Medienethische Perspektiven auf generative Künstliche Intelligenz: Nachhaltigkeit und das „gute Leben" im Fokus

Sigrid Kannengießer, Anne Mollen

Abstract

In diesem Beitrag diskutieren wir generative Künstliche Intelligenz aus einer medienethischen Perspektive und arbeiten Werte heraus, die in der Produktion und Aneignung generativer KI, aber auch in ihre Materialität, also den Infrastrukturen generativer KI, eingeschrieben sind. Dabei setzen wir solche Werte in den Fokus, die mit Blick auf Nachhaltigkeit und der Frage, wie generative KI nachhaltig(er) gestaltet werden kann, besonders relevant sind: Gerechtigkeit, Freiheit und Selbstbestimmung sowie Würde. Um diese Werte zu diskutieren, knüpfen wir an bestehende Forschung zu Nachhaltigkeit und digitalen Medientechnologien sowie ethischen Ansätzen aus dem Diskurs um generative KI an. Medienethische Fragen sind im wissenschaftlichen und gesellschaftspolitischen Diskurs zunehmend relevant, da generative KI komplexe negative sozial-ökologisch-ökonomische Folgen verursacht, die es gilt, zu adressieren, fragt man nach einem „guten Leben" mit generativer KI und danach, wie generative KI nachhaltig(er) gestaltet werden kann.

1. Einleitung

Seit der Antike beschäftigen sich Philosophen mit der Frage, was das „gute Leben" ist (Horn 2021). Fragen wir heute nach dem „guten Leben", so gerät Nachhaltigkeit ins Zentrum, stehen Gesellschaften doch vor der aktuellen und dringenden Herausforderung, durch eine nachhaltige Entwicklung die Lebensgrundlage heutiger und zukünftiger Lebewesen zu schützen (UN WCED 1987) und ein gutes Leben für alle zu ermöglichen. Die entsprechende politische Zielsetzung findet sich in den Zielen für nachhaltige Entwicklung der Vereinten Nationen (United Nations 2015). Nachhaltigkeit verstehen wir als gesellschaftlichen Wert, also gesellschaftliches Ziel (Funiok 2016: 322), das auf Gerechtigkeit in der sozialen, ökologischen, ökonomischen und kulturellen Dimension abzielt.

In der Kommunikations- und Medienwissenschaft wird das „gute Leben" überwiegend als individuelles Wohlbefinden aus medienpsychologischer Perspektive untersucht (z.B. Brey et al. 2012) und Nachhaltigkeit wurde vor allem als Thema in Medieninhalten in den Blick genommen (z.B. Bonfadelli 2007), bzgl. Kommunikator*innen (Weder und Eriksen 2023) und Rezipierenden (Kannengießer 2021). Zunehmend wird auch die Materiali-

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

tät digitaler Medientechnologien fokussiert und gefragt, wie diese nachhaltiger gestaltet und genutzt werden können (z.B. Kannengießer 2022). Diese Frage kann mit der Verbreitung von Technologien automatisierter Kommunikation, also generativer Künstliche Intelligenz wie z.B. ChatGPT, modifiziert werden. Denn es ergeben sich medien- und kommunikationsethische Fragestellungen, die sich direkt aus der Materialität generativer KI, ihren Infrastrukturen (Mollen/Kannengießer *im Druck*) und den zugrundeliegenden Sprachmodellen (LLMs), als sogenannte string-prediction-Systeme (Bender et al. 2021), die auf der Basis riesiger Datenmengen Wortabfolgen berechnen und neue Texte generieren, ableiten lassen. Negative Folgen generativer KI zeigen sich mit Blick auf Fragen von ökologischer, sozialer und ökonomischer Nachhaltigkeit, beispielsweise durch große Mengen an Kohlenstoffdioxid, die durch das Training der Modelle und der Nutzung (und damit verheerenden Folgen für den Klimawandel) emittiert werden, Bias und Diskriminierung in Datensätzen sowie Formen kultureller Hegemonie und prekären Arbeitsbedingungen der „Click-Worker:innen" (oftmals Frauen), die die Datensätze moderieren (Mollen/Kannengießer *im Druck*)[1]. Daraus resultierend ergeben sich viele Forschungsfragen für die Kommunikations- und Medienwissenschaft, die zum einen die sozialen, ökologischen und ökonomischen Folgen der Produktion und Nutzung generativer KI verstehen und Praktiken der nachhaltigen Gestaltung generativer KI analysieren und auch transformativ entwickeln muss. In diesem Kapitel spitzen wir diese Fragen aus einer medienethischen Perspektive zu und arbeiten die gesellschaftlichen Werte heraus, die sich in der Materialität generativer KI sowie deren Gestaltung und Aneignung manifestieren. Dabei fokussieren wir die folgenden Werte: Gerechtigkeit, Freiheit und Selbstbestimmung sowie Würde. Um diese Werte zu diskutieren, knüpfen wir an bestehende Forschung zu Nachhaltigkeit und digitalen Medientechnologien (Kannengießer 2022) sowie ethischen Ansätzen aus dem Diskurs um die Nachhaltigkeit generativer KI an (Rohde et al. 2024).

