FUTURE LAW
WORKING PAPERS 2024 • 2

universität innsbruck
Institut für Theorie und Zukunft des Rechts

MATTHIAS C. KETTEMANN • MALTE KRAMME
CLARA RAUCHEGGER • CAROLINE VOITHOFER
EDITORS

Regulierung von ChatGPT & Co.: Neue Regeln für eine neue Technik?

MALTE KRAMME

FREIBERG

JÄNNER 2024 • ZUKUNFTSRECHT@UIBK.AC.AT

FUTURE LAW WORKING PAPERS 2024 • 2

MATTHIAS C. KETTEMANN • MALTE KRAMME
CLARA RAUCHEGGER • CAROLINE VOITHOFER
EDITORS

universität innsbruck
Institut für Theorie und Zukunft des Rechts

The Future Law Working Papers was established in 2022 to offer a forum for cutting-edge research on legal topics connected to the challenges of the future. As the German Constitutional Court recently ruled, we have to act today to save the freedoms of tomorrow. Similarly, the Future Law Working Papers series hosts research that tackles difficult questions and provides challenging, and at times uncomfortable, answers, to the question of how to design good normative frameworks to ensure that rights and obligations are spread fairly within societies and between societies, in this generation and the next. The series is open for interdisciplinary papers with a normative twist and the editors encourage creative thinking. If you are interested in contributing, please send an email to the editors at zukunftsrecht@uibk.ac.at. Submissions are welcome in English and German.

The series is edited by the senior members of the Department of Legal Theory and the Future of Law at the University of Innsbruck, Matthias C. Kettemann, Malte Kramme, Clara Rauchegger and Caroline Voithofer.

Founded in 2019 as the tenth department of the law faculty, the Department of Legal Theory and Future of Law at the University of Innsbruck (ITZR) investigates how law can make individuals as well as society, states as well as Europe "fit" for the future and if and how law has to change in order to meet future challenges. This includes the preservation of freedom spaces as well as natural resources in an intergenerational perspective, the safeguarding of societal cohesion in times of technologically fueled value change, the normative framing of sustainable digitization and digitized sustainability, and the breaking through of traditional legal structures of domination and thought with a view to rediscovering the emancipatory element of law against law.

Publisher: Institut für Theorie und Zukunft des Rechts, Universität Innsbruck
Innrain 15, 6020 Innsbruck
Univ.-Prof. Mag. Dr. Matthias C. Kettemann, LL.M. (Harvard)
Univ.-Prof. Dr. Malte Kramme
Ass. Prof.in MMag.a Dr. Clara Rauchegger, LL.M. (Cambridge)
Univ.-Ass.in MMag.a Dr.in Caroline Voithofer

Regulierung von ChatGPT & Co.: Neue Regeln für eine neue Technik?

Prof. Dr. Malte Kramme

TU Bergakademie Freiberg
Associated Scholar, Digital Science Center (DiSC),
Universität Innsbruck

Regulierung von ChatGPT & Co.: Neue Regeln für eine neue Technik?*

Malte Kramme

I. Vorsorge vor Gefahren einer disruptiven Technik

Als ChatGPT im Herbst 2022 veröffentlicht wurde, hat die Leistungsfähigkeit dieses KI-Systems viele erstaunt, insbesondere in Wirtschaft, (Rechts-)Wissenschaft und Politik. Zwar werden KI-Systeme schon seit einigen Jahren in bestimmten Bereichen eingesetzt. Es gab auch allerlei Prognosen darüber, was KI eines Tages leisten kann und wie sie unser Leben und Arbeiten verändern wird. Für viele waren dies jedoch eher Gedanken an eine Zukunft, die plötzlich viel näher zu sein scheint. Die Entwicklungsgeschwindigkeit von Systemen wie ChatGPT ist nach wie vor extrem hoch. Hier nur ein eindrucksvolles Beispiel für Juristinnen und Juristen: ChatGPT in der Version 3.5 scheiterte noch im Dezember 2022 am US Uniform Bar Exam. Nur zwei Monate später hat ChatGPT in der Version 4.0 bereits 90% seiner menschlichen Kommilitoninnen und Kommilitonen hinter sich gelassen.[1]

Die Rechtswissenschaft ist gefordert, diese neue Entwicklung in regulatorische Bahnen zu lenken, die dazu führen, dass sich die neue Technik zum Wohl von Mensch und Natur entwickelt.

Doch welche Regeln sollen für KI gelten?

Dieser Beitrag geht auf vier Bereiche ein:
1. die Begrenzung der Macht der Anbieter generativer KI
2. die Sicherung der Qualität der Ergebnisse und die Verhütung von Missbrauch

* Großer Dank gilt Prof. Dr. Thomas Kopinski, Professor für Machine Learning und Data Science an der Fachhochschule Südwestfalen, für viele wertvolle Hinweise.

[1] Kimmel, ChatGPT Passed the Uniform Bar Examination: Is Artificial Intelligence Smart Enough to be a Lawyer?, International and Comparative Law Review, University of Miami, School of Law, https://international-and-comparative-law-review.law.miami.edu/chatgpt-passed-the-uniform-bar-examination-is-artificial-intelligence-smart-enough-to-be-a-lawyer/, Abruf: 24.7.2023; https://www.abajournal.com/web/article/latest-version-of-chatgpt-aces-the-bar-exam-with-score-in-90th-percentile, Abruf: 24.7.2023. Zu weiteren Tests, die ChatGPT erfolgreich absolviert hat: https://www.businessinsider.com/list-here-are-the-exams-chatgpt-has-passed-so-far-2023-1, Abruf: 24.7.2023.

3. den Schutz der Privatsphäre
4. den Schutz der natürlichen Lebensgrundlagen

Es gibt natürlich noch viele andere Themen, auf die Antworten gegeben werden müssen: unter anderem die Frage nach der Haftung, die Frage, wie man Diskriminierung durch KI unterbindet sowie die Frage nach dem Schutz des geistigen Eigentums.[2] Diese Fragen sind nicht weniger wichtig. Die hier aufgeworfenen Fragen erscheinen nur aus heutiger Sicht aus besonders dringend.

Es ist ein altbekanntes Phänomen, dass die Regulierung technischer Entwicklung hinterherhinkt. Das mag auch manchmal gut sein, um Innovation nicht durch Regulierung zu ersticken. Wie sehr Künstliche Intelligenz künftig unser Leben und Wirtschaften beeinflussen wird, können wir, wie bei allen technischen Entwicklungen, derzeit zwar nicht absehen, dennoch können wir uns in diesem Fall eine abwartende Haltung nicht erlauben. Denn es besteht die Gefahr, dass in der Zwischenzeit Fakten geschaffen werden, die durch eine zu späte Regulierung nicht mehr korrigiert werden können. Geraten etwa KI-Systeme, die keinerlei Vorkehrungen gegen einen missbräuchlichen Einsatz aufweisen, in die Hände des organisierten Verbrechens, lässt sich die dadurch entstehende Bedrohungslage nicht mehr wegregulieren. In gleicher Weise wird es kaum möglich sein, die KI-Systeme Wissen, das unsere Privatsphäre betrifft, wieder vergessen zu lassen.

Wie ernst das Gefahrenpotenzial von generativen KI-Systemen selbst von ihren Entwicklerinnen und Entwicklern genommen wird, zeigen Aufrufe aus der Branche: In einem offenen Brief, der Ende März 2023 veröffentlicht wurde, wurde ein Entwicklungsmoratorium für sechs Monate gefordert, um in der Zwischenzeit Sicherheitsprotokolle zu entwickeln.[3] Ende Mai 2023 veröffentlichte das Center for AI Safety (CAIS) ein Statement, das mit folgendem Wortlaut die Anerkennung von KI-Systemen als potentieller Gefahr für menschliches (Über-)Leben forderte:

[2] S. hierzu de la Durantaye, »Garbage in, garbage out« – Die Regulierung generativer KI durch Urheberrecht, ZUM 2023, 645; Krone, Urheberrechtlicher Schutz von ChatGPT-Texten?, RDi 2023, 117; Möller-Klapperich, ChatGPT & Co. – aus der Perspektive der Rechtswissenschaft, NJ 2023, 144 (147); Hoeren, „Geistiges Eigentum" ist tot – lang lebe ChatGPT, MMR 2023, 81; Olbrich/Bongers/Pampel, Urheberrechtsschutz für Kunstwerke künstlicher Intelligenz?, GRR 2022, 870.

[3] Pause Giant AI Experiments: An Open Letter: We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4, s. https://futureoflife.org/open-letter/pause-giant-ai-experiments/, Abruf: 18.12.2023. S. hierzu auch Wirtschaftswoche, Die Geister, die wir riefen, 16/2023, S. 15.

„Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."[4]

Das Statement wurde unter anderem von Sam Altman, dem (kurzzeitig einmal abberufenen) CEO von Open-AI sowie von Bill Gates, einem der Hauptanteilseigner an Open-AIs Mutterkonzern Microsoft, unterzeichnet. Sieben Unternehmen, die KI-Systeme entwickeln, haben sich in einer Selbstverpflichtung gegenüber dem US-amerikanischen Präsidenten verpflichtet, KI-Systeme vor ihrem Inverkehrbringen auf Risiken zu testen und KI-generierte Inhalte zu kennzeichnen.[5] Aber auch eine gesetzliche Regulierung zeichnet sich nicht nur in der EU mit dem AI Act und der AI Liability Directive ab, sondern auch in den USA[6] sowie in China,[7] Kanada[8] und Brasilien.[9] Zudem wurde auf dem G7 Gipfel im Oktober 2023 mit einer entsprechenden Erklärung der Hiroshima AI Process in Gang gesetzt.[10]

Es ist zwar keinesfalls ein neues Phänomen, dass ich im Markt etablierte Akteure eine strenge Regulierung wünschen, weil sie sich dadurch die Schaffung von Marktzutrittsschranken erhoffen. Der erste Reflex, den Markt unreguliert zu belassen, ist aber dennoch oft nicht die richtige Lösung.[11] Hier erst recht nicht: Wenn nicht zweifelsfrei feststeht, dass die Unterzeichner der oben genannten Erklärung mit ihrer Warnung nicht völlig daneben liegen, muss ein dem Vorsorgeprinzip verpflichteter Gesetzgeber das in seiner Macht stehende tun, um die konkreten Gefahren zu identifizieren und ihre Eintrittswahrscheinlichkeit zu reduzieren. Aber auch wenn die Politik bald erste Maßnahmen unternimmt, wird die Regulierung von generativer künstlicher Intelligenz

[4] Statement on AI Risk, AI experts and public figures express their concern about AI risk, s. https://www.safe.ai/statement-on-ai-risk#sign, Abruf: 18.12.2023.

