

Oliver Bendel

KI-basierte Textgeneratoren aus Sicht der Ethik

Abstract: This chapter deals with AI-based text generators from an ethical point of view. It begins with a brief overview of generative AI (section 2) and AI-based text generators in particular. This is followed by a compilation of relevant areas of applied ethics (section 3), especially information ethics and its relatives, and then a structured, detailed ethical discussion (section 4). Opportunities and risks of operation and use are addressed. The chapter concludes with a summary and outlook (section 5). Overall, numerous risks are identified, as well as opportunities such as supporting work. The transformative potential for an important part of the creative and scientific professions is obvious. It is clearly a second wave of digitization that is taking hold.

1 Einführung

Die Disziplin der Künstlichen Intelligenz (KI) hat seit ihrer Gründung in den 1950er-Jahren viele Höhen und Tiefen erlebt. Um das Jahr 2000 herum herrschte unter ihren Vertretern verbreitet Pessimismus. Die Fortschritte bei bestimmten Formen künstlicher Intelligenz, etwa bei Übersetzungsprogrammen, deuteten jedoch eine Renaissance an. Ab 2020 erregten Sprachmodelle zumindest die Aufmerksamkeit von Experten. Der Durchbruch kam, als OpenAI im Jahre 2022 den Bildgenerator DALL-E 2 aufschaltete und den Chatbot und Textgenerator ChatGPT der Allgemeinheit zur Verfügung stellte. Ihnen folgten bald etliche Anwendungen generativer KI. Insbesondere KI-gestützte Text- und Bildgeneratoren werden heute privat und beruflich häufig verwendet, mehr und mehr auch Videogeneratoren. Dabei ergeben sich in Arbeitswissenschaft, Sozialwissenschaft und Ethik vielfältige Fragen.

Der Blick aus der Perspektive der Ethik fällt bei generativer KI, wenn man die Literatur studiert, vor allem auf Textgeneratoren. Als Anwendungsgebiete sind hierbei die Wissenschaft oder die Wirtschaft im Fokus.¹ Zudem interessieren Aus- und Weiterbildung hinsichtlich der Textproduktion. Die moralischen Implikationen

¹ Hazem Zohny et al., „Ethics of generative AI“, *Journal of Medical Ethics*, Bd. 49, 2023, 79–80; Steven C. H. Hoi, „Responsible AI for Trusted AI-powered Enterprise Platforms“, *Proceedings of the Sixteenth ACM International Conference on Web Search and Data*, New York 2023.

von Bild- und Videogeneratoren sind vorrangig in den Medien ein Thema, zumal man mit den Resultaten – etwa Bildern von spärlich bekleideten jungen Frauen – gleich die Artikel schmücken kann.² Trotz der vorhandenen und zunehmenden Beschäftigung mit Textgeneratoren fehlen nach wie vor überzeugende systematische Abhandlungen. Diese Lücke versucht dieser Beitrag ein Stück weit zu schließen.

Der vorliegende Beitrag schafft zunächst in aller Kürze die Grundlagen zu generativer KI (Abschnitt 2), insbesondere zu KI-basierten Textgeneratoren. Es folgt eine Zusammenstellung der zuständigen Bereiche der angewandten Ethik (Abschnitt 3), vor allem der Informationsethik und ihrer Verwandten, und daraufhin eine strukturierte, ausführliche ethische Diskussion (Abschnitt 4). Dabei werden Chancen und Risiken des Betriebs und der Nutzung behandelt.³ Eine Zusammenfassung mit Ausblick (Abschnitt 4) rundet den Beitrag ab.

2 Grundlagen von Textgeneratoren

In diesem Abschnitt werden KI-gestützte Textgeneratoren als Formen generativer KI erklärt und ihre Funktionen genannt. „Text“ bezieht sich auf das Format, das von ihnen ausgegeben wird, wobei auch Text (evtl. neben Bild) eingegeben werden kann. Es wird entsprechend geklärt, welche Form dieser annehmen kann, ohne dass dabei in die Tiefe gegangen wird.

Textgeneratoren sind eine Ausprägung generativer KI. Im Englischen spricht man von „generative AI“, ein Wort, das ins Deutsche als „Generative AI“ übernommen wurde.⁴ Es handelt sich um einen Sammelbegriff für KI-basierte Systeme, mit denen auf scheinbar professionelle und kreative Weise alle möglichen Ergebnisse produziert (oder eben generiert) werden können, etwa Texte (Textgeneratoren), Bilder (Bildgeneratoren) und Videos (Videogeneratoren), ferner Codes, Formeln, Simulationen etc.⁵ Grundlage sind Eingaben der Benutzer, sogenannte Prompts.

Bei Generative AI wird Machine Learning (maschinelles Lernen) verwendet, insbesondere Deep Learning, bei dem Big Data eine Rolle spielt, wobei unter-

2 Adrian Lobe, „Sieht unheimlich echt aus! Wie die künstliche Intelligenz Erotik-Models überflüssig macht“, *Luzerner Zeitung*, 2023, in [<https://t1p.de/9fj15>] (Zugriff: 28.09.2023).

3 Amanda Hetler, „Pros and cons of AI-generated content“, *TechTarget*, 2023, in [www.techtarget.com/whatis/feature/Pros-and-cons-of-AI-generated-content] (Zugriff: 28.09.2023).

4 „AI“ steht für „artificial intelligence“.