1 Die Strukturierung der Folgen der Produktion und Nutzung generativer KI entlang des Drei-Säulen-Modells der Nachhaltigkeit (Hauff und Clauss 2018) ist relevant, um u.a. aufzuzeigen, dass diese nicht nur in der ökologischen Dimension durch die zunehmend diskutierte Energieintensität und damit Kohlenstoffdioxidemissionen verortet sind.

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

2. Nachhaltige generative KI und Medienethik

Wenn generative KI zunehmend in den Gegenstandsbereich der Kommunikations- und Medienwissenschaften gerät, dann wird sie auch Gegenstand medienethischer Betrachtungen. Die Technologie rückt hierbei auf mindestens dreierlei Art in den Fokus: Erstens, als Medientechnologie, mit der Nutzer:innen direkt interagieren und über die sie Medieninhalte rezipieren können – was ethische Fragen der Nutzung aufwirft. Zweitens, als Technologie, die zunehmend in der Medienproduktion eingesetzt wird, beispielsweise in journalistischen Produktionsprozessen. In diesem Zusammenhang fordert generative KI nicht nur bestehende Wertvorstellungen, beispielsweise mit Blick auf eine journalistische Berufsethik, heraus. Drittens, führt die spezifische Materialität generativer KI, die sich in ihren Infrastrukturen manifestiert, zu teils völlig neuen Herausforderungen, die sie von anderen digitalen Medientechnologien unterscheidet. Mitunter erfordert generative KI eine Neu-Aushandlung moralischer Ordnungen.

Auch wenn sich an den Schnittstellen von Disziplinen wie dem Machine Learning, der Philosophie und auch der Rechtswissenschaft bereits eine intensive Diskussion zur „Ethik von Künstlicher Intelligenz" im Allgemeinen und „Ethik von generativer KI" im Speziellen ausgebildet hat (Hagendorff 2020; Mittelstadt et al. 2016), bedarf es einer kommunikations- und medienethischen Perspektive, die die drei benannten Bereiche aus einer medienethischen Perspektive kritisch reflektiert und dabei Fragen von Nachhaltigkeit als zentralem gesellschaftlichen Wert fokussiert. Diese kann nicht nur drängende medienethische Fragen mit Blick auf klassische medienbezogene Anwendungsfelder von generativer KI adressieren, sondern auch eine kontextualisierte ethische Perspektive abbilden, die sich nicht nur auf die technischen Systeme von generativer KI beschränkt, sondern über die Rolle generativer KI als Medien in der Gesellschaft auch weiterführende ethische Implikationen in den Blick nimmt.

Mit Fokus auf die Rolle generativer KI für die Produktion von Medieninhalten steht in den bereits bestehenden ethikbezogenen Diskussionen zu Algorithmen und generativer KI ein soziales Nachhaltigkeitsverständnis im Vordergrund. Dies beschreibt Aspekte generativer KI, die eine Rolle dafür spielen, inwiefern Menschen Freiheiten und Möglichkeiten haben, eine gute Lebensgrundlage für sich umsetzen zu können oder inwiefern sie sozial ausgeschlossen sind (Sen 2000, Nussbaum 2006). Die Diskussion rund um generative KI verweist hier auf eine etablierte Diskussion rund um mögliche Formen der Diskriminierung und Reproduktion von gesellschaftlichen

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Stereotypen (Barocas et al. 2019; Zuiderveen Borgesius 2018) insbesondere aufgrund von Bias in den KI-Trainingsdaten (Bender et al. 2021). Darüber hinaus zeigt sich eine Dominanz anglo-amerikanischer Sichtweisen und Wertvorstellungen in generativer KI, die sich ebenfalls in nicht repräsentativer Darstellung bzw. der fehlenden Repräsentanz bestimmter Kulturen (beispielsweise gewisser Sprachen) äußert (Bellio 2023; Luo/Puett 2024).