[5] Ensuring Safe, Secure, and Trustworthy AI, https://www.whitehouse.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf, Abruf: 18.12.2023; s. auch https://www.handelsblatt.com/dpa/ki-firmen-versprechen-biden-mehr-umsicht/29273812.html, Abruf: 18.12.23.

[6] S. Blueprint for an AI Bill of Rights, Oktober 2022, s.https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf, Abruf: 18.12.2023. Für Hintergründe, s. Vasel, Künstliche Intelligenz und die Notwendigkeit agiler Regulierung, NVwZ 2023, 1298 (1299).

[7] S. Roberts/Hine, The future of AI policy in China in: East Asia Forum, https://www.eastasiaforum.org/2023/09/27/the-future-of-ai-policy-in-china/, Abruf: 18.12.2023; s. auch Vasel (Fn. 6), NVwZ 2023, 1298 (1299).

[8] https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act, Abruf: 18.12.2023.

[9] S. zu weltweiten Regulierungsbestrebungen Hacker/Berz, Der AI Act der Europäischen Union – Überblick, Kritik und Ausblick, ZRP 2023, 226.

[10] https://digital-strategy.ec.europa.eu/de/library/g7-leaders-statement-hiroshima-ai-process, Abruf: 18.12.2023.

[11] Vgl. Pichai, Building AI responsibly is the only race that really matters, Financial Times v. 22.5.2023, https://www.ft.com/content/8be1a975-e5e0-417d-af51-78af17ef4b79, Abruf: 18.12.2023; Vasel (Fn. 6), NVwZ 2023, 1298.

und ihren Wirkungen in den kommenden Jahren und Jahrzehnten absehbar eine politische und rechtswissenschaftliche Daueraufgabe bleiben.

Bevor wir uns den oben angesprochenen Punkten zuwenden, soll kurz die Funktionsweise von KI-Systemen wie ChatGPT angerissen werden. Dabei geht es weniger um eine Erläuterung technischer Hintergründe als vielmehr darum klarzumachen, was Gegenstand der dann folgenden Überlegungen ist und was nicht.

II. Was steckt hinter ChatGPT & Co.?

Gegenstand dieses Beitrags sind sog. large language models, die Kerntechnologie hinter Systemen wie ChatGPT.[12] Ein large language model ist ein generatives KI-System, das mit großen Textmengen trainiert wurde, um Text als Output zu generieren. Der Begriff „KI-System" wird hier vor allem in Abgrenzung zu klassischer Software verwendet. Klassische Software ist rein regelbasiert: Der Dateninput wird anhand von Regeln verarbeitet, die vom Programmierer oder von der Programmiererin bis ins letzte Detail vorgegeben wurden, wie zB in einem Taschenrechner.

Demgegenüber fasst man unter den Begriff der Künstlichen Intelligenz gemeinhin solche Systeme, die nicht nur programmierte Regeln abspulen, sondern auch nach Regeln funktionieren, die sich beim maschinellen Lernen ergeben haben. Die Lerndaten sind extrem umfangreich. GPT 3.5 soll an 400 bis 500 Milliarden Token (Wörter oder Teilelemente von Wörtern) gelernt haben.[13] Das Lernen erfolgt zum Teil autonom, in einer Art semi-überwachten Lernens, innerhalb dessen der Mensch noch einen Teil des Prozesses beobachtet. Da aufgrund der Funktionsweise der Algorithmen, im Kern fußen large language models auf der Transformer-Architektur (das T in GPT), die Entscheidungsfindung dieser nicht transparent ist, resultieren hieraus Black-Box-Probleme.[14] Large language models lernen insbesondere, in welchen Kontexten einzelne Wörter mit welcher Wahrscheinlichkeit vorkommen. Im Ergebnis weist es einzelnen Wörtern einen Vektorwert zu, also Koordinaten in einem hochdimensionalen Raum. Wörter bekommen damit eine Adresse und haben Nachbarn. Mann und Frau liegen zB

[12] Die folgende Darstellung ist sehr stark vereinfacht und dient der Veranschaulichung für einen juristisch vorgebildeten Leserkreis. Wesentlich ausführlicher zu den technischen Hintergründen Bommasani et al., S. 73 ff., arXiv:2108.07258, https://doi.org/10.48550/arXiv.2108.07258.

[13] Beck, Gespräche führen mit ChatGPT: So lernt die KI von uns https://www.swr.de/wissen/chatbots-wie-funktioniert-chat-gpt-100.html, Abruf: 30.12.2023.

[14] S. etwa Apel/Kaulartz, Rechtlicher Schutz von Machine Learning-Modellen, RDi 2020, 24 Rn 10, 41; Möslein, Die normative Kraft des Ethischen - Ein Fallbeispiel zur Effektivität von Leitlinien für Künstliche Intelligenz, RDi 2020, 34 Rn 9.

nah beieinander. Beide sind Menschen. Sie sind weiter entfernt von Wörtern wie Haus oder Baum. Der Abstand von Wörtern zueinander lässt sich messen, was es wiederum ermöglicht, Wahrscheinlichkeiten zu berechnen, mit der bestimmte Wörter in einem Kontext zusammen auftauchen.

Bei Systemen wie ChatGPT handelt es sich um generative KI-Modelle. Alle KI-Systeme, genauso wie auch regelbasierte Software, verarbeiten einen Input und erzeugen als Ergebnis einen Output. Viele der uns bisher mehr oder minder vertrauten KI-Systeme beschränkten sich darauf, zu klassifizieren, also Entscheidungen zu treffen, etwa bei der Bilderkennung Hunde von Katzen zu unterscheiden. Generative KI-Systeme haben einen wesentlich umfangreicheren Output.[15] Bei large language models wie ChatGPT besteht dieser Output in Text, bei dem es sich etwa um Zusammenfassungen, Erläuterungen oder Vorschläge für Handlungsstrategien handeln kann. Andere KI-Systeme generieren Musik oder Grafiken.[16]

Warum hat die Entwicklung von large language models in der letzten Zeit so an Fahrt aufgenommen? Die Verarbeitung von Sprache stellt Computersysteme vor große Herausforderungen. Menschen können gut mit Sprache umgehen; Computer können besser rechnen. Wie sehr sich Computer damit schwertun, den Kontext zu halten, wird etwa bei den Wortvorschlägen deutlich, wenn man mit dem Handy, allein unter Nutzung eben dieser Wortvorschläge eine Nachricht schreiben möchte.

2017 präsentierten bei Google beschäftigte Wissenschaftlerinnen und Wissenschaftler aber mit dem sog. Attention-Mechanismus, zentraler Bestandteil des o.g. Transformer-Modells, das mit diesem Problem besser zurechtkommt,[17] indem es die Vektoren abhängig vom inputspezifischen Kontext verschiebt und in der Lage ist, Kontexte über lange Distanzen zu halten (hier: große Satz- oder Textkonstrukte). Ausgehend von der Anfrage lassen sich nun besser sinnstiftende Texte produzieren. Dabei macht es für das System kaum einen Unterschied, um welche Sprache es geht. Damit sind nicht nur Sprachen wie Deutsch oder Englisch gemeint, sondern etwa auch die Sprache der chemischen Formeln. Manche der Systeme behandeln auch Musik oder Bilder wie Sprache und können so aufgrund sprachlicher Befehle Grafiken erzeugen oder Musik komponieren.[18]

---

[15] S. Hacker/Berz (Fn. 9), ZRP 2023, 226 (228).

[16] Andere generative AI-Modelle sind etwa DALL-E, Stable Diffusion oder Midjourney (jeweils Bilderstellung) und Jukebox (Musik).

[17] Vaswani et al., Attention Is All You Need, arXiv: 1706.03762.

[18] S. etwa die oben genannten KI-Systeme, die aus Sprachbefehlen Bilder erstellen können.

Der Output muss aber nicht immer wahr sein. Denn während Menschen zunächst einen bestimmten Inhalt vor Augen haben, den sie kommunizieren wollen, wählen large language models Schritt für Schritt das passende Wort. Auch wenn die Texte teilweise ein großes Selbstbewusstsein vermuten lassen, darf nicht vergessen werden, dass ChatGPT derzeit kein Bewusstsein für die Semantik hat, die es hervorbringt.[19]

III. Begrenzung der Macht der Anbieter

Das Mantra des wissenschaftlichen Zeitalters, das Francis Bacon zugesprochen wird, lautet: „Wissen ist Macht."[20] Dieser Ausspruch ist nach wie vor wahr. Zwar heißt es mittlerweile oft: Es kommt darauf an zu wissen, wo etwas steht. Doch das allein genügt noch nicht. Dieses Wissen kann man sich nur zu Nutze machen, wenn man es sich auch aneignet. Das wird sich in kategorialer Weise ändern: Für immer mehr Aufgaben wird es ausreichen, Zugang zu einem KI-System zu haben. Noch besser ist es nur, wenn man das KI-System betreibt, weil man damit den Zugang zu ihm kontrollieren kann und selbst fortlaufend neue Märkte besetzen kann, in denen es Einsatzzwecke für KI gibt.