5 Oliver Bendel, „Generative KI“, *Gabler Wirtschaftslexikon*, Wiesbaden 2023b, in [www.wirtschaftslexikon.gabler.de/definition/generative-ki-124952] (Zugriff: 28.09.2023).

schiedliche Datenquellen – vor allem Internetquellen wie Foren und Aggregatoren – und Trainingsmethoden herangezogen werden.⁶ Bei Reinforcement Learning from Human Feedback (RLHF), einer Erweiterung des Reinforcement Learning, nutzt man die Klassifikation und Evaluation durch entsprechend instruierte menschliche Arbeitskräfte, die oft irgendwo auf der Welt schlecht bezahlt ihr Dasein fristen. Mit deren Feedback wird ein Belohnungssystem trainiert, das wiederum – um ein Beispiel zu nennen – einen Chatbot trainiert.

Textgeneratoren wie ChatGPT können auf der Basis von text- und bildbasierten Prompts Texte aller Art erstellen, zusammenfassen und übersetzen, redigieren und paraphrasieren oder in Stil und Ausdruck verändern, u. a. Bücher, Studienarbeiten, Fachartikel, Werbetexte, Gedichte oder Rezepte. Dabei können sie über sich hinauswachsen, dialogische Strukturen erlernen und zum Chatbot werden.⁷ Einen solchen kann man mit einem Text-to-Speech- und einem Speech-to-Text-System verbinden und in einen sozialen Roboter integrieren, der damit weitreichende natürlichsprachliche Fähigkeiten erlangt, oder in Suchmaschinen, wie es Microsoft und Google gemacht haben.⁸ Seit 2023 verwendet man dafür verstärkt GPT-4 oder vergleichbare Sprachmodelle.

3 Ausgewählte Bereiche der angewandten Ethik

Dieser Abschnitt klärt in aller Kürze, was Ethik und speziell angewandte Ethik ist. Im Zusammenhang mit der angewandten Ethik geht er vor allem auf die klassischen Bereichsethiken ein. Dabei wird der Begriff der Informationsethik weit gefasst. Die Maschinenethik wird nicht als klassische Bereichsethik gesehen, sondern (ausgehend vom Subjekt der Moral, das gerade in ihr freilich kein vollständiges ist, lediglich ein abgeschautes und nachgebildetes) als Ethik neben der Menschenethik (der klassischen Ethik, die sich auf den Menschen als Subjekt richtet) beziehungsweise als Sonderbereich der angewandten Ethik.⁹

Die Ethik ist eine Disziplin der Philosophie und entstand vor zweieinhalb Jahrtausenden, zunächst hauptsächlich auf der Grundlage der Werke von Aristot-

⁶ Bendel 2023b.

⁷ Ebd.

⁸ Frederic Lardinois, „Microsoft launches the new Bing, with ChatGPT built in“, *TechCrunch*, 2023, in [www.techcrunch.com/2023/02/07/microsoft-launches-the-new-bing-with-chatgpt-built-in/] (Zugriff: 28.09.2023).

⁹ Oliver Bendel, *Handbuch Maschinenethik*, Wiesbaden 2019a.

teles,¹⁰ eingebettet in eine fertile Phase antiker Hochkultur. Sie begründet und hinterfragt die Moral, zum Beispiel was als gut und böse beziehungsweise gerecht und ungerecht gilt. Dabei konzentriert sie sich auf das Handeln und nicht auf das Denken. Die angewandte Ethik bezieht sich auf mehr oder weniger klar abgrenzbare Themenfelder und bildet bestimmte Bereichsethiken heraus, die im Wesentlichen Reflexionsdisziplinen sind. Im vorliegenden Kontext sind Informationsethik, Technikethik, Medienethik und Wirtschaftsethik relevant. Zudem wird innerhalb der angewandten Ethik die Maschinenethik herangezogen.

Technikethik bezieht sich auf moralische Fragen der Anwendung von Technik und Technologie, auf die Handlungen des Homo faber, des technischen Menschen.¹¹ Sie kann sich mit Automobil- oder Waffentechnik ebenso befassen wie mit Nanotechnologie oder Nukleartechnik. In der Informationsgesellschaft, in der immer mehr Produkte Computertechnologien enthalten und Services virtuell abgewickelt werden, ist die Technikethik besonders eng mit der Informationsethik verbunden und verschmilzt teilweise mit ihr. Nach einer anderen Definition ist die Informationsethik ein Teil der Technikethik. Allerdings ist ihr Gegenstandsbereich inzwischen sehr groß, sodass viel für ihre Eigenständigkeit spricht.

Die Moral der Informationsgesellschaft ist Gegenstand der Informationsethik.¹² Diese erforscht, wie man sich – als Homo faber des Informationszeitalters – bei der Bereitstellung und Nutzung von Internetdiensten, Informations- und Kommunikationstechnologien (IKT), Informationssystemen, digitalen Medien, KI-Systemen und Robotern moralisch verhält beziehungsweise verhalten sollte.¹³ Sie ist im Zentrum der Bereichsethiken angesiedelt: Alle müssen sich mit ihr verständigen, da alle Anwendungsbereiche von computerbasierten Technologien durchdrungen sind. Die KI-Ethik kann als Teil der Informationsethik oder als eigenständiger Bereich betrachtet werden, ebenso die Roboterethik.

Der Gegenstand der Medienethik ist die Moral in den Medien und der Medien.¹⁴ Dabei interessieren sowohl die Methoden der Massenmedien und sozialen Medien als auch das Verhalten der Benutzer sozialer und digitaler Medien und ihre Rolle als Prosumenten (d. h. Produzenten und Konsumenten). Automatismen und Manipulationen durch KI-basierte Technologien und Systeme rücken in den Fokus und sind eng mit der Informationsethik verknüpft. Es erwachsen enge Verbindungen zur

¹⁰ Oliver Bendel, *400 Keywords Informationsethik: Grundwissen aus Computer-, Netz- und Neue-Medien-Ethik sowie Maschinenethik*, Wiesbaden 2019b.

