bedeutet einen Verlust an Meinungsvielfalt (siehe Abbildung links+Mitte). Zusätzlich führt die zuvor erwähnte Bevorzugung (in Reichweite und Verbreitungsgeschwindigkeit) verschwörerischer und (rechts-)populistischer Inhalte in den sozialen Medien zu einer politisch rechts gerichteten Verzerrung in den Trainingsdaten der nächsten Generation von Sprachmodellen. Dadurch erwarten wir in der Überlagerung beider Effekte einen *rechtslastigen Diversitätsverlust* (Abbildung rechts).

Eine solche Deformation öffentlicher Diskursräume über die Wechselwirkung großer Sprachmodelle mit den sozialen Medien zugunsten einer a) hegemonial-konservativen Meinungseinfalt und b) einer zentralen Machtposition eines Technologie-Oligopols, welches die Verzerrung algorithmisch codiert, muss aus der Sicht einer progressiven Position als Rückschritt und als politische Sackgasse zurückgewiesen werden. Die Unzulänglichkeit der sich neu ergebenden Informationsinfrastruktur bestehend aus *großen Sprachmodellen + Social Media-Plattformen + Suchalgorithmen* wird sich wohl kaum durch eine gesellschaftlich legitimierte, besser ausbalancierte Inhalte-Moderation abfedern lassen.

Ein emanzipatorischer Zugang zu einer grundlegenden Technologiekritik darf nicht auf der Ebene kosmetischer Korrekturen einer zahnlosen „Technikfolgenabschätzung“ verharren. Anstatt große Sprachmodelle unkritisch als unausweichlichen technologischen Fortschritt hinzunehmen, sollten wir die Frage aufwerfen, *ob*, und nicht *wie*, wir diese Technologien überhaupt gesellschaftlich akzeptieren sollten. Die langfristigen gesellschaftlichen Folgen dieser Modelle innerhalb einer dominanten KI-Empfehlungs- und Entscheidungs-Assistenz insbesondere für den Prozess der politischen Willensbildung, tauchen in einer nun allseits geforderten *technischen Sicherheitsforschung von KI-Systemen* als ‚schwer zu quantifizieren‘ gar nicht auf.<sup>18</sup>

Wir sollten unsere Haltung in Bezug auf die politische Schädigung KI-basierter Sprachmodelle ausrichten an unserer Haltung gegenüber KI-basierten, autonomen Waffensystemen: Warum sollte eine Gesellschaft einen derart rückwärts gewandten technologischen ‚Fortschritt‘ hinnehmen?

„Marx sagt, die Revolutionen sind die Lokomotive der Weltgeschichte. Aber vielleicht ist dem gänzlich anders. Vielleicht sind die Revolutionen der Griff des in diesem Zuge reisenden Menschengeschlechts nach der Notbremse.“<sup>19</sup>  
(Walter Benjamin)



18 Siehe dazu: *Nudging – die politische Dimension psychotechnologischer Assistenz*, DISS-Journal#43 (2022) <http://www.diss-duisburg.de/2022/08/nudging-die-politische-dimension-psychotechnologischer-assistenz/>

19 Walter Benjamin: MS 1100. In: Ders.: *Gesammelte Schriften*. I, hg. v. R. Tiedemann und H. Schweppenhäuser, Frankfurt am Main: Suhrkamp 1974, 1232.