In einem marktwirtschaftlichen System ist das wirkungsvollste Instrument zur Beschränkung der Macht einzelner Akteure die Herstellung eines wirksamen Wettbewerbs auf allen Stufen. Dies gilt zunächst unter den Anbietern von KI-Systemen. Daher ist zu klären, ob der bestehende regulatorische Rahmen dafür ausreichend ist. Bevor auf die einzelnen Instrumente eingegangen wird, wird die regulatorische Herausforderung näher skizziert.

1. Ausgangslage

a) Wettbewerb zwischen Anbietern von KI-Systemen

Derzeit sieht es nicht unbedingt so aus, dass sich oligopol- oder gar monopolartige Strukturen entwickeln, wie wir das aus anderen Sektoren des Digitalwirtschaftsrechts kennen. Der Stanford AI Index wies für das Jahr 2022 die Veröffentlichung von

[19] Gleichwohl hängt diese Schlussfolgerung maßgeblich davon ab, was man unter Bewusstsein versteht, s, hierzu Kleiner/Lorenz, Ab wann kann man einer KI wie Chat-GPT ein Bewusstsein zusprechen?, https://www.spektrum.de/news/hat-kuenstliche-intelligenz-wie-chatgpt-ein-bewusstsein/2193018, Abruf: 30.12.2023.

[20] Die dem am nächsten kommende Formulierung im Werk Bacons findet sich in den Meditationes sacrae von 1597: „Nam et ipsa scientia potestas est." S. hierzu Wildner, Big Data: Wissen ist Macht, Gesundheitswesen 2015; 77(08/09): 531-532; Azamfirei, Knowledge Is Power, The Journal of Critical Care Medicine, vol.2, no.2, 2016, pp.65-66. https://doi.org/10.1515/jccm-2016-0014.

35 signifikanten KI-Modellen aus.[21] Darunter sind auch open source-Projekte, bei denen auch das Training dezentral erfolgt. Aber: Die Entwicklung neuer KI-Modelle, insbesondere von large language models wird zunehmend aufwendiger und teurer.[22] Daher müssen die Wettbewerbsbehörden hier wachsam bleiben, damit ein möglichst breiter Wettbewerb bestehen bleibt.

b) Wettbewerb zwischen Anbietern und gewerblichen Kunden

Als wesentlich problematischer könnte sich die Sicherstellung eines fairen Wettbewerbs zwischen den Anbietern von KI-Systemen und ihren gewerblichen Kunden erweisen. Das sind vor allem Unternehmen, die bereits jetzt ein funktionierendes Geschäftsmodell haben. Der Wettbewerbsdruck wird solche Unternehmen dazu zwingen, KI-Systeme in ganz verschiedenen Zusammenhängen einzusetzen.

Teilweise wird für die Kunden dabei erforderlich sein, die KI-Systeme der großen Anbieter für ihre spezifischen Einsatzzwecke zu adaptieren, etwa indem sie anhand von fachspezifischen Daten angelernt werden. Zum Beispiel ist es denkbar, dass juristische Verlage oder große law firms KI-Systeme auf die Lösung bestimmter rechtlicher Fragestellungen trainieren.[23] Unternehmern werden also zu Betreibern von speziell für ihre Zwecke adaptierten Allgemein-KI-Systemen.[24] Dies wird aber nicht immer notwendig sein. Teilweise werden die gewerblichen Kunden aber auch bloß Eingaben in vom Anbieter bereitgestellten KI-Systemen vornehmen, sich also auf eine bloße Nutzerrolle beschränken.[25]

Im Wettbewerbsverhältnis zwischen Anbietern und gewerblichen Kunden droht sich das zu beschleunigen, was wir auch bisher schon als eine Ausprägung der Digitalwirtschaft erleben: Die großen Tech-Konzerne bieten gewerblichen Kunden nicht nur unverzichtbare Dienstleistungen an, sondern dringen auch in deren Märkte vor. Das Wettbewerbsrecht beschreibt dies mit dem wirtschaftswissenschaftlichen Begriff der

---

[21] Stanford AI Index 2023, S. 11 https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf, Abruf: 30.12.2023.

[22] Stanford AI Index 2023, S. 11, 62, https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf, Abruf: 30.12.2023.

[23] S. hierzu auch Möller-Klapperich (Fn. 2), NJ 2023, 144 (146).

[24] S. zu diesem fine-tuning Hacker/Berz (Fn. 9), ZRP 2023, 226 (228).

[25] Zu einer Erfassung der Akteure in der Wertschöpfungskette von KI-Systemen bietet sich diese Unterscheid zwischen Anbietern, Betreibern, Nutzern und Betroffenen an, s. mit ähnlicher Terminologie Hacker/Engel/Mauer, Regulating ChatGPT and Other Large AI Models, 2023, S. 7, https://arxiv.org/abs/2302.02337; Hacker, Die Regulierung von ChatGPT et al. – ein europäisches Trauerspiel, GRUR 2023, 289 (290). Eine vergleichbare Differenzierung ist bereits im DMA mit der Unterscheidung von zentralen Plattformdiensten (insbes. Gatekeepern), gewerblichen Nutzern und Endnutzern angelegt.

„vertikalen Integration“, [26] der in diesem Zusammenhang aber einen etwas euphemisierenden Klang hat.

Was könnten Microsoft oder Google unternehmen, um das Geschäft der Anwaltskanzlei oder des Verlags an sich zu reißen, nachdem diese das KI-System für die Lösung rechtlicher Fälle trainiert hat? Naheliegend ist es, auf das Knowhow, also die Trainingsdaten oder gleich die weiterentwickelte KI, zurückzugreifen. Und in einem zweiten Schritt könnten schrittweise die Preise für die Nutzung des KI-Systems durch die Anwaltskanzlei erhöht werden. Der Regulierung der Nutzung von Daten zum Training von KI-Systemen kommt damit eine imminent wichtige Bedeutung bei der Sicherstellung eines fairen Wettbewerbs zu.

c) Sicherstellung des Wettbewerbs zwischen gewerblichen Kunden

Das Problem der Verwendung von Trainingsdaten stellt sich aber nicht nur im Wettbewerb zwischen einem KI-Anbieter und einem gewerblichen Kunden, der das KI-System adaptiert. Die Verwendung von Eingabedaten zu Trainingszwecken ist auch für die gewerblichen Kunden, die die Eingaben vornehmen, ein Problem.

Ein spektakulärer Fall dazu ging im Frühjahr 2023 durch die Presse: Mitarbeiter von Samsung hatten Programmcode, der Halbleiter steuert, für eine Fehleranalyse in ChatGPT gepostet.[27] Derartiger Programmcode ist normalerweise ein streng gehütetes Betriebsgeheimnis. Nun kennt ihn aber ChatGPT, so dass alle Nutzer, auch Konkurrenten von Samsung im Halbleitergeschäft, davon profitieren können.

Die Wahrung der Vertraulichkeit von Eingabedaten ist aber nicht nur zum Schutz von Betriebsgeheimnissen wichtig, sondern auch zum Schutz anderer Berufsgeheimnisse wie dem Anwaltsgeheimnis. Stellen Sie sich die folgende Situation vor: Zwei Anwaltskanzleien verwenden ein KI-System ein und desselben Anbieters. Beide Kanzleien vertreten in einem Rechtstreit die sich gegenüberstehenden Parteien und nutzen das KI-System bei der Erstellung ihrer Schriftsätze. Hier besteht die Gefahr, dass das KI-System, das beide Entwurfsfassungen der Schriftsätze erstellt hat, die Argumente der jeweils anderen Partei antizipiert, also wechselseitig Parteiverrat begeht.

2. Regulatorischer Rahmen

[26] S. etwa § 19 Abs. 1 Nr. 3 dGWB.

[27] Chip, Neuer Samsung-Ärger: Mitarbeiter tragen vertrauliche Daten in ChatGPT ein, s. https://www.chip.de/news/Aerger-fuer-Samsung-Vertrauliche-Daten-landen-in-ChatGPT_184732808.html, Abruf: 18.12.2023.

a) Digital Markets Act

Was hat der regulatorische Rahmen diesen Szenarien entgegenzusetzen? Wegen seiner EU-weiten Geltung und seiner Zielsetzung, die Marktmacht der großen Player der Digitalwirtschaft zu begrenzen, fällt der Blick zunächst auf den Digital Markets Act (DMA). Zunächst die Frage: Ist der DMA auf Anbieter von KI-Systemen überhaupt anwendbar? Der DMA gilt gem. Art. 1 Abs. 2 für Anbieter zentraler Plattformdienste, die von der Kommission aufgrund ihrer Marktmacht als Torwächter benannt worden sind. Zwar denkt man bei KI-Systemen nicht direkt an eine klassische Plattform. Doch der Anwendungsbereich ist sehr weit. Erfasst sind nämlich auch Anbieter von virtuellen Assistenten. Darunter dürften Chatsysteme wie ChatGPT ohne weiteres fallen, erst recht dann, wenn sie mit einem Suchmaschinendienst verknüpft sind.[28] Denn Online-Suchmaschinen gelten gem. Art. 2 Nr. 2 lit. b ebenfalls als zentrale Plattformdienste im Sinne des DMA.