¹¹ Ebd.

¹² Ebd.

¹³ Rainer Kuhlen, *Informationsethik. Umgang mit Wissen und Informationen in elektronischen Räumen*, Konstanz 2004.

¹⁴ Bendel 2019b.

Wirtschaftsethik, zumal sich die Medienlandschaft im Wandel befindet und der wirtschaftliche Druck wächst.

Die Wirtschaftsethik beschäftigt sich mit der Moral in der Wirtschaft.¹⁵ Im Mittelpunkt steht das Individuum, das produziert, handelt, führt und ausführt sowie konsumiert (Konsumentenethik), und das Unternehmen, das Verantwortung gegenüber Mitarbeitern, Kunden und der Umwelt trägt (Unternehmensethik). Damit sind auch Mikro- und Mesoebene angesprochen. Darüber hinaus sind die moralischen Implikationen wirtschaftlicher Prozesse und Systeme sowie von Globalisierung und Monopolisierung von Interesse (Wirtschaftsethik mit Bezug zur Makroebene). In der Informationsgesellschaft ist die Wirtschaftsethik mit ihren Haupt- und Nebenzweigen eng mit der Informationsethik verwachsen.

Gegenstand der Maschinenethik ist die Moral von Maschinen, meist die von teilautonomen und autonomen Systemen wie Chatbots, Sprachassistenten, bestimmten Robotern, bestimmten Drohnen und selbstfahrenden Autos.¹⁶ In ihr und mit ihr entstehen sogenannte moralische Maschinen mit einer sogenannten künstlichen Moral. Meist implementiert man moralische Regeln, die strikt eingehalten werden. Es gibt zudem Ansätze mit Machine Learning, und bei der aufkommenden Constitutional AI werden Sprachmodelle genutzt.¹⁷ Die Maschinenethik ist keine klassische Bereichsethik und mehr als eine Reflexionsdisziplin. Sie positioniert sich zwischen Philosophie auf der einen Seite und Informatik beziehungsweise Robotik auf der anderen. Wie diese technischen Disziplinen ist sie eine Gestaltungsdisziplin – sie untersucht, mit anderen Worten, nicht nur ihre Gegenstände, sondern bringt sie auch (mit Hilfe technischer Mittel) hervor.

4 Eine ethische Diskussion von Textgeneratoren

In diesem Abschnitt werden ethische Fragen im Zusammenhang mit KI-basierten Textgeneratoren aufgeworfen. Dabei handelt es sich um eine Systematisierung von Überlegungen in der Literatur sowie um eigene Erkenntnisse, unter Fortführung früherer Arbeiten.¹⁸ Ferner bilden eigene Tests mit Textgeneratoren und Beob-

¹⁵ Bendel 2019b.

¹⁶ Wendell Wallach/Colin Allen, *Moral Machines: Teaching Robots Right from Wrong*. Oxford 2009; Michael Anderson/Susan Leigh Anderson, *Machine Ethics*, Cambridge 2011; Oliver Bendel, „Maschinenethik“, *Gabler Wirtschaftslexikon*, Wiesbaden 2012, in [www.wirtschaftslexikon.gabler.de/Definition/maschinenethik.html] (Zugriff: 28.09.2023); Bendel 2019a.

¹⁷ Yuntao Bai et al., „Constitutional AI. Harmlessness from AI Feedback“, *ArXiv*, Ithaca 2022, in [www.arxiv.org/abs/2212.08073] (Zugriff: 28.09.2023).

¹⁸ Bendel 2019a.

achtungen in den Communities eine Grundlage, wobei sie kaum repräsentativ sein können und die Tools sich mit ihren stochastischen Prozessen gerade dadurch auszeichnen, dass sie jedes Mal einen anderen Output liefern. Es kann sich um keine vollständige Liste von Themen handeln. Die Ausführungen des Autors an anderer Stelle werden zusammengefasst oder fortgeführt.¹⁹

4.1 Urheberrecht und Fremdnutzung

Die Sprachmodelle, die Textgeneratoren zugrunde liegen, werden zu einem gewissen Anteil mit urheberrechtlich geschütztem Material trainiert. Dies ist aus rechtlicher wie aus ethischer Sicht problematisch.²⁰ So wird die Leistung eines Urhebers genutzt, ohne dass dieser um Erlaubnis gefragt, vergütet oder entschädigt wird.²¹ Zudem hat man keinen Einfluss darauf, in welche Ergebnisse die eigenen Werke einfließen. Eventuell verstoßen sie gegen die eigenen Ziele, Ideen und Werte, oder sie verfehlen den eigenen Geschmack.²² Wenn es sich um geschütztes (und oft aufwändig erarbeitetes) Material von Unternehmen und Organisationen handelt, wird gegen deren wirtschaftliche und anderweitige Interessen verstoßen. Zudem haben sie kaum Einfluss darauf, in welche Resultate die entsprechenden Werke münden. So könnte sich die Konkurrenz in gewisser Weise und in unterschiedlicher Hinsicht an den Leistungen bereichern. Seit 2023 mehren sich die Klagen gegen Anbieter generativer KI, etwa von betroffenen oder angeblich betroffenen Autoren und Künstlern.²³

Informationsethik und Medienethik widmen sich urheberrechtlichen Fragen und schöpferischen Prozessen rund um die Textgeneratoren.²⁴ Die Wirtschaftsethik untersucht die Verschiebungen, die sich durch die Rechtsverletzungen ergeben, u. a. den Aspekt des Missbrauchs durch die unentgeltliche und unrechtmäßige Nutzung

19 Bendel 2023d und Bendel 2024 (s. unten Anm. 23).

20 Jan Smits/Tijn Borghuis, „Generative AI and Intellectual Property Rights“, in *Law and Artificial Intelligence. Information Technology and Law Series*, hg. von Bart Custers/Eduard Fosch-Villaronga, Bd. 35, The Hague 2022, in [www.doi.org/10.1007/978-94-6265-523-2_17] (Zugriff: 28.09.2023).