Rückt die Materialität von Medientechnologien, und in diesem Fall generativer KI, in den Vordergrund, so zeigt sich eine dezidierte Nachhaltigkeitsdiskussion, die auch unter eben jenem Begriff firmiert, vor allem mit Blick auf die ökologischen Folgen von generativer KI. Hier sind die teils immensen Ressourcenverbräuche, insbesondere Energie (und entsprechend Emissionen) sowie der Wasserverbrauch in Rechenzentren, in den Fokus technisch-ausgerichteter Disziplinen gerückt (z.B. Kaack et al. 2022; Li et al. 2023). Sozialwissenschaftliche Analysen verweisen darüber hinaus unter anderem auf die Kontinuität kolonialer Ausbeutungsstrukturen, mit Blick auf Umwelt, Ressourcen, Arbeitskraft etc. (Bresnihan/Brodie 2021; Gray/Siddharth 2019; Lehuedé 2024), für die für generative KI notwendigen Infrastrukturen.

Aus einer ökonomischen Nachhaltigkeitsperspektive wird ebenfalls die Materialität von generativer KI relevant. Denn insbesondere die immens hohe Rechenleistung, die für die Entwicklung und das Training generativer KI notwendig ist, sowie der Zugang zu Trainingsdatensätzen als auch der Zugang zu Arbeitskräften mit der entsprechenden Expertise für die Entwicklung von generativer KI schafft extrem hohe Eintrittsbarrieren in den generativen KI-Markt (Gray Widder et al. 2023). Dies führt zu einem sehr konzentrierten Markt (Vipra/Korinek 2023). Da beispielsweise journalistische Texte eine wichtige Rolle für das Training von generativer KI und somit für Fragen des Zugangs und der Marktkonzentration spielen, gilt es hier medienethische Reflexionen anzustoßen. Diese können in Kontinuität zu lang etablierten Diskussionen rund um Medienkonzentration im Fach stehen.

Die politisch institutionalisierten Werterahmen, auf die sich die Diskussion rund um generative KI und Nachhaltigkeit in der Regel bezieht, sind die Ziele für nachhaltige Entwicklung (SDGs). Dies passiert jedoch in der Regel sehr funktional gedacht, unter der Prämisse, inwiefern KI potenziell zur Erreichung der SDGs beitragen kann (Vinuesa et al. 2020). Eine Perspektive, die vor allem Potenziale, nicht aber die vielfältigen negativen Auswirkungen, Hindernisse und Zielkonflikte eines Einsatzes von KI für

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Nachhaltigkeitszwecke berücksichtigt, ist notwendigerweise unzureichend und muss sich den Vorwurf des Technosolutionsimus gefallen lassen (Sætra 2022, 2023). Für eine medienethische Betrachtung ist es daher umso entscheidender, auch die möglichen Werterahmen zur Bewertung von generativer KI in den Blick zu nehmen. Diese können anwendungsspezifische Aushandlungen umfassen, wie sie aktuell beispielsweise mit der Paris Charter[2] für den Journalismus formuliert wurden. Aber auch die SDGs sollten weiterhin als Werterahmen herangezogen werden und die Nachhaltigkeitsimplikationen von generativer KI in ihrer konkreten Umsetzung und in der Verhandlung von Zielkonflikten vor diesem Rahmen analysiert werden. Auch wenn die SDGs weder ausreichend noch widerspruchsfrei sind, so bilden sie dennoch einen relevanten Kontext für medienethische Perspektiven auf generative KI und insbesondere ihre Infrastrukturen, nicht zuletzt, weil die Ziele einen politischen Wertekanon implizieren, der handlungsleitend für gesellschaftliche Entwicklungen sein sollte.

3. Zentrale Werte für eine nachhaltige Gestaltung generativer KI

Die Kommunikations- und Medienwissenschaft beschäftigt sich in medienethischen Arbeiten mit Moral, Werten und Normen in den Bereichen der Medienproduktion, Medieninhalte und Medienrezeption. Digitale Medien und Onlinemedien haben den Forschungsgegenstand der Medienethik verändert: Charles Ess spricht von einer „digitalen Medienethik" (Ess 2014) und meint damit eine Medienethik, die digitale Medien in den Fokus setzt, denn die Ethik selber wird nicht digital. Prinzing et al. konstatieren gar eine „Neuvermessung der Medienethik" (Prinzing et al. 2015). Mit der Verbreitung generativer KI gilt es, eine Ethik generativer KI zu entwickeln, da sich zum einen zwar medienethische Fragen aus dem Diskurs um digitale Medientechnologien auch mit Blick auf generative KI fortsetzen, diese neuen Technologien aber auch durch ihre Spezifika neue medienethische Fragen provozieren.