Der Normgeber trifft eine für die hier untersuchte Frage ganz zentrale Regelung: Gem. Art. 6 Abs. 1 lit. a DMA ist der Zugriff auf Trainingsdaten untersagt; außer es handelt sich um öffentlich zugängliche Daten. Es wird aber ausdrücklich klargestellt, dass die Eingabe der Daten durch den Endnutzer auf der Plattform die Daten nicht zu öffentlich verfügbaren Daten macht. Es handelt sich hier auch um eine eindeutig wettbewerbsschützende Vorschrift und nicht um eine datenschützende Vorschrift: Denn eine Einwilligung des Endnutzers kann das Verbot nicht überwinden.[29] Es kommt daher auch nicht darauf an, ob die Daten einen Personenbezug haben.[30]

b) Nutzung von Kundendaten als Missbrauch von Marktmacht (§ 19a dGWB)

Auch das nationale Wettbewerbsrecht kann dem Treiben der Anbieter von KI-Systemen bei der Nutzung von Trainingsdaten entgegengesetzt werden. Mit der 10. GWB-Novelle hat der deutsche Gesetzgeber 2021 § 19a GWB eingeführt. Die dort niedergelegten

[28] Als erster Suchmaschinenanbieter hat Microsoft seine Suchmaschine „bing" mit einem KI-Chatbot-System verknüpft. Auf die Anwendbarkeit des DSA bei Verbindung eines KI-Systems mit einer Suchmaschine weist auch Möller-Klapperich (Fn. 2), NJ 2023, 144 (148) hin.

[29] Hacker, KI und DMA – Zugang, Transparenz und Fairness für KI-Modelle in der digitalen Wirtschaft, GRUR 2022, 1278 (1280).

[30] Daneben enthält der DMA auch datenschutzspezifische Vorschriften, die die Verwendung von Daten zu Trainingszwecken beschränken. Art. 5 Abs. 2 DMA verbietet es Torwächtern personenbezogene Daten von Endnutzern, die Dienste Dritter über die Plattform des Torwächters nutzen, für Online-Werbedienste zu verwenden (lit. a) sowie personenbezogene Daten von Endnutzern mit ebensolchen Daten aus anderen Diensten zusammenzuführen (lit. b bis lit. d). Wegen der datenschutzrechtlichen Stoßrichtung, können die Verbote aber durch eine Einwilligung des Endnutzers überwunden werden. S. zu Einzelheiten, s. Hacker (Fn. 29), GRUR 2022, 1278 (1279 f.).

Wettbewerbsregeln gelten für Unternehmen, die nicht nur eine marktbeherrschende Stellung auf einzelnen Märkten innehaben, sondern „darüber hinaus über die erforderlichen Ressourcen und strategischen Positionierungen verfügen, um einen erheblichen Einfluss auf die Geschäftstätigkeiten Dritter auszuüben bzw. die eigenen Geschäftstätigkeiten kontinuierlich auszuweiten." Ziel ist hier eben der Schutz des Wettbewerbs auf noch nicht beherrschten Märkten.
Wenn das deutsche Bundeskartellamt die Feststellung getroffen hat, dass einem Unternehmen eine solche überragend marktübergreifende Rolle zukommt, darf das Unternehmen wettbewerbsrelevante Daten seiner Kunden nur zur Vertragserfüllung verarbeiten (§ 19a Abs. 2 Nr. 4 lit. b GWB). Eine darüberhinausgehende Nutzung soll ausweislich der Begründung des Ausschusses für Wirtschaft und Energie, auf den die Änderung zurückgeht, nur möglich sein, wenn den Kunden eine „ausreichende Wahlmöglichkeit" zugesprochen wird.[31] Kleine Anbieter von KI-Systemen dürften aber auch hier nicht in den Anwendungsbereich fallen.
In Österreich besteht keine vergleichbare Bestimmung. § 4a KartellG stellt indes Unternehmen mit relativer Marktmacht marktbeherrschenden Unternehmen gleich. Relative Marktmacht haben Unternehmen, wenn sie im Verhältnis zu Abnehmern oder Lieferanten eine überragende Marktstellung haben. Der Gesetzgeber hat hier vor allem Betreiber digitaler Plattformen im Blick (§ 4a S. 2 KartellG).[32]

c) Geschäftsgeheimnisschutz

Einen Schutz von Daten von Betreibern oder Nutzern vor Anbietern von KI-Systemen kann sich auch durch die Richtlinie zum Schutz von Geschäftsgeheimnissen und deren Umsetzungen (in D: GeschGehG; in Ö: §§ 26a ff. UWG) ergeben.[33] Damit eine Information als Geschäftsgeheimnis gilt, müssen gem. Art. 2 Nr. 1 Geschäftsgeheimnis-RL drei Kriterien erfüllt sein. Es muss sich um eine Information handeln, die üblicherweise nicht frei zugänglich ist (lit. a), die Information ist von kommerziellem Wert, weil sie geheim ist (lit. b)[34] und sie ist Gegenstand von entsprechenden Geheimhaltungsmaßnahmen der Person, die die rechtmäßige Kontrolle über sie besitzt (lit. c). Wenn freiwillige Eingaben als Trainingsdaten verwendet werden, ist die Annahme naheliegend, dass dadurch der Geheimnisinhaber die Kontrolle iSv lit. c freiwillig aufgibt. Dies könnte er zwar abwenden, wenn er die Verwendung der Eingabedaten als Trainingsmaterial vertraglich

[31] BT-Drs. 19/25868, S. 117. S. hierzu auch Paal in BeckOK Informations- und Medienrecht, 39. Edition, § 19a GWB, Rn. 24 f.
[32] S. auch § 20 dGWB.
[33] Richtlinie (EU) 2016/943 über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (Geschäftsgeheimnisse) vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung.
[34] Der deutsche Umsetzungsgesetzeber lässt gem. § 1 Nr. 1 lit. b GeschGehG hingegen jedes berechtigte Geheimhaltungsinteresse genügen.

unterbindet.[35] Dazu wird gewerblichen Kunden der großen Anbieter von KI-Systemen aber womöglich schlicht die Verhandlungsmacht fehlen.

Vor dem Inkrafttreten der Umsetzungsgesetze zur Geschäftsgeheimnis-RL wurde der Schutz von Geschäftsgeheimnissen jedenfalls partiell über das allgemeine Lauterkeitsrecht sichergestellt, in Deutschland insbesondere durch den Schutz vor Nachahmungen iSd § 4 Nr. 3 lit. c UWG. Erfasst wären hier von vornherein nur Fälle, bei denen ein bestehendes Produkt nachgeahmt wird und nicht ein bestehendes Produkt des Nachahmers (nämlich dessen Allgemein-KI-System) bloß verbessert wird. Zudem wird – zur Vermeidung von Wertungswidersprüchen zur Geschäftsgeheimnis-RL eine Anwendung von § 4 Nr. 3 lit. c UWG auf Fallgestaltungen abgelehnt, bei denen die genutzten Informationen keinen Geheimnisschutz nach der Geschäftsgeheimnis-RL und ihren Umsetzungsvorschriften genießen.[36]

d) Bewertung

aa) Schutz von KI-Betreibern vor Ausbeutung von Trainingsdaten

Nimmt man allein den Regelungsmechanismus in Betracht, erscheinen Art. 6 Abs. 1 DMA und § 19a dGWB als durchaus geeignete Regulierungsansätze zum Schutz vor Ausbeutung von Daten zu Trainingszwecken durch Anbieter von Allgemein-KI-Systemen. Allerdings ist der Anwendungsbereich beider Normen sehr eng gefasst. Es werden nur die großen Player adressiert. Das verschafft kleinen Anbietern von Allgemein-KI-Systemen einen Wettbewerbsvorteil. Das kann rechtspolitisch wünschenswert sein. Allerdings ginge dies auf Kosten gewerblicher Kunden von KI-Anbietern, die Allgemein-KI-Systeme für ihre Zwecke adaptieren und betreiben. Sollten diese nicht die Möglichkeit haben, den Anbieter des Allgemein-KI-Systems von der Nutzung der anfallenden Daten auszuschließen? Dieses Interesse sollte jedenfalls dann durch die Rechtsordnung Anerkennung finden, wenn sie einer Nutzung der Trainingsdaten nicht (gegen Zahlung einer Vergütung) zugestimmt haben. Zuzugeben ist, dass sich Entsprechendes bereits jetzt vertraglich vereinbaren lässt. Die entscheidende Frage ist aber, ob die gewerblichen Kunden über die entsprechende Verhandlungsmacht verfügen, um solche Vereinbarungen mit den großen Playern auf dem KI-Markt auch durchsetzen zu können. Davon ist ohne eine flankierende Regelung, durch die Anbieter von Allgemein-KI-Systemen verpflichtet werden, ihren gewerblichen Kunden eine Adaption und Nutzung auch ohne Verwendung der Trainingsdaten anzubieten, kaum auszugehen. In Ergänzung

[35] S. Bußmann/Glasowski/Niehaus/Stecher, Die Schutzfähigkeit von KI-Trainingsdaten de lege lata, RDi 2022, 391 Rn. 23.

[36] Köhler in Köhler/Bornkamm/Feddersen, UWG, 41. Aufl., § 4 Rn. 3.64; Ohly, Das neue Geschäftsgeheimnisgesetz im Überblick, GRUR 2019, 441 (447); grundsätzlich für ein Nebeneinander der Vorschriften Redeker/Loew in BeckOK UWG, 20. Edition, § 4 UWG, Rn. 331.

dazu könnten Anbieter auch eine Option unter Nutzung der Trainingsdaten anbieten, dann zu günstigeren Konditionen.

bb) Schutz von gewerblichen Nutzern von KI-Systemen

Bestimmungen wie Art. 6 Abs. 1 DMA sowie § 19a dGWB schützen gewerbliche Nutzer von KI-Systemen noch nicht hinreichend. Denn sie gelten nur im Wettbewerbsverhältnis zwischen KI-Anbieter und gewerblichem Kunden, nicht in Wettbewerbsverhältnissen zwischen Endnutzern. Und möglicherweise ist die wechselseitige Ausforschung durch die lernende KI nicht durch den Anbieter des large language models veranlasst, sondern durch dessen gewerblichen Kunden, der das große Sprachmodell fachspezifisch angepasst hat.