21 Copyright Office, *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*, 2023, in [<https://t1p.de/mdob8>] (Zugriff: 28.09.2023).

22 Oliver Bendel, „Generative KI aus ethischer Sicht“, in *Generative Künstliche Intelligenz (KI) und Mensch-Maschine-Augmentation*, hg. von Sabine Seufert/Siegfried Handschuh, Stuttgart (erscheint Februar 2024).

23 Johanna Schwanitz, „Für ChatGPT 300 Millionen Wörter aus dem Internet gestohlen“, *FAZ*, 2023, in [<https://t1p.de/naa71>] (Zugriff: 28.09.2023).

24 Oliver Bendel, „Image Synthesis from an Ethical Perspective“, *AI & SOCIETY*, 2023, in [<https://link.springer.com/article/10.1007/s00146-023-01780-4>] (Zugriff: 28.09.2023).

von Daten. Zudem ist die Rechtsethik im Spiel, zumal es sich um Grenzfragen im Recht handelt.

4.2 Urheberschutz von Werken

Mit Blick auf Textgeneratoren wird diskutiert, in einem jeweils unterschiedlichen rechtlichen Rahmen, ob die damit verfertigten Texte Urheberschutz erhalten können.²⁵ Wenn ausreichende menschliche Beteiligung vorhanden ist, bei einer gewissen Schöpfung- oder Gestaltungshöhe, kann dies im Prinzip der Fall sein.²⁶ Möglicherweise reichen schon spezifische Prompts dafür aus, dass ein Urheberschutz entsteht, oder ein komplexes System davon, wie es für Romane oder Sachbücher notwendig ist.²⁷ Hier wären dann die Inputs selbst und die durch sie verursachten Werke schützenswert. Zudem kann eine nachträgliche Bearbeitung von KI-Werken durch Autoren oder Lektoren dazu führen, etwa wenn diese Kapitel umschreiben respektive ergänzen. Es muss letztlich etwas Eigenständiges und Unterscheidbares vorliegen. Auch ein Kuratieren kann wohl in Urheberschutz münden, wenn das damit gebildete Ganze als eigenes kreatives Werk angesehen werden kann.²⁸ Dazu gehören die Sammlung, die Anordnung und die Besprechung von Texten.

Informationsethik und Medienethik widmen sich dem Schutz und der Eigenständigkeit oder Nichteigenständigkeit von Werken.²⁹ Die Wirtschaftsethik untersucht die Möglichkeiten und Missbräuche, die sich durch die Nutzung der Textgeneratoren – etwa durch Mitbewerber – ergeben. Zudem ist die Rechtsethik im Spiel, zumal es sich wiederum um Grenzfragen im Recht handelt.

²⁵ Bendel 2024.

²⁶ Sebastian Grüner, „Urheberschutz bei KI-Nutzung unter Umständen möglich“, *Golem*, 2023, in [www.golem.de/news/us-copyright-office-urheberschutz-bei-ki-nutzung-unter-umstaenden-moeglich-2303-172726.html] (Zugriff: 28.09.2023).

²⁷ Barth 2023.

²⁸ Grüner 2023.

²⁹ Bendel 2023d.

4.3 Datenschutz und informationelle Autonomie bei Prompts

Textgeneratoren verlangen meist nach text-, manchmal auch nach bildbasierten Prompts.³⁰ Zuweilen kann man bei ihnen ganze Dokumente oder Bilder hochladen, auf die sie dann Bezug oder die sie als Ausgangsmaterial nehmen. Prompts verraten viel über die Gesinnung, die Interessen, den Geschmack und letztlich die Person des Benutzers, ganz abgesehen davon, dass häufig durch den Aufruf der Website Aufenthaltsort, Betriebssystem, Browser etc. bekannt werden.³¹ Darin ist der Vorgang den Eingaben in Suchmaschinen nicht unähnlich. Zudem können Prompts persönliche oder betriebliche Daten enthalten. So tippen Mitarbeiter etwa Kundendaten ein oder Daten über das Unternehmen – eine Praxis, die schon zu etlichen betrieblichen Sperren in den USA und in Europa geführt hat.³² Entsprechend werden teilweise Betriebsgeheimnisse und die informationelle Autonomie verletzt.

Informationsethik und Medienethik werfen Fragen in Bezug auf den Datenschutz, den Schutz von Bildern und die informationelle Autonomie bei der Eingabe von Prompts auf.³³ Die Wirtschaftsethik untersucht in moralischer Hinsicht die Möglichkeiten beim Handel von Prompts über Plattformen. Zudem ist die Rechtsethik im Spiel, etwa in Bezug auf die Verletzung von Persönlichkeitsrechten.

4.4 Verantwortung und Haftung

Textgeneratoren können als teilautonome Systeme aufgefasst werden. Die Prompts der Benutzer führen dazu, dass sie etwas generieren, also autonom bearbeiten und automatisch erzeugen. Im Prinzip können sich Textgeneratoren in autonome Systeme verwandeln, etwa wenn die Prompts von der einen Maschine erstellt sowie übertragen und von der anderen Maschine – einer virtuellen wie im Falle von Textgeneratoren oder einer physischen wie im Falle eines Roboters – abgearbeitet werden. Einen auf einem Sprachmodell basierenden Chatbot, bei dem der Dialog in den Vordergrund rückt, kann man ebenfalls als autonom betrachten. Wie bei allen (teil-)autonomen Maschinen ergeben sich Fragen zu Verantwortung im morali-

³⁰ Oliver Bendel, „Prompt“, *Gabler Wirtschaftslexikon*, Wiesbaden 2023a, in [www.wirtschaftslexikon.gabler.de/definition/prompt-125087] (Zugriff: 28.09.2023).