Bringen wir unsere eigenen medienethischen Arbeiten zu Nachhaltigkeit und digitalen Medientechnologien (Kannengießer 2022) und generativer KI (Rohde et al. 2024) zusammen, so zeigen sich auf Grundlage der bestehenden Forschung zu Nachhaltigkeit und Digitalisierung sowie generativer KI, dass folgende Werte besonders relevant sind, da sie zum einen durch

2 https://rsf.org/en/paris-charter-ai-and-journalism

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

die jetzige Art und Weise der Produktion, Bereitstellung und Nutzung generativer KI nicht erfüllt sind: Gerechtigkeit, Freiheit und Selbstbestimmung sowie Würde. Wir zeigen in den folgenden Abschnitten, wo aus medienethischer Perspektive Probleme und Herausforderungen bei der Erfüllung dieser Werte in der Produktion und Aneignung generativer KI liegen. Dabei gehen wir im Folgenden nur auf Herausforderungen für die genannten Werte durch generative KI ein, die sich in der Produktion, Aneignung und in der Materialität der Systeme manifestiert. Wir sprechen nicht über den missbräuchlichen Gebrauch von generativer KI, beispielsweise durch Akteure, die KI für Wahlmanipulation, politische Falschinformation oder Deep Fakes einsetzen. Dieser Missbrauch von generativer KI hat ebenfalls umfassende Auswirkungen auf Gerechtigkeit, Selbstbestimmung, Freiheit und Würde – steht in diesem Beitrag allerdings nicht im Fokus.

3.1 Gerechtigkeit

Gerechtigkeit ist einer der zentralen Werte, wenn Nachhaltigkeit im Fokus steht, definieren wir diese doch als sozial-ökologisch-ökonomische Gerechtigkeit (s.o.). Krainer (2018: 320-321) betont, dass es verschiedene medienethische Betrachtungsmöglichkeiten des Gerechtigkeitsbegriffs gebe. Zu diesen gehören z.B. medien- und kommunikationsrelevante Gerechtigkeiten wie gerechte Zugangsmöglichkeiten zu Medien sowie ein gerechter Besitz medialer Produktionsmittel und eine gleichberechtigte Repräsentation unterschiedlicher Personen oder Meinungen innerhalb der Medieninhalte (ebd.). Weitere Studien, die die Produktion und Aneignung digitaler Technologien aus einer Nachhaltigkeitsperspektive analysieren, betonen die Relevanz von Fragen nach Gerechtigkeit in Produktions- und Entsorgungsprozessen digitaler Medientechnologien (Kannengießer 2022: 174-185).

Unterscheiden wir im Sinne Frasers (2003) zwischen Anerkennungs- und Verteilungsgerechtigkeit, so werden mit Blick auf generative KI sowohl Fragen gerechter Repräsentationen in den Outputs generativer KI relevant als auch Fragen nach globaler Verteilungsgerechtigkeit und der ungleichen Verteilung von Produktion und Nutzung generativer KI.

Die Diskussion um Gerechtigkeitsfragen bei den Outputs generativer KI schließt an kommunikations- und medienwissenschaftliche Perspektiven an, die die Inhalte in den Blick nehmen, die durch generative KI produziert werden. So zeigen sich am Beispiel von diskriminierenden und stereotypen

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Outputs die Ungleichbehandlungen von oft bereits marginalisierten Gruppen durch generative KI. Sexualisierende Darstellungen von nicht-weißen Frauen[3] oder die stereotype Darstellung von Geschlechtern beispielsweise im Kontext ihrer Berufe[4], weisen auf eine Schlechterstellung von bestimmten Personengruppen im Vergleich mit anderen. Die Ursachen liegen unter anderem in den Datensätzen, in denen die diskriminierenden und stereotypen Darstellungen begründet liegen (Barocas et al. 2019). Ebenfalls ausgehend von den Datensätzen, zeigt sich eine Ungleichbehandlung bestimmter Wissensbestände in generativer KI (Luo/Puett 2024) – aufgrund der Dominanz anglo-amerikanisch geprägter Datensätze, die zu einer fehlenden Repräsentation nicht westlicher und marginalisierter Wissensbestände führt und diese auch technisch manifestiert.