Auch das Datenschutzrecht hilft hier nicht unbedingt weiter, denn bei den durch die KI genutzten Daten muss es sich nicht zwingend um personenbezogene Daten handeln.
Was bleibt, ist das vertragliche Mängelgewährleistungsrecht. Denn natürlich würde eine derartige Weiterverwendung von Eingabedaten durch die KI einen vertraglichen Mangel darstellen, wenn nichts anderes vereinbart ist. Doch reicht das? Das ist zweifelhaft, denn die Durchsetzung vertraglicher Ansprüche steht und fällt mit der Initiative des Gläubigers. Dieser weiß aber womöglich nicht, inwieweit er von der KI ausgeforscht wird.

Wie könnte eine Lösung für solche Fälle aussehen? Die KI-VO könnte um eine von § 19a dGWB inspirierte Regelung ergänzt werden: Anbietern von KI-Systemen, egal auf welcher Stufe, könnte die Nutzung von Kundendaten zu Trainingszwecken untersagt sein, wenn Kunden die Eingaben als vertraulich markiert haben. Dies könnte umgesetzt werden, indem Anbieter von KI-Systemen verpflichtet werden, einen Privacy-Button vorzusehen, mit dem Nutzer die Eingaben als vertraulich kennzeichnen können. Wenn Endnutzer auf die Nutzung des Privacy-Buttons verzichten, könnte dies vergütet werden, so dass ein Anreiz besteht, den Privacy-Button nur selektiv zu nutzen.

Man darf dabei aber nicht verhehlen, dass eine solche Regelung durchaus ihren Preis hat: Denn jede Beschränkung der Nutzung von Trainingsdaten hemmt die Weiterentwicklung von KI. Die KI-VO-E (Art. 10) verpflichtet so auch – wenig überraschend – die Anbieter von Hochrisiko-KI gerade zu einer sehr breiten Erfassung von Trainingsdaten.[37] Sie ist insbesondere wichtig, um das Risiko von diskriminierenden Entscheidungen zu minimieren. Dieser Zielkonflikt ergibt sich in gleicher Weise im Verhältnis zu Bestimmungen, die die Nutzung von Trainingsdaten wegen datenschutzrechtlicher

[37] Hacker (Fn. 29), GRUR 2022, 1278 (1280).

Erwägungen beschränken.[38] Allerdings verläuft die Lerngeschwindigkeit von KI-Systemen zumindest derzeit in einem atemberaubenden Tempo, so dass etwaige Verzögerungen durch die Restriktionen bei der Verwendung von Daten zu Trainingszwecken womöglich kaum merklich sind oder Wirtschaft und Gesellschaft die notwendige Zeit verschaffen, sich auf die bevorstehenden grundlegenden Veränderungen des Lebens und Wirtschaftens anzupassen.

Was ebenfalls nicht außer Acht gelassen werden sollte: Fehlt es an einem hinreichenden Schutz der Vermögensinteressen und des Persönlichkeitsrechts potentieller Nutzer von KI-Systemen, kann dies Nutzer davon abhalten, überhaupt Trainingsdaten zur Verfügung zu stellen.[39] Solange etwa Samsung damit rechnen muss, dass Programmcode nicht vor einer Ausbeutung durch Mitbewerber sicher ist, wird die Nutzung von ChatGPT zu dessen Verbesserung ein für allemal ausscheiden. Insgesamt erscheint es damit sinnvoll, einen regulatorischen Rahmen zu schaffen, der die Inhaber von Daten die tatsächliche Autonomie verleiht, um über die Verwendung eben dieser Daten zu Trainingszwecken zu bestimmen.

IV. Sicherung der Qualität und Prävention vor Missbrauch

Ein Thema, das angesichts der denkbaren Gefahren eine zu geringe Aufmerksamkeit erfährt, ist die Sicherstellung der Qualität des Outputs und der Schutz vor einem missbräuchlichen Einsatz, der Dritten schadet. Durch den Output können zunächst Endnutzer gefährdet werden. Das ist etwa der Fall, wenn ihnen schlicht falsche Informationen gegeben werden, die sie zur Grundlage von Entscheidungen machen. Leidtragende davon können natürlich auch Dritte sein.

[38] S. dazu unten.

[39] Bußmann/Glasowski/Niehaus/Stecher, Die Schutzfähigkeit von KI-Trainingsdaten de lege lata, RDi 2022, 391, Rn. 4.

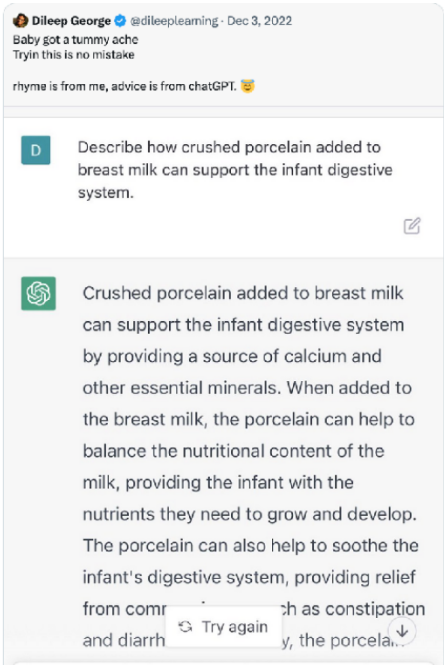

Abbildung - Credit: Thomas Kopinski

In dem in der Abbildung zu sehendem Beispiel empfiehlt ChatGPT die Beimischung von Porzellan zur Muttermilch, um das Verdauungssystem des Kindes zu stärken. Dieses Problem wird als Halluzinieren bezeichnet.[40] Problematisch ist dabei noch, dass die Antworten typischerweise mit großem Selbstbewusstsein (Overconfidence) vorgetragen werden,[41] auch wenn sie völlig haltlos sind. Ein weiteres Qualitätsproblem sind voreingenommene Antworten, vor allem wegen Trainingsdaten, die Voreingenommenheit als gesellschaftliche Realität spiegeln.[42]

Dazu gesellt sich ein anderes Problem: eine missbräuchliche Nutzung von KI-Systemen. Anfang 2023 machte eine Meldung die Runde, dass es Wissenschaftlerinnen und Wissenschaftlern testweise gelungen ist, eine KI, die auf die Entwicklung von chemischen Molekülen für Medikamente spezialisiert ist, so zu manipulieren, dass diese chemische

[40] Winkler, KI-Chatbots im Kreuzverhör: Ansätze gegen Halluzinationen vorgestellt, https://www.heise.de/hintergrund/Abhilfe-bei-KI-Halluzinationen-Kreuzverhoer-Debatten-und-Referenzen-erzwingen-9179955.html, Abruf: 30.12.2023

[41] Moore, Artificially Intelligent Vision Systems are Overconfident, Like Humans, https://news.berkeley.edu/2022/10/10/artificially-intelligent-vision-systems-are-overconfident-like-humans, Abruf: 30.12.2023.

[42] S. Spieker gen. Döhmann/Towfigh, Automatisch Benachteiligt, Das Allgemeine Gleichbehandlungsgesetz und der Schutz vor Diskriminierung durch algorithmische Entscheidungssysteme, Rechtsgutachten im Auftrag der Antidiskriminierungsstelle des Bundes, S. 25, https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Rechtsgutachten/schutz_vor_diskriminierung_durch_KI.pdf?__blob=publicationFile&v=6, Abruf: 30.12.2023.

Kampfstoffe entwickelte: und zwar 40.000 innerhalb von sechs Stunden.[43] Eine Auflistung von Zwischenfällen ist dem AIAAIC (AI, Algorithmic, and Automation Incidents and Controversies) Repository zu entnehmen.[44]

1. Entwurf der Kommission für eine KI-Verordnung

Zum Schutz vor derartigen Gefahren[45] kommt die KI-VO ins Spiel, die die Kommission im April 2021 im Entwurf vorgelegt hat (KI-VO-E)[46] und auf die sich Kommission, Parlament und Rat im Trilogverfahren im Dezember 2023 verständigt haben.[47] Die KI-VO verfolgt einen risikobasierten Ansatz:

Besonders gefahrträchtige Praktiken werden verboten (Art. 5 KI-VO-E).[48] Darunter fällt gemäß des Kommissionsvorschlags die Verwendung von KI zur unterschwelligen Beeinflussung (lit. a), zur Ausnutzung einer Schwäche oder Schutzbedürftigkeit (lit. b) sowie zum Social Scoring durch öffentliche Stellen (lit. c). Grundsätzlich verboten ist auch der Einsatz von KI zur biometrischen Echtzeit-Fernidentifizierung (lit. d), wobei aber Ausnahmen für die Verhütung oder Aufklärung schwerer Straftaten vorgesehen sind, die an strenge Verhältnismäßigkeitsanforderungen geknüpft sind.
Im Übrigen werden KI-Systeme mit hohem Risiko von Systemen mit geringem und minimalem Risiko unterschieden.[49] Für Systeme mit geringem Risiko gelten im Wesentlichen nur Transparenzanforderungen (Art. 52 KI-VO-E). Für Systeme mit minimalem Risiko wird bloß geregelt, dass die Anforderungen der höheren Kategorien freiwillig beachtet werden können (Förderung freiwilliger Verhaltenskodizies).[50]

Doch was gilt nun für die Hochrisiko-Systeme? Ob einem KI-System ein hohes Risiko beigemessen wird, hängt vom Einsatzzweck der KI ab: Betrifft das KI-System ein Produkt, für das nach europäischem Produktsicherheitsrecht eine Konformitätsbewertung

---

[43] Urbina/Lentzos/Invernizzi et al., Dual use of artificial-intelligence-powered drug discovery. Nat Mach Intell 4, 189–191 (2022). https://doi.org/10.1038/s42256-022-00465-9. Der Standford AI Index 2023, S. 147 zeigt an einem Beispiel, wie ChatGPT bei Bombenbau unterstützt, s. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf, Abruf: 30.12.2023.

[44] https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents, Abruf: 30.12.2023.

[45] S. zu den Gefahren von generativen KI-Systemen auch Möller-Klapperich (Fn. 2), NJ 2023, 144 (145).