³¹ Bendel 2024.

³² Benjamin Weinmann, „In den USA sperren Grossfirmen ChatGPT: Nun reagieren auch Novartis, Post und Co. – auf unterschiedliche Art“, *Luzerner Zeitung*, 2023, in [https://t1p.de/li3ls] (Zugriff: 28.09.2023).

³³ Bendel 2023d.

schen und rechtlichen Sinne und Haftung im rechtlichen Sinne.³⁴ So kann ein Text, der Falschinformationen oder Halbwahrheiten enthält, im entsprechenden Kontext Schaden anrichten. Auch diskriminierende, rassistische und sexistische Aussagen werfen Fragen auf.

Die Informationsethik (mitsamt KI-Ethik und Roboterethik) fragt nach der Verantwortung von KI-Systemen. Es herrscht Einigkeit darüber, dass diese keine Verantwortung tragen können.³⁵ Diese trägt stets der Mensch, wobei es sich um zahlreiche Entwickler, Vermittler und Betreiber handeln kann, die zudem schwer zu identifizieren sind. Die Rechtsethik mag zu anderen Schlüssen kommen und dabei helfen, eine elektronische Person zu etablieren, die eine Haftung im Zivilrechtlichen übernehmen kann.³⁶

4.5 Stereotype, diskriminierende, rassistische und sexistische Darstellungen

Bei der Bias-Diskussion geht es um Vorurteile und Verzerrungen, die – absichtlich oder unabsichtlich – in die KI-Systeme einfließen und von diesen wiedergegeben werden. So reproduzieren manche Texte zum Beispiel Rollenklischees zu Frauen und Männern oder haben eine politische oder weltanschauliche Schlagseite. Einigen Programmen ist zudem eine kulturelle Voreingenommenheit zu eigen, insofern sie auf der englischen Sprache beruhen und die Texte zum Training zum größten Teil aus dem westlichen Kulturkreis stammen und generierte Outputs deshalb stereotype Konzepte widerspiegeln können.³⁷ Stereotype, Vorurteile und Verzerrungen sind nicht per se ein Problem (beziehungsweise gehören in gewisser Weise zu unserem Alltag), können aber in Beleidigungen und Benachteiligungen münden. Zudem können sie, selbst wenn sie Wirklichkeit widerspiegeln, auch Wirklichkeit formen, und zwar in unerwünschter Weise.

Informationsethik und Medienethik fragen nach dem Entstehen von Stereotypen und diskriminierenden, rassistischen und sexistischen Aussagen bei der Text-

³⁴ Bendel 2024.

³⁵ Claudia Schwarz, „Maschinen können keine moralische Verantwortung tragen“, *Amos international*, 2022, 41–46, in [<https://t1p.de/l62tn>] (Zugriff: 28.09.2023).

³⁶ Susanne Beck, „Künstliche Intelligenz – ethische und rechtliche Herausforderungen“, in *Philosophisches Handbuch Künstliche Intelligenz*, hg. von Klaus Mainzer, 2020, in [https://link.springer.com/referenceworkentry/10.1007/978-3-658-23715-8_29-1] (Zugriff: 28.09.2023).

³⁷ Cris Martin P. Jacoba et al., „Bias and Non-Diversity of Big Data in Artificial Intelligence. Focus on Retinal Diseases“, 2023; Jörg Breithut, „Künstliche Intelligenz. Die Text-zu-Quatsch-Generatoren“, *DER SPIEGEL*, 2022, in [<https://t1p.de/qte9h>] (Zugriff: 28.09.2023).

generierung.³⁸ Zudem ist die Rechtsethik im Spiel, etwa in Bezug auf die Diskriminierung. Die Maschinenethik kann Methoden zur Begrenzung von Anfragen und zur Ausgabe von Inhalten bereitstellen, zum Beispiel Metaregeln und Restriktionen.

4.6 Falschdarstellungen von Sachverhalten

Textgeneratoren neigen auch in der zweiten und dritten Generation zum „Halluzinieren“, wie der mehr oder weniger anerkannte Fachbegriff für das Verbreiten von Unwahrheiten lautet. So tauchen nicht nur falsche Angaben im Text auf, etwa zu Personen und Unternehmen, sondern auch falsche Quellen, sofern welche – meist geschieht dies auf Anfrage – angegeben werden. Es werden zum Beispiel Webadressen erfunden, bei denen die Domain stimmt, der Pfad dagegen nicht. Neuere Generatoren beziehungsweise Versionen versuchen solche Fehler mit verschiedenen Strategien zu vermeiden. Sie fassen sich kürzer, greifen auf Angaben in ihren Wissensdatenbanken zurück oder auf Online- und Livedaten zu. Gerade europäische Anbieter wie Aleph Alpha aus Heidelberg werben mit solchen Ansätzen. Damit schränkt man zum Teil wiederum die Mächtigkeit der Sprachmodelle ein, die – so merkwürdig es klingt – nicht zuletzt in ihrer überbordenden Fantasie besteht.