In diesem Zusammenhang lassen sich auch weitere Fragen von globaler Verteilungsungerechtigkeit und der ungleichen Verteilung von Produktion und Nutzung generativer KI diskutieren. Der Umstand, dass generative KI vor allem für Märkte im globalen Norden mit westlichen Wissensbeständen produziert wird, während prekäre und nicht-nachhaltige Produktionspraktiken in den globalen Süden verlegt werden, knüpft an die bereits skizzierten Diskussionen zu Nachhaltigkeit und Materialität in der Kommunikations- und Medienwissenschaft an. So bleibt eine Perspektive auf gerechte Produktionsbedingungen, unter denen diese Technologien entwickelt und hergestellt werden, aber auch gerechte Bedingungen bei der Entsorgung von Technologien, die für die Entwicklung, Produktion und Aneignung generativer KI genutzt wurden, weiterhin relevant – insbesondere angesichts ihres immensen Hardware- und Unterhaltbedarfs (Luccioni et al. 2022; Strubell et al. 2019).

Aus Nachhaltigkeitsperspektive werden dann nicht nur die direkt zuschreibbaren globalen Verteilungsungerechtigkeiten mit Blick auf generative KI relevant, sondern auch generelle Klimaauswirkungen. Denn während generative KI mit massiven Kohlenstoffdioxidemissionen einhergeht, die große Tech-Unternehmen bereits jetzt zur Anpassung ihrer Klimaziele gezwungen haben, profitiert ein Markt im globalen Norden, während die massivsten Auswirkungen des Klimawandels im globalen Süden zu spüren sind. Hier kam aus der Machine-Learning Forschung bereits vor einigen Jahren der kritische Zwischenruf, die Größe von Sprachmodellen, die die

3 https://www.technologyreview.com/2022/12/13/1064810/how-it-feels-to-be-sexually-objectified-by-an-ai/

4 https://huggingface.co/spaces/society-ethics/DiffusionBiasExplorer

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Grundlage von generativer KI bilden, vor dem Hintergrund dieser globalen Auswirkungen zu reduzieren (Bender et al. 2021). Kleinere Sprachmodelle und kleinere Datensätze reduzieren nicht nur den Energieverbrauch drastisch, sondern können auch durch sorgfältig kuratierte Datensätze Formen der Diskriminierung reduzieren. Die Kommunikations- und Medienwissenschaft beginnt diese Thematik und insbesondere den technischen Systemen nachgelagerte Nachhaltigkeitsauswirkungen in den Blick zu nehmen (Mollen et al. in Vorbereitung)[5].

3.2 Selbstbestimmung und Freiheit

In der Kommunikations- und Medienwissenschaft wird Selbstbestimmung meist als informationelle Selbstbestimmung, also als Kontrolle der Menschen über Informationen gedacht (Heesen 2017). Dabei geht es zum einen um die Kontrolle im Sinne eines Zugangs zu Informationen, zum anderen aber auch um die Kontrolle im Sinne eines Datenschutzes über Informationen der eigenen Person. Neben den Diskursen zur informationellen Selbstbestimmung haben sich allgemeine Diskussionen zu Selbstbestimmung insbesondere mit Blick auf digitale Medien entwickelt. Hier werden zunehmend die Begriffe digitale Souveränität, Autonomie oder Selbstbestimmung sowohl austauschbar als auch in sich widersprüchlich verwendet (Couture/Toupin 2019; Müller et al. 2022). Die gemeinsame Grundlage dieser unterschiedlichen Ansätze ist die Vorstellung, dass individuelle und kollektive Akteure „das Recht, die Fähigkeiten und die Chancen Einzelner, ihr Leben nach eigenen Vorstellungen zu gestalten" (Wessler et al. 2018) haben. Es geht also um die Fähigkeit, freien Willen mit Blick auf digitale Medientechnologien auszuüben in Abwesenheit von Fremdbestimmung oder -kontrolle.

Damit liegen Fragen der Selbstbestimmung ganz nah bei medienethischen Fragen der Freiheit. Rath (2018: 193) betont, dass „Freiheit die Bedingung moralischer Beurteilung überhaupt [ist] und damit auch die Voraussetzung für eine ethische Reflexion auf die Prinzipien, die einer Moral zugrunde liegen". Freiheit im Sinne eines nicht-manipulativen oder fremdbestimmten Umgangs mit digitalen Technologien, greift die oben

5 Der Beitrag befindet sich derzeit in Vorbereitung: Mollen, Anne/Kannengießer, Sigrid/Velkova, Julia/Jansen, Fieke: AI Infrastructures and Sustainability: Expanding perspectives on automation, communication and media. London: Palgrave Macmillan.