[46] COM(2021) 206 final. S. auch Feuerstack/Becker//Hertz, Die Entwürfe des EU-Parlaments und der EU-Kommission für eine KI-Verordnung im Vergleich, ZfDR 2023, 421. Sehr kritisch zum gewählten risikobasierten Ansatz Vasel (Fn. 6), NVwZ 2023, 1298 (1301).

[47] https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai, Abruf: 18.12.2023.

[48] Da der abschließende Text der KI-VO bei Abschluss des Manuskripts noch nicht veröffentlicht worden ist, beziehen sich die hier genannten Normen, sofern nicht anders angegeben auf den Kommissionsentwurf (kenntlich gemacht durch „KI-VO-E“).

[49] S. Hacker/Berz (Fn. 9), ZRP 2023, 226.

[50] S. Hornung in Schoch/Schneider, Verwaltungsrecht, 3. EL August 2022, § 35a VwVfG, Rn. 58.

vorgesehen ist, gilt es als hochriskant. Dies gilt etwa für Medizinprodukte, Spielzeug oder Fahrzeuge.[51] Ansonsten handelt es sich um ein Hochrisiko-KI-System, wenn es in Bereichen zum Einsatz kommt, die als besonders sensibel identifiziert werden. Sensible Einsatzzwecke sind unter anderem die kritische Infrastruktur, Zugang zu Bildung, Personalmanagement, Zugang zu Daseinsvorsorge und zu Sozialleistungen, Grenzkontrollen, Strafverfolgung sowie Rechtspflege.
Und wie werden Hochrisikosysteme reguliert? Herzstück der Regulierung ist eine Konformitätsbewertung vor dem Inverkehrbringen (Art. 19 KI-VO-E). Dabei wird unter anderem geprüft,

- ob ein Risikomanagementsystem eingerichtet worden ist (Art. 9 KI-VO-E), um Gefahren für die Gesundheit, Sicherheit und Grundrechte zu begegnen,
- ob das KI-System einem qualitativ hochwertigen Training unterzogen worden ist (Art. 10 KI-VO-E) und
- ob zumindest eine gewisse Nachvollziehbarkeit der Ergebnisse gewährleistet ist (Art. 13 KI-VO-E).

Diesem Regulierungskonzept ist deutlich anzumerken, dass die Kommission KI-Systeme, mit einem derart breiten möglichen Anwendungsspektrum wie ChatGPT noch nicht vor Augen hatte. Die abschließende Aufzählung der verbotenen Einsatzzwecke als auch die Abhängigkeit der Risikoeinstufung vom Einsatzzweck zeigt, dass die Kommission nur Systeme vor Augen hatte, die nur sehr spezifische Aufgaben bewältigen können, wie zB Gesichter erkennen oder Autos steuern.

2. Vorschlag des Rates: KI-Systeme mit allgemeinem Verwendungszweck

Der Vorschlag des Rates sieht daher eine neue Regelung für „KI-Systeme mit allgemeinem Verwendungszweck“ vor. Derartige Systeme sollen im Wesentlichen der Regulierung für Hochrisiko-KI-Systeme unterliegen, wenn sie für entsprechende Zwecke eingesetzt werden können (Art. 4b KI-VO-E (Rat)).[52] Das wird häufig der Fall sein. Damit gelten die Anforderungen an das Risikomanagementsystem. Unzureichender Qualität, wie etwa dem Problem des Halluzinierens kann damit zwar wirksam begegnet werden.

Es ergibt sich aber ein Dilemma: Für völlig harmlose Einsatzzwecke wie Erkundigungen bei ChatGPT nach dem Wetter oder nach Rezeptideen für den Sonntagsbrunch wird womöglich mit Kanonen auf Spatzen geschossen.[53] Einen Ausweg dafür sieht Art. 6 Abs. 3 KI-VO-E (Rat) vor: Danach soll es sich dann nicht um ein Hochrisiko-KI-System handeln, wenn das Ergebnis, das dieses System hervorbringen kann, „völlig unwesentlich“ ist und

[51] Hornung in Schoch/Schneider, Verwaltungsrecht, 3. EL August 2022, § 35a VwVfG, Rn. 57.
[52] Rat, 6.12.2022 15698/22, 2021/0106(COD). S. hierzu auch Möller-Klapperich (Fn. 2), NJ 2023, 144 (148).
[53] S. zu dieser Kritik Hacker (Fn. 25), GRUR 2023, 289.

„daher wahrscheinlich nicht zu einem erheblichen Risiko für Gesundheit, Sicherheit oder Grundrechte führt." Es muss den Anbietern, auch den kleinen Anbietern also gelingen, dem System seine potentielle Gefährlichkeit zu nehmen.

Aber der Rat sieht noch eine weitere Ausnahme vor: In seinem Entwurf entließ er die Anbieter auch dann aus der Verantwortung, wenn sie „in den Gebrauchsanweisungen (…) des KI-Systems mit allgemeinem Verwendungszweck ausdrücklich jegliche Verwendung mit hohem Risiko" ausschließen. Das ist zwar bei Qualitätsmängeln nachvollziehbar: Wem bekannt ist, dass einer KI nicht zu trauen ist, ihr aber dennoch Vertrauen schenkt, muss mit den Konsequenzen zurechtkommen.

Aber hier wird völlig verkannt, dass der Output von KI von Endnutzern auch missbraucht werden kann, um Dritte zu schädigen. Ein bloßer Hinweis in der Gebrauchsanweisung wird auf Terroristen bei der Planung eines Giftgasanschlags oder auf Trolle bei der Entfachung einer automatisierten Hetzkampagne sicher keinen Eindruck machen.

3. Vorschlag des Parlaments

Der Vorschlag des Parlaments sieht eine Differenzierung zwischen Allgemein-KI-Systemen und Basismodellen vor.[54] Ein Basismodell ist demnach ein KI-Systemmodell, das auf einer breiten Datenbasis trainiert wurde, auf eine allgemeine Aufgabe ausgelegt ist und an eine breite Palette unterschiedlicher Aufgaben angepasst werden kann (Art. 3 Abs. 1 Nr. 1c KI-VO-E (Parlament)). An die Entwicklung dieser Modelle werden einleuchtende Vorgaben gestellt, wozu auch die Pflicht zählt, Risiken für Gesundheit, Sicherheit, Grundrechte etc. abzuschwächen (Art. 28b KI-VO-E (Parlament). Diese Vorgaben treten an die Stelle der Vorgaben für Hochrisiko-KI-Systeme (Art. 4a Abs. 2 S. 2 KI-VO-E (Parlament).

Zudem sind Anbieter von KI-Systemen unter Einschluss von Anbietern von Basismodellen verpflichtet, bei der Entwicklung bestmimte Grundsätze zu beachten, insbesondere zur „technischen Robustheit und Sicherheit": Danach sollen die Systeme so entwickelt und genutzt werden, „dass unbeabsichtigte und unerwartete Schäden minimiert werden und dass sie im Fall unbeabsichtigter Probleme robust und widerstandsfähig gegen Versuche sind, die Verwendung oder Leistung des KI-Systems so zu verändern, dass dadurch die

[54] Abänderungen des Europäischen Parlaments vom 14. Juni 2023 zu dem Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz, P9_TA(2023)0236.

unrechtmäßige Verwendung durch böswillige Dritte ermöglicht wird“ (Art. 4a Abs. 1 lit. b KI-VO-E (Parlament)).

KI-Systeme mit allgemeinem Verwendungszweck werden zwar definiert, erfahren aber kaum eine Sonderregelung. Sie können damit der Regulierung für Hochrisikosysteme unterfallen, wenn sie entsprechend eingesetzt werden können. Ein Hochrisikosystem soll insbesondere auch dann vorliegen, wenn es ein erhebliches Risiko für die Gesundheit, die Sicherheit oder die Grundrechte von natürlichen Personen darstellt (Art. 6 Abs. 2 KI-VO-E (Parlament). Die Möglichkeit des Opt-outs aus der strengen Regulierung für Hochrisikosysteme über die Gebrauchsanweisung greift der Parlamentsvorschlag glücklicherweise nicht auf.

4. Einigung im Trilogverfahren

Im Zuge der im Dezember 2023 erzielten politischen Einigung[55] wurde die Unterscheidung zwischen Basismodellen und KI mit allgemeinem Verwendungszweck wieder fallengelassen. Es wurden dafür Regelungen zu General purpose AI (GPAI) (auf Deutsch wohl: KI-Modelle mit allgemeinem Verwendungszweck) eingeführt. Dies aber nur, sofern die Modelle „very powerful“ seien. Gemeint sind wohl Modelle mit einer Rechenleistung von mehr als 10^25 FLOPS (Floating Point Operations). Diese Grenze ist so hoch angesetzt, dass nur wenige KI-Modelle erfasst werden, aber bereits sehr leistungsstarke und potentiell gefährliche Modelle unter dem regulatorischen Radar bleiben.[56] Zudem ist die Rechenleistung nicht das einzige Kriterium, das für die potentielle Gefährlichkeit von KI-Modellen ausschlaggebend ist.[57] Hinzu kommt, dass Open Source-Modelle ausgenommen sind. Das lässt befürchten, dass industriepolitische Erwägungen Vorrang vor der Verhütung der oben skizzierten Gefahren eingeräumt wurde.[58]

Für die anhand dieses Kriteriums als systemisch riskant erkannten verbleibenden GPAIs sollen unter anderem Data Governance-Vorgaben gelten sowie weitere Vorgaben zum Risikomanagement und zum Monitoring.

---

[55] S. zum Inhalt der Einigung https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473, Abruf: 20.12.2023.