Informationsethik und Medienethik interessieren sich für Fälschungen, Falschdarstellungen und Falschinformationen im textuellen Bereich.³⁹ Die Politikethik untersucht die Implikationen im Politischen, etwa mit Blick auf die Schwächung der Demokratie, die Wirtschaftsethik diejenigen im Wirtschaftlichen. Zudem ist die Rechtsethik im Spiel, etwa in Bezug auf falsche Tatsachenbehauptungen, Rufmord und Ehrverletzung.

4.7 Weigerungen, Zurückweisungen und Einschränkungen

Bei vielen Textgeneratoren sind Restriktionen im Hinblick auf die Prompts vorhanden.⁴⁰ Wenn diese bestimmte Stichwörter oder Aussagen expliziter oder sexueller Art enthalten, werden sie nicht ausgeführt.⁴¹ Damit werden (meist absichtlich

³⁸ Bendel 2023d.

³⁹ Ebd.

⁴⁰ Bendel 2024.

⁴¹ Bei Bildgeneratoren sind noch andere Methoden vorhanden. So lässt Ideogram alle möglichen Prompts zu, begutachtet dann aber seine Bilder vor der Finalisierung. Wenn es zum Schluss kommt, dass das Ergebnis problematisch sein könnte, wird dieses zensiert. Dabei kann es sich wohl gemerkt

hineinprogrammierte) moralische Regeln der Anbieter angewandt, ohne dass diese im Sinne des Benutzers sein müssten. Manche Anbieter berufen sich darauf, dass in verschiedenen Kulturen und bei verschiedenen Individuen unterschiedliche moralische Vorstellungen existieren, denen man gerecht werden muss. Allerdings könnte man ja in seinem Account oder Profil vermerken, zu welcher Kultur oder Gruppe man gehört oder was man als Individuum will und erträgt, mit der damit eröffneten Option der Personalisierung. Solche Geschäftsmodelle sind aus vielen anderen und auch verwandten Bereichen bekannt, etwa von Chatbots wie Replika mit einer Adultvariante, bei der Dirty Talk und Rollenspiele aller Art möglich sind.

Informationsethik und Medienethik interessieren sich in diesem Zusammenhang für die Zensur der Textgeneratoren, die Wirtschaftsethik speziell für die Verweigerung von Dienstleistungen.⁴² Zudem ist die Rechtsethik im Spiel, die Übergriffigkeit und Beeinflussung beurteilen muss. Es werden moralische Regeln angewandt, sodass die Hersteller implizit oder explizit Maschinenethik treiben.⁴³ Sie werden dies fortführen, wenn sie Constitutional AI anwenden.⁴⁴

4.8 Wissenschaftlichkeit und Referenzierbarkeit

Ungeklärt ist das Problem, wie Texte von Textgeneratoren in journalistischen, belletristischen oder wissenschaftlichen Arbeiten und Artikeln verwendet werden können.⁴⁵ Wenn sie in diese ohne Verweis eingebaut werden, ist die Selbstständigkeit der Leistung nicht mehr vorhanden. Wenn referenziert wird, ist die Frage, worauf – denn es handelt sich nicht um eine eigentliche Veröffentlichung, nicht einmal um graue Literatur.⁴⁶ Die Ergebnisse, die sich von Mal zu Mal ändern, bekommt in der Regel nur der Benutzer angezeigt (lediglich bei Bildgeneratoren sind Communities verbreitet, wo man auch die Ergebnisse der anderen Mitglieder sieht, wie im Falle von Midjourney und Ideogram). Manche Verbände und Hochschulen schlagen vor, sie wie Aussagen aus einem persönlichen Gespräch zu verwenden und entsprechend darauf zu verweisen. So etwas ist jedoch in studentischen und wissenschaftlichen Arbeiten sparsam zu verwenden. Zudem ist in diesem Zusam-

um völlig unverfängliche Prompts handeln, die die generative KI in einer Weise umzusetzen beginnt, die sie selbst nicht zulässt.

42 Bendel 2023d.

43 Bendel 2019a.

44 Bai et al. 2022.

45 Oliver Bendel, „ChatGPT“, *Gabler Wirtschaftslexikon*, Wiesbaden 2023c, in [www.wirtschaftslexikon.gabler.de/definition/chatgpt-124904] (Zugriff: 28.09.2023).

46 Bendel 2024.

menhang nach dem Sinn und Zweck von Zitaten zu fragen – diese sollen ja eigentlich die eigenen Ausführungen ergänzen und belegen, nicht ersetzen.

Die Informationsethik beschäftigt sich mit den moralischen Implikationen der Integration computergenerierter Texte in Texte von Schülern, Studenten und Wissenschaftlern, wie auch die Wissenschaftsethik, die die „neuen Plagiate“ – die vielleicht keine im rechtlichen Sinne sind, indes die Selbstständigkeit der Leistung aushebeln – untersucht, zudem die Veränderung der Originalität und Seriosität wissenschaftlicher Texte durch KI-Systeme.

4.9 Verbesserung und Veränderung der Arbeit

Textgeneratoren haben das Potenzial, Arbeit beziehungsweise Arbeitsergebnisse in inhaltlicher Hinsicht zu verbessern und zu verändern. Dies gilt für Journalisten, Autoren, Texter, zunächst aber für Laien (darunter Schüler und Studenten), die ohne einschlägige Fähigkeiten mit Hilfe geeigneter Prompts hochwertige und ansprechende Werke in die Welt bringen können, die denen der genannten Profis ähneln. Sie erstellen Dokumente, Artikel und Bücher, bei genügender Kontrolle des Outputs. Zudem können sie mit dem KI-System brainstormen und sich eine Struktur vorschlagen lassen. Die Arbeit von Experten kann sich gravierend verändern.⁴⁷ Sie benötigen passende Prompts und damit in jedem Falle sprachliche und metasprachliche Fähigkeiten (die sie natürlich genauso auf entsprechenden Plattformen einkaufen können). Sie müssen die Werke nachbearbeiten, erweitern, zusammenstreichen, neu anordnen etc., damit diese zum Kontext passen. Bei einigen wird all dies die Arbeit produktiver machen, bei anderen unproduktiver. Vermutlich hat die Arbeit nicht mehr die gleiche Eigenständigkeit, Vollständigkeit und Wertigkeit wie vorher. Das Ergebnis kann sich freilich auch verschlechtern, bei Laien wie bei Experten.