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

genannten Diskussionen unter dem Begriff der digitalen Autonomie bzw. Souveränität auf (Pohle/Thiel 2020; Sahebi/Formosa, 2022). Der Fokus liegt hier nicht auf Freiheit als Grundbedingung einer ethischen Reflexion, sondern Freiheit im Umgang mit digitalen Technologien. Bei Fragen der nachhaltigen Gestaltung von generativer KI werden beide Aspekte des hier skizzierten Freiheitsbegriffs relevant – sowohl Freiheit in der Anwendung als auch in der Reflexion.

Das Thema der Kontrolle über die eigenen Daten (informationelle Selbstbestimmung) wird beispielsweise auch mit Blick auf generative KI wieder aufgegriffen. Aktuell beobachten wir juristische Auseinandersetzungen rund um die Nutzung von sowohl personenbezogenen als auch urheberrechtlich geschützten Daten für das Training von generativer KI. Problematisch ist in diesem Zusammenhang die fehlende Transparenz vieler Anbieter generativer KI, die weder Auskunft über die verwendeten Datensätze noch detaillierte Informationen zur ihren generativen KI-Systemen veröffentlichen. Dies erschwert Versuche von individuellen und kollektiven Akteur:innen, selbstbestimmt zu handeln, ihre Selbstbestimmung zu verteidigen sowie zu juristischen Neuaushandlungen des Rechts auf Urheberschaft und informationeller Selbstbestimmung.

Die Intransparenz von generativer KI beeinflusst Freiheit und Selbstbestimmung auch mit Blick auf die fehlende Diversität von Angeboten. Während manche Anbieter wie Meta, mehr oder weniger freiwillig, ihre generativen KI-Systeme Open Source zur Verfügung stellen, lehnen andere Anbieter offene Systeme vehement ab. Wenngleich weiterhin unklar bleibt, was eigentlich genau Offenheit mit Blick auf generative KI bedeutet (Solaiman 2023), so führt Intransparenz und Geschlossenheit der Systeme vor allem zu noch höheren Markteintrittsbarrieren (Gray Widder et al. 2023), so dass die Entwicklung generativer KI-Systeme in der Hand weniger Tech-Unternehmen bleibt. Die Konsequenz ist, dass trotz einer zunehmenden Zahl von generativen KI-Modellen, die ihnen zugrunde liegenden Werte und moralischen Ordnungen eingeschränkt bleiben. So werden sowohl Selbstbestimmung mit Blick auf alternative Entwicklungen von generativen KI-Systemen als auch die Auswahlmöglichkeiten der Endnutzenden beschränkt.

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

3.3 Würde

Ebenfalls muss man Überschneidungen beim Wert der Würde und Aspekten von Freiheit und Selbstbestimmung konstatieren. In der neuzeitlichen Philosophie sowie in aktuellen politischen Normen wie den Menschenrechten bezieht sich Würde vor allem auf die Autonomie und Selbstbestimmung des Einzelnen (Bohrmann 2018: 55, siehe hier auch eine Begriffsgeschichte). Mit Blick auf generative KI lassen sich zu diesem Punkt die bereits kurz angerissenen Fragen zu gerechten Produktions- und Entsorgungsprozessen der Technologie nennen. Selbstbestimmung wird hier in Hinblick auf die Selbstbestimmung von Menschen bezogen, die an den Produktions- und Entsorgungsprozessen von (digitalen) Medientechnologien beteiligt sind (Kannengießer 2022). Die Produktions- und Entsorgungsprozesse tangieren im Kern letztlich auch Fragen der Würde, da es hier im Sinne der Kriterien der International Labour Organization um „decent work“[6] also würdevolle Arbeitsbedingungen geht. Statt von würdevoller Arbeit lässt sich aber nicht nur mit Blick auf den Abbau der Ressourcen für die für generative KI notwendige Hardware von prekären Arbeitsverhältnissen sprechen. Auch die Art und Weise der Arbeit von sogenannten Click-Worker:innen, die beispielsweise Daten für das Training von KI oder die Outputs von generativer KI kuratieren, ist extrem belastend für die Akteur:innen (oftmals Frauen), die oft unterbezahlt, ohne entsprechende Arbeitsschutz-Maßnahmen wie psychologische Betreuung arbeiten (Gray/Siddharth 2019). Durch die Marktmacht von großen Technologieunternehmen als Auftraggeber für Subunternehmen, oftmals im globalen Süden, sehen sich Arbeitende in Machtabhängigkeiten (Miceli et al. 2022a; Miceli et al. 2022b) und somit in ihrer Selbstbestimmung und Würde eingeschränkt.