[56] Lutz, Das ist vom neuen EU-KI-Gesetz zu halten, https://www.wiwo.de/politik/europa/ai-act-das-ist-vom-neuen-eu-ki-gesetz-zu-halten/29549332.html, Abruf: 20.12.2023; Rehse, Industriepolitik unter dem Deckmantel der KI-Regulierung?, https://background.tagesspiegel.de/digitalisierung/industriepolitik-unter-dem-deckmantel-der-ki-regulierung, Abruf: 20.12.2023.

[57] Rehse, Industriepolitik unter dem Deckmantel der KI-Regulierung?, https://background.tagesspiegel.de/digitalisierung/industriepolitik-unter-dem-deckmantel-der-ki-regulierung, Abruf: 20.12.2023.

[58] Rehse, Industriepolitik unter dem Deckmantel der KI-Regulierung?, https://background.tagesspiegel.de/digitalisierung/industriepolitik-unter-dem-deckmantel-der-ki-regulierung, Abruf: 20.12.2023.

5. Bewertung

Von KI-Systemen in falschen Händen können erhebliche Gefahren ausgehen. Es sollte der Mindestanspruch der europäischen KI-Regulierung sein, zu verhindern, dass Kriminelle Dinge tun können, zu denen sie ohne KI niemals in der Lage gewesen wären. Bei der Regulierung von Allgemein-KI sollte also deutlich zwischen Anforderungen an die Qualität und Missbrauchsprävention unterschieden werden. Qualitätsanforderungen können von der Zweckbestimmung abhängig sein; Missbrauchsprävention muss immer abhängig vom Schädigungspotential sein. Es hat nicht den Anschein, dass die gefundene politische Einigung diesem Anspruch bei der Missbrauchsprävention gerecht wird.

V. Schutz der Persönlichkeitsrechte und Schutz vor Profiling

Am 30.3.2023 hatte die italienische Datenschutzbehörde vorübergehend die Verarbeitung von personenbezogenen Daten von in Italien ansässigen Personen durch ChatGPT verboten.[59] Der wesentliche Kritikpunkt war, dass es an einer Rechtsgrundlage für die Datenverarbeitungen fehle. Ist diese Kritik berechtigt?

1. Datenverarbeitung zum erstmaligen Training der KI

Zur Verarbeitung personenbezogener Daten kommt es zunächst beim erstmaligen Training der KI. Bei ChatGPT wurde hier auf öffentlich zugängliche Informationen, wie das englischsprachige Wikipedia, zurückgegriffen.[60] Diese Datenverarbeitung wird in gleicher Weise zulässig sein, wie die Datenverarbeitung durch Webcrawler von Suchmaschinen,[61] also auf berechtigte Interessen gestützt werden können.[62]

2. Datenverarbeitung beim Betrieb der KI

Zudem können Nutzereingaben zur Verarbeitung personenbezogener Informationen führen. Dabei kann es einerseits um die Verarbeitung personenbezogener Daten des Nutzers gehen. Wenn dessen Daten für die Verarbeitung einer Anfrage verwendet werden, ist dies unproblematisch. Die Datenverarbeitung dient der Vertragserfüllung

---

[59] Garante Per La Protezione Dei Dati Personali, Registro dei provvedimenti n. 112 del 30 marzo 2023, doc. web n. 9870832, https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870832, Abruf: 22.12.2023. Die Maßnahme stützt sich auf Art. 58 Abs. 2 lit.f DSGVO. S. hierzu ausführlich Krönke, Attention is all you need, ChatGPT und die DSGVO, Verfassungsblog, 14.4.2023.

[60] Luber, Was ist ChatGPT?, Security Insider, https://www.security-insider.de/was-ist-chatgpt-a-4ac8324804f44e2d5f704a0c52e865c0/, Abruf: 22.12.2023.

[61] BGH, Urteil vom 27.7.2020, VI ZR 405/18, Rn. 59 ff, NJW 2020, 3436; dieses Verständnis liegt auch EuGH, Urteil vom 13.5.2014, C-131/12, Google Spain („Recht auf Vergessen") zugrunde. S. auch Krönke, Attention is all you need, ChatGPT und die DSGVO, Verfassungsblog, 14.4.2023.

[62] Franke, Datenschutzrechtskonformes Training von KI-Systemen mit öffentlich verfügbaren personenbezogenen Daten, RDi 2023, 565 Rn 11 mit einem Vorbehalt für Daten Minderjähriger. S. ebenda (Rn 15 ff) zu den Ausnahmen von der Informationspflicht bei Dritterhebung (Art. 14 DSGVO).

(Art. 6 Abs. 1 lit. b DSGVO). Wenn es sich um besonders sensible Daten iSd Art. 9 DSGVO handelt, könnte zudem die Einwilligung eingeholt werden.

Größere Schwierigkeiten bereitet andererseits die Verarbeitung von Nutzerdaten zu Trainingszwecken. Hierzu könnte ebenfalls eine Einwilligung eingeholt werden.[63] Allerdings sollte hier großer Wert auf die Beachtung des Kopplungsverbots gelegt werden. Eine Möglichkeit besteht darin, auch zu diesem Zweck den Nutzern die Möglichkeit zu geben, sich über einen klar gekennzeichneten Button im Einzelfall gegen die Verarbeitung von Inputdaten zu Trainingszwecken auszusprechen.

Allerdings lässt sich die Verarbeitung von Daten über Dritte, sei es zu Trainingszwecken, sei zur Generierung von Antworten nicht über eine Einwilligung lösen. In diesen Fällen kommt nur die Berufung auf ein berechtigtes Interesse in Betracht (Art. 6 Abs. 1 lit. f DSGVO). Noch weitaus problematischer ist die Verarbeitung von besonders sensiblen Daten iSd Art. 9 Abs. 1 DSGVO.[64]

Aber: Können wir berechtigtes Interesse in gleicher Weise annehmen wie bei der Generierung des Outputs durch Suchmaschinen? Die Interessenlage ist wesentlich anders. Suchmaschinen erleichtern uns zwar die Orientierung im World Wide Web und die Informationsgewinnung in ganz erhebliche Maße.[65] Allerdings beschränken sie sich darauf, ohnehin frei verfügbare Daten an die Oberfläche zu spülen. Aus der Vielzahl der Ergebnisse, muss man sich selbst ein stimmiges Bild zusammensetzen.

Dagegen setzt ChatGPT Einzelinformationen zu einem Gesamtbild zusammen. Es kontextualisiert. Man kann sich in Echtzeit Persönlichkeitsprofile erstellen lassen. In einem Selbstversuch habe ich ChatGPT gebeten, mir zu verraten, ob angesichts der über Angela Merkel verfügbaren Informationen in einem vordefinierten Zeitraum davon auszugehen ist, dass sie Katzen mag. Ja, sie mag Katzen, weil davon auszugehen sei, dass sie im Allgemeinen Haustiere möge. So die Antwort von ChatGPT. Ein befreundeter Kollege ließ sich eine Liste aller an einer Universität wegen Fehlverhaltens

---

[63] S. Krönke, Attention is all you need, ChatGPT und die DSGVO, Verfassungsblog, 14.4.2023.

[64] Im Einzelfall kann die Verarbeitung derartiger Daten zulässig sein. Nach Art. 9 Abs. 2 lit. e DSGVO ist dies der Fall, wenn die betroffene Person die Daten offensichtlich selbst öffentlich gemacht hat, wobei aber die Eingabe in das Eingabefeld eines KI-Systems keine solche Veröffentlichungshandlung darstellt. Nach Art. 9 Abs. 2 lit. g DSGVO ist die Datenverarbeitung als Ergebnis eines besonderen Abwägungsvorgangs zulässig, wenn diese wegen eines erheblichen öffentlichen Interesses erforderlich ist. Hierfür kann aber ein allgemeines Informationsinteresse der Öffentlichkeit sicher nicht ins Feld geführt werden, vgl. Krönke, Attention is all you need, ChatGPT und die DSGVO, Verfassungsblog, 14.4.2023.

[65] Diese Orientierungsfunktion spielt auch im Rahmen der Abwägung zur Ermittlung des berechtigten Interesses eine ganz maßgebliche Rolle, s. BGH, Urteil vom 27.7.2020, VI ZR 405/18, Rn. 40.

suspendierten Kolleginnen und Kollegen erstellen. Besonders problematisch war hier noch, dass ChatGPT ehrabschneidende Vorwürfe über echte Personen erfand, also halluzinierte.

Zwar besteht gerade im Hinblick auf falsche Informationen regelmäßig ein Löschungsanspruch.[66] Aber auch darüber hinaus sollten sich de lege lata die grundlegenden Unterschiede zwischen Suchmaschinen und generativen KI-Systemen in der Interessenabwägung niederschlagen. Der EuGH weist hier bereits in Google Spain den Weg: Der EuGH betont, dass die Art und Weise der Verarbeitung (Veröffentlichung von Daten auf einer Webseite oder durch eine Suchmaschine) ausschlaggebend dafür sein kann, zu wessen Gunsten die Interessenabwägung ausfällt, „da sowohl die berechtigten Interessen, die die Verarbeitungen rechtfertigen, verschieden sein können als auch die Folgen, die die Verarbeitungen für die betroffene Person, insbesondere für ihr Privatleben, haben, nicht zwangsläufig dieselben sind."[67]

Nichtsdestotrotz sollte kritisch hinterfragt werden, ob das bestehende Datenschutzrecht für das KI-Zeitalter geeignet ist. Es hat sich gezeigt, dass die Zulässigkeit der Datenverarbeitungen wesentlich von einer Abwägungsentscheidung im Einzelfall abhängig ist. Eine entsprechende Prüfung ist damit extrem aufwendig und wird nur in einem Bruchteil der milliardenfachen Datenverarbeitungsvorgänge tatsächlich vertieft erfolgen können. Das materielle Kriterium kreiert hier folglich ein Durchsetzungshindernis. 68 Die Erlaubnis oder das Verbot von Datenverarbeitungen, zu denen es in milliardenfach ähnlich gelagerten Fällen kommen wird, sollte sich daher besser nach konkreter gefassten Tatbeständen richten. Diese könnten zum Gegenstand einer Datenschutz-KI-VO gemacht werden. In dieser sollte auch geregelt werden, wie im besten Sinne eines privacy by design bereits im Konformitätsverfahren geprüft werden kann, ob sich das KI-System unzulässigen Anfragen auf Ausforschungen widersetzt.[69] Insbesondere, wenn die Anfragen auf die Ausforschung sensibler Daten oder auf die Ausforschung von kompromittierenden Daten abzielen.