Die Medienethik interessiert sich für die Verschiebungen bei den Kompetenzen der Texterstellung und -nutzung, wobei auch das Zusammenspiel von Text- und Bildgeneratoren von Bedeutung ist.⁴⁸ Die Wirtschaftsethik, insbesondere die Unternehmensethik, untersucht die Veränderung der Arbeit (nicht zuletzt der wissenschaftlichen Arbeit und der Aus- und Weiterbildung) und der Arbeitsbeziehungen.

⁴⁷ Thomas H. Davenport/Nitin Mittal, „How Generative AI Is Changing Creative Work“, *Harvard Business Review*, 2022, in [www.hbr.org/2022/11/how-generative-ai-is-changing-creative-work] (Zugriff: 28.09.2023).

⁴⁸ Bendel 2023d.

4.10 Effizienzsteigerung bei der Arbeit

Mit der Verbesserung und Veränderung der Arbeit hängen die Steigerung der Produktivität bzw. die Effizienzsteigerung zusammen. Man kann zum Beispiel als Student oder Wissenschaftler eine wesentlich größere Menge an Texten produzieren oder zusammenfassen und rezipieren. Von ähnlichen Effekten profitieren Texter und Werbefachleute, Marketing- und Kommunikationsabteilungen sowie Autoren, die mit Hilfe komplexer Prompts ganze Bücher erstellen können. Dies führt nicht zuletzt zu Zeit- und Kostenersparnis. Man bewältigt das Studium oder den Berufsalltag schneller und leichter. Zugleich wächst die Gefahr, dass unnötige und oberflächliche Arbeiten entstehen, die etwa das Wissenschaftssystem mit seinem Peer-Review an die Grenze der Belastbarkeit bringen. Auch die Effektivität, die in anderer Hinsicht bereits angesprochen wurde, lässt sich verbessern, etwa hinsichtlich der korrekten Rechtschreibung, die von der Hochschule oder der Community verlangt und bewertet wird, oder hinsichtlich der Sprache eines Organs.

Die Medienethik interessiert sich wiederum für die Verschiebungen bei den Kompetenzen der Texterstellung und -nutzung. Die Wirtschaftsethik untersucht die Veränderung der Arbeit und der Arbeitsbeziehungen, die Wissenschaftsethik die Veränderung der wissenschaftlichen Arbeit und der Aus- und Weiterbildung, zudem die Überlastung des Wissenschaftssystems mit all ihren Konsequenzen.

4.11 Ersetzung und Abschaffung der Arbeit

Bereitstellung und Nutzung von Textgeneratoren gefährden im Extremfall ganze Berufe und Branchen. So ergeben sich Risiken für Journalisten, Autoren und Texter. Lange Zeit galten kreative und künstlerische Berufe als geschützt, was weder im Textbereich noch im Bildbereich mehr der Fall ist, zumindest dort nicht, wo Massenprodukte und Durchschnittsware ausreichen.⁴⁹ Je nach Geschäftsmodell und je nach Verhalten von Privaten und Unternehmen werden etliche Betroffene ihre Arbeit einschränken müssen oder gar verlieren. Ebenso sind Marketing- und Werbeprofis sowie Abteilungen für Öffentlichkeitsarbeit und Kommunikation bedroht. Werbetexte, Wortmarken, CD-Vorlagen etc. lassen sich auf Knopfdruck realisieren, wobei nicht unbedingt Spezialisten in diesem Bereich vonnöten sind, allenfalls für das Überprüfen und Genehmigen. Auch ganze Strategien, etwa zur Markteinführung eines Produkts, vermögen ChatGPT und Co., ein geschicktes

⁴⁹ Davenport/Mittal 2022.

Prompten vorausgesetzt, zu entwickeln. Nicht zuletzt könnten Programmierer – also die Anwender künstlicher Sprachen – verdrängt werden.⁵⁰

Dieser Problematik widmet sich die Medienethik, etwa in Bezug auf die Verschiebung bei den Kompetenzen bei der Medienerstellung.⁵¹ Die Wirtschaftsethik, insbesondere die Unternehmensethik, untersucht die Ersetzung der Arbeit und die Entstehung von Arbeitslosigkeit in den kreativen und administrativen Berufen.

4.12 Unselbstständigkeit des Menschen

Als Taschenrechner in den 1980er-Jahren in den Schulen verteilt wurden, trat die Befürchtung auf, dass die Kinder und Jugendlichen das Rechnen verlernen und es als Erwachsene im Privat- und Berufsleben nicht mehr anwenden könnten.⁵² Textgeneratoren betreffen weniger das Rechnen, sondern mehr das Schreiben, eine andere Kulturtechnik, die sicherlich noch wesentlicher und grundlegender ist. Bei ihrer Nutzung fällt der Akt des Erarbeitens auf dem echten oder virtuellen Papier weg, der für viele Menschen einem lauten Nachdenken und einem erst vorsichtigen, dann immer sichereren Erkunden entspricht. Nun übernimmt das KI-System – nach einer recht kurzen Anweisung – sowohl Prozess als auch Resultat. Ohne großen Aufwand steht man vor einer erheblichen Menge von Text, die es nur noch händisch oder maschinell nachzubearbeiten gilt. Man könnte davon sprechen, dass die Unselbstständigkeit des Menschen durch den Gebrauch dieser Werkzeuge zunimmt.