Würde wird darüber hinaus in der Kommunikations- und Medienwissenschaft v.a. mit Blick auf die Medieninhalte diskutiert und nach der Menschenwürde in den Repräsentationen der Medieninhalte gefragt (Bohrmann 2018). Die Diskussion um diskriminierende und stereotype Darstellungen durch generative KI betrifft somit nicht nur Fragen der Gerechtigkeit im Sinne einer Ungleichbehandlung in den Outputs generativer KI (siehe oben). Sie wird ebenfalls vor dem Hintergrund einer Diskussion um eine würdevolle Darstellung und gesellschaftliche Repräsentation relevant. Hierbei geht es nicht nur um Fragen der Schlechterstellung bestimmter

6 https://www.ilo.org/topics/decent-work

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Gruppen, sondern auch um die kulturell unangemessene Darstellung bestimmter Personen und Personengruppen. Ein prägendes Beispiel ist die Diskussion um „The American Smile" – die sich auf generierte Bilder von historischen Darstellungen beispielsweise von Native Americans mit breitem Lachen bezieht. Dieses Lachen ist typisch für moderne nordamerikanische Portraits, aber völlig unpassend in der historischen Darstellung von indigenen Gruppen[7].

Versuche von Anbietern generativer KI, Repräsentationen durch Anpassung der technischen Systeme angemessener zu gestalten, haben in der jüngsten Vergangenheit zu historisch nicht korrekten Darstellungen geführt[8], die Fragen der Diskriminierung und Misrepräsentation noch einmal neu angefacht haben. Letztlich zeigen diese gescheiterten Versuche, dass generative KI-Systeme in ihrer Architektur grundlegende Probleme mit kultureller Diversität und Sensibilität haben – und somit relevant sind für eine Diskussion von Würde auf der Ebene von Medieninhalten.

4. Schlussgedanken

Fragt man heute mit einer medienethischen Perspektive auf generative KI nach dem „guten Leben", so stehen Fragen zur Nachhaltigkeit im Fokus. Denn generative KI verursacht komplexe negative soziale-ökologische-ökonomische Folgen, die sichtbar gemacht und adressiert werden müssen, will man generative KI nachhaltiger gestalten. In diesem Beitrag wurden auf der Basis einer Skizzierung der negativen Folgen generativer KI, Werte herausgearbeitet, die aus einer medienethischen Perspektive für den Zusammenhang von Nachhaltigkeit und generativer KI besonders relevant werden: Gerechtigkeit, Freiheit und Selbstbestimmung sowie Würde. Dabei haben wir exemplarische Herausforderungen für die genannten Werte aufgezeigt, die sich in der Produktion, Aneignung und in der Materialität der Systeme manifestieren. Diese Herausforderungen gilt es zu überwinden, wenn generative KI in der Produktion, Aneignung und ihrer Materialität nachhaltig(er) und im sozial-ökologisch-ökonomischen Sinne gerechter gestaltet werden soll.

Der Schlüssel einer solchen nachhaltigeren Gestaltung liegt letztendlich in einer Gemeinwohlorientierung, die besagt, dass „überindividuelle (ge-

7 https://medium.com/@socialcreature/ai-and-the-american-smile-76d23a0fbfaf
8 https://www.nytimes.com/2024/02/22/technology/google-gemini-german-uniforms.html

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

meinsame, öffentliche) Interessen Maßstäbe des Handelns sein können und sollen" (Filipović 2017). Dem Konzept der Gemeinwohlökonomie ist die Idee inhärent, dass Unternehmen nicht mehr kapitalistisch orientiert sind (also gewinnorientiert im Sinne der Akkumulation von Kapital), sondern orientiert an dem „Streben nach dem allgemeinen Wohl" (Felber 2010: 24). Es bedarf einer solchen Orientierung an dem Wohl der Allgemeinheit, um die oben benannten Werte in der Gestaltung der Infrastrukturen generativer KI zu manifestieren.

Literatur

Barocas, Solon/Hardt, Moritz/Narayanan, Arvind (2019): *Fairness and machine learning: limitations and opportunities.* Online verfügbar unter: https://fairmlbook.org/ (Abruf am: 18.11.2024)

Bellio, Naiara (2023): *ChatGPT-like models boom, but small languages remain in shadows. AlgorithmWatch.* Online verfügbar unter: https://algorithmwatch.org/en/chatgpt-models-and-small-languages/ (Abruf am: 18.11.2024)

Bender, Emily M./Gebru, Timnit/McMillan-Major, Angelina/Shmitchell, Shmargaret (2021): *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, S. 610–623. Online verfügbar unter doi: 10.1145/3442188.3445922

Bohrmann, Thomas (2018): *Würde* (Teil 11). In: Communicatio Socialis 51 (1), S. 54–59. Online verfügbar unter doi: 10.5771/0010-3497-2018-1-54

Bonfadelli, Heinz (2007): *Nachhaltigkeit als Herausforderung für Medien und Journalismus.* In: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.), Nachhaltigkeitsforschung - Perspektiven der Sozial- und Geisteswissenschaften. Bern, S. 255–297.

Bresnihan, Patrick/Brodie, Patrick (2021): *New extractive frontiers in Ireland and the moebius strip of wind/data.* In: Environment and Planning E: Nature and Space 4 (4), S. 1645–1664. Online verfügbar unter doi: 10.1177/2514848620970121

Brey, Philip/Briggle, Adam/Spence, Edward (Hg.) (2012): *The Good Life in a Technological Age.* New York: Routledge. Online verfügbar unter doi: 10.4324/9780203124581

Constantinides, Panos/Henfridsson, Ola/Parker, Geoffrey G. (2018): *Introduction—Platforms and Infrastructures in the Digital Age.* In: Information Systems Research 29 (2), S. 381–400. Online verfügbar unter doi: 10.1287/isre.2018.0794

Couture, Stephane/Toupin, Sophie (2019): *What does the notion of "sovereignty" mean when referring to the digital?* In: New Media & Society 21 (10), S. 2305–2322. Online verfügbar unter doi: 10.1177/1461444819865984

Ess, Charles (2014): *Digital media ethics.* 2. Auflage. o.O.: Polity Press.

Felber, Christian (2010): *Gemeinwohl-Ökonomie. Das Wirtschaftsmodell der Zukunft.* Wien: Deuticke.

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Filipović, Alexander (2017): *Gemeinwohl als medienethischer Begriff. Über öffentliche Kommunikation und gesellschaftliche Mitverantwortung.* In: Communicatio Socialis 50 (1), S. 9–19. Online verfügbar unter doi: 10.5771/0010-3497-2017-1-9

Fraser, Nancy (2003): *Soziale Gerechtigkeit im Zeitalter der Identitätspolitik. Umverteilung, Anerkennung und Beteiligung.* In: Fraser, Nancy/Honneth, Axel (Hg.): Umverteilung oder Anerkennung? Eine politisch-philosophische Kontroverse. Frankfurt a.M.: Suhrkamp, S. 13-128.

Funiok, Rüdiger (2016): *Werte* (Teil 5). In: Communicatio Socialis 49 (3), S. 322–326. Online verfügbar unter doi: 10.5771/0010-3497-2016-3-322

Gray, Mary L./Siddharth, Suri (2019): *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass.* Houghton Mifflin Harcourt.

Gray Widder, David/West, Sarah/Whittaker, Meredith (2023): *Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI.* Akzeptiert für eine Veröffentlichung in Nature. Online verfügbar unter doi: 10.2139/ssrn.4543807

Hagendorff, Thilo (2020): *The Ethics of AI Ethics: An Evaluation of Guidelines.* In: Minds and Machines 30 (1), S. 99–120. Online verfügbar unter doi: 10.1007/s11023-020-09517-8

Heesen, Jessica (2017): *Informationelle Selbstbestimmung* (Teil 10). In: Communicatio Socialis 50 (4), S. 495–500. Online verfügbar unter doi: 10.5771/0010-3497-2017-4-495

Horn, Christoph (2021): *Das gute Leben mit Platon und Aristoteles: Ursula Wolfs philosophische Aufsätze.* In: Deutsche Zeitschrift Für Philosophie 69 (2), S. 317–325. Online verfügbar unter doi: 10.1515/dzph-2021-0026

Kaack, Lynn H./Donti, Priya L./Strubell, Emma/Kamiya, George/Creutzig, Felix/Rolnick, David (2022): *Aligning artificial intelligence with climate change mitigation.* In: Nature Climate Change 12(6), S. 518–527. Online verfügbar unter doi: 10.1038/s41558-022-01377-7

Kannengießer, Sigrid (2021). *Media Reception, Media Effects and Media Practices in Sustainability Communication: State of Research and Research Gaps.* In: Weder, Franzisca/Krainer, Larissa/Karmasin, Matthias (Hg.) The Sustainability Communication Reader. Springer VS, Wiesbaden. Online verfügbar unter doi: 10.1007/978-3-658-31883-3_18

Kannengießer, Sigrid (2022): *Digitale Medien und Nachhaltigkeit: Medienpraktiken für ein gutes Leben.* Wiesbaden: Springer