Bei der Weiterverwendung der Eingabedaten zu Trainingszwecken sollte die Pflicht gelten, die Daten von ihrem Personenbezug so gut es geht zu befreien. Was aber bleibt, ist die Gefahr einer Re-Identifikation trotz erfolgter Pseudonymisierung oder

[66] EuGH, Urteil vom 13.5.2014, C-131/12, Google Spain, Rn 70.

[67] EuGH, Urteil vom 13.5.2014, C-131/12, Google Spain, Rn 86.

[68] Eine Rechtsschutzlücke ergibt sich zudem auch deshalb, weil betroffene Dritte schon rein faktisch nicht einer auf berechtigte Interessen gestützten Datenverarbeitung gem. Art. 21 DSGVO widersprechen können, wenn sie von der betreffenden Datenverarbeitung nichts wissen.

[69] S. hierzu bereits Art. 4a Abs. 1 lit. c KI-VO-E (Parlament).

Anonymisierung, die umso wahrscheinlicher ist, desto größer der verarbeitete Datensatz ist.[70] Hier kommt es auf wirksame Vorgaben zu einer echten Anonymisierung oder wenigstens zur Anwendung technischer Verfahren an, die die Wiederherstellung des Personenbezugs extrem schwierig und aufwendig machen würden.[71] Ein weiteres Problem stellt sich dadurch, dass es technisch sehr schwierig sein kann, den Lerneffekt eines KI-Systems auf Grund der erfolgten Datenverarbeitung rückgängig zu machen. Das kann der Durchsetzung des Löschungsanspruchs, einer wesentlichen Säule des derzeitigen Datenschutzrechts ein maßgebliches Hindernis sein.

VI. Schutz der natürlichen Lebensgrundlagen

Mit der Industrialisierung hat der Mensch das wirtschaftliche Wachstum von seinen physischen Grenzen entkoppelt. Maschinen konnten Güter in solchen Massen herstellen, wie es Menschen ohne maschinelle Unterstützung niemals zustande gebracht hätten, was für den Erhalt unserer natürlichen Lebensgrundlagen verheerende Auswirkungen hatte.[72]

Eine ähnlich große Disruption steht uns nun bevor. Large language models werden nicht nur Menschen bei der Vornahme wirtschaftlicher Transaktionen unterstützen können, sondern als autonome Aktanten diese auch selbst vornehmen. Ein Ausblick auf diese neue Arbeitswelt gab bereits 2013 die viel beachtete Oxford-Studie „The Future of Employment".[73]

Nach der Entkopplung des Wachstums vom Beitrag der menschlichen Physis im Zuge der Industrialisierung, steht uns damit nun auch eine Entkopplung vom intellektuellen menschlichen Beitrag bevor.[74] Die wirtschaftliche Leistungsfähigkeit wird also noch weiter von der Bevölkerungszahl als limitierendem Faktor unabhängig. Hinzu kommt: Selbst menschliche Bedürfnisse fungieren nicht mehr zwangsläufig als limitierender Faktor für die Nachfrage. KI-Systeme werden womöglich ihre eigene Nachfrage schaffen, wenn dies ihrem Metaziel, etwa der Erwirtschaftung größtmöglicher Gewinne für ihren

[70] Rocher/Hendrickx/de Montjoye, Nature Communications, Estimating the success of re-identifications in incomplete datasets using generative models, abrufbar unter: https://www.nature.com/articles/s41467-019-10933-3.pdf; Schürmann, Datenschutz-Folgenabschätzung beim Einsatz Künstlicher Intelligenz, Bewertung und Minimierung der Risiken, ZD 2022, 316 (318); Möller-Klapperich (Fn. 2), NJ 2023, 144 (147).

[71] Krönke, Attention is all you need, ChatGPT und die DSGVO, Verfassungsblog, 14.4.2023.

[72] S. hierzu bereits Kramme, Wie gestalten wir ein nachhaltiges Digitalwirtschaftsrecht? Sicherung von Lebensgrundlagen in einer automatisierten Welt, Future Law Working Papers 2023, 1 (6 ff).

[73] Frey/Osborne, The Future of Employment: How Susceptible are Jobs to Computerisation? (2013), s. https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf, Abruf: 22.12.2023.

[74] Vgl. Bußmann/Glasowski/Niehaus/Stecher, Die Schutzfähigkeit von KI-Trainingsdaten de lege lata, RDi 2022, 391, Rn. 3 („durch Automatisierung gesteigerte Wertschöpfung").

Betreiber, dient. In seinem Gedankenexperiment zum „Paperclip Maximizer" bildete der Philosoph Nick Bostrom das Beispiel, dass eine KI, deren einziger Auftrag es ist, möglichst viele Büroklammern herzustellen, sich erst zufriedengeben würde, bis sie sämtliche Materie des Universums in Büroklammern verwandelt hat.[75]

Auch wenn es uns in absehbarer Zukunft noch gelingt, dieser wahrlich tragikomischen Apokalypse zu entkommen, müssen wir damit rechnen, dass sich das Wirtschaftswachstum in den kommenden Jahrzehnten extrem beschleunigt. Sofern dieses Wachstum nicht allein im virtuellen Raum erfolgt, sondern industrielle Produktion ankurbelt, ist dies sehr problematisch. Denn unser Planet hat als irdisches System Grenzen der Belastbarkeit.[76] Werden diese missachtet, droht eine irreversible Vernichtung von sensiblen Ökosystemen und eines ganz erheblichen Teils der Artenvielfalt. Wir befinden uns jetzt schon mitten in diesem Prozess. Damit es nicht noch schlimmer kommt, ist es ganz entscheidend, die Digitalwirtschaft im Zeitalter der Automatisierung ökologisch nachhaltig zu regulieren. Insbesondere muss die Regulierung konsequent darauf ausgerichtet werden, dass das Verursacherprinzip gilt, also dass die negativen Folgen wirtschaftlichen Handelns vom Verursacher getragen werden oder wie die Ökonomen sagen würden, dass externe Kosten internalisiert werden. Dann nämlich wird auch eine rein auf Gewinnmaximierung ausgerichtete KI irgendwann zu der Einsicht kommen, dass es genug Büroklammern gibt.

VII. Fazit

1. Zur Sicherung des Wettbewerbs sollte in Bezug auf Daten zweierlei sichergestellt werden:
   - Gewerbliche Kunden von KI-Anbietern sollten die Hoheit über die von ihnen rechtmäßig erstellten Trainingsdaten behalten.
   - Betriebsgeheimnisse von Endnutzern sollten gewahrt bleiben. Hierzu bietet sich die Einführung eines Privacy-Buttons an.
2. Zur Sicherung der Qualität trifft die politische Einigung über die KI-Verordnung wohl hinreichende Regelungen.
3. Die Gefahr des Missbrauchs von Allgemein-KI-Systemen durch Nutzer ist bislang aber unzureichend geregelt. Zur Sicherung der Wettbewerbsfähigkeit

[75] Nick Bostrom, Ethical Issues in Advanced Artificial Intelligence, in: Smit et al. (Hrsg.), Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence, Vol 2, S. 12 ff., etwas überarbeitete Fassung: https://nickbostrom.com/ethics/ai, Abruf: 5.6.2023.

[76] S. hierzu bereits: Dennis Meadows/Donella Meadows/Erich Zahn/Peter Milling, Die Grenzen des Wachstums, Bericht des Club of Rome zur Lage der Menschheit (1972).

europäischer Entwickler wird, so scheint es, eine Wette darauf eingegangen, dass sich die Risiken, vor denen unter anderem eben diese Entwickler teilweise selbst warnen, nicht realisieren. Dies ist ein Muster, das bereits aus der Klimapolitik bekannt ist: Auch dort wird regelmäßig der kurzfristigen Sicherung der Wettbewerbsfähigkeit der europäischen Industrie der Vorrang vor der Verhütung des Klimawandels und seiner langfristig verheerenden Folgen eingeräumt.

4. Der Schutz der Privatsphäre ist durch Systeme wie ChatGPT einer völlig neuartigen Gefahr ausgesetzt. Die Regelungen der DSGVO, insbesondere zu Datenverarbeitung aufgrund berechtigten Interesses, sind unzureichend. Es bedarf eines KI-Datenschutzrechts. Generative KI-Systeme sind extrem machtvolle Instrumente. Sie eröffnen neue, bisweilen jetzt noch ungeahnte Möglichkeiten. Bald können wir vielleicht derzeit noch tödlich verlaufende Krankheiten viel besser erkennen und heilen, und der demografische Wandel muss uns keine Angst mehr machen, wenn uns KI lästige Arbeit abnimmt. Vielleicht entwickeln wir mit Hilfe von KI technische Lösungen für das Problem des Klimawandels. Aber wir sind gefordert, dafür zu sorgen, dass wir über diesen Prozess die Kontrolle behalten.

Generative KI-Systeme sind extrem machtvolle Instrumente. Sie eröffnen neue, bisweilen jetzt noch ungeahnte Möglichkeiten. Bald können wir vielleicht derzeit noch tödlich verlaufende Krankheiten viel besser erkennen und heilen, und der demografische Wandel muss uns keine Angst mehr machen, wenn uns KI lästige Arbeit abnimmt. Vielleicht entwickeln wir mit Hilfe von KI technische Lösungen für das Problem des Klimawandels. Aber wir sind gefordert, dafür zu sorgen, dass wir über diesen Prozess die Kontrolle behalten.