Die Technikethik stellt zusammen mit der Informationsethik die Frage, wie das Verhältnis von Mensch und Werkzeug ist, ob der Mensch durch die Technik gewinnt oder verliert und wie sich in diesem Zusammenhang die Kulturtechnik des Schreibens entwickelt und verändert.⁵³ Die Informationsethik nimmt speziell die Zunahme der Unselbstständigkeit durch Textgeneratoren unter die Lupe, die mit der Zunahme der Automatisierung einhergeht.

5 Zusammenfassung und Ausblick

Im abschließenden Abschnitt erfolgen eine kurze Zusammenfassung und ein kleiner Ausblick. Der vorliegende Beitrag gab zunächst eine Einführung zu Textgene-

⁵⁰ Patrick Hannemann, „ChatGPT gefährdet viele Berufe: Diese Jobs werden sich völlig verändern“, *CHIP*, 2023, in [<https://t1p.de/y7fdu>] (Zugriff: 28.09.2023).

⁵¹ Bendel 2023d.

⁵² Bendel 2024.

⁵³ Bendel 2023d.

ratoren. Dann wurden ethische und am Rande soziale und rechtliche Aspekte behandelt. Es konnten insgesamt zahlreiche Risiken herausgearbeitet werden, jedoch auch Chancen wie die Unterstützung der Arbeit. Das Transformationspotenzial für einen wichtigen Bereich der kreativen und wissenschaftlichen Berufe wurde offensichtlich. Es handelt sich förmlich um eine zweite Digitalisierungswelle, die den Text erfasst.

Weitere damit zusammenhängende Punkte sind ebenfalls von Interesse. So sind die meist textbasierten Prompts ein Grund, dem Kulturwerkzeug Sprache neue Aufmerksamkeit zu widmen.⁵⁴ Gefragt sind äußerst präzise und prägnante Beschreibungen, sodass die Sprachfähigkeit des Benutzers gefordert ist und geschult werden muss. Eine hier nicht behandelte Frage ist, wie damit umzugehen ist, dass Übersetzungs- und Redigierprogramme wie DeepL und DeepL Write sowie manche Textgeneratoren (auch amerikanischer Provenienz) häufig gegenderte Texte mit Gendersternchen oder Doppelpunkt im Wortinneren nicht als Fehler deklarieren oder sogar propagieren beziehungsweise produzieren – und damit die durch den Rechtschreibrat fixierte Normsprache schwächen.⁵⁵

Alle Themen können Gebieten der angewandten Ethik zugeordnet werden, die sich spezialisiert und eigene Begriffe (z.T. sogar eigene Methoden) herausgebildet haben, also den Bereichsethiken⁵⁶ und der Maschinenethik.⁵⁷ Natürlich sind genauso Begriffe und Konzepte der allgemeinen, empirischen wie normativen Ethik relevant. So kann man mit der Menschenwürde argumentieren, die mit Texten verletzt wird, oder mit Gleichheit und Gleichberechtigung. Begriffe wie Diskriminierung, Rassismus und Sexismus können auf die genannten Bereichsethiken bezogen werden, haben aber eine Verankerung in der Ethik insgesamt und sind Thema weiterer Bereichsethiken wie der Politikethik.

Neue Versionen von Sprachmodellen könnten die Mächtigkeit der Textgeneratoren erhöhen. Es ist freilich damit zu rechnen, dass die Qualität abnimmt, wenn KI-Systeme immer mehr mit KI-Content trainiert werden. Es gibt unterschiedliche Möglichkeiten, dieser Gefahr zu begegnen, etwa durch die Hinzunahme von Bild- und Audiodaten und von Livedaten – sozusagen frisches Blut für den KI-Vampir. Die

54 Bendel 2024.

55 Das Ergebnis eigener Tests ist, dass DeepL Write in wesentlichen Aspekten versagt, in Bezug auf Groß- und Kleinschreibung, Zusammen- und Getrennschreibung sowie die amtlich fixierten Regeln. Es werden zum Teil eigene Regeln angewandt, die nicht einmal transparent gemacht werden. Damit ist der Nutzen des Tools grundsätzlich in Frage gestellt. Auf der Website heißt es: „Schreiben Sie klar, präzise und fehlerfrei“. Aber das kann man offenbar besser, wenn man sich auf sein Sprachgefühl und den Rechtschreibrat verlässt.

56 Bendel 2019b.

57 Bendel 2019a, 2012a.

Multimodalität von ChatGPT seit Herbst 2023 ist in genau diesem Sinne: Man kann mit der Anwendung sprechen und ihr Bilder zeigen (und von ihr das Dargestellte und damit zusammenhängende Hintergründe und Vorgänge erklären lassen). Eine andere Bedrohung ist, dass die Textgeneratoren zu sehr eingehegt werden, sei es aus rechtlichen, sei es aus ethischen oder moralischen Gründen. Es braucht zumindest Geschäftsmodelle, die auch riskante und explizite Ein- und Ausgaben erlauben. Konkurrenz dürfte das Geschäft insgesamt beleben, und wenn mehr und mehr Open-Source-Generatoren entstehen, werden die Marktmacht der Konzerne und die damit einhergehende Abhängigkeit ein Stück weit zurückgedrängt. Die Risiken von Textgeneratoren sind erheblich, die Chancen aber so offensichtlich, dass Homo faber auf solche Werkzeuge kaum verzichten wird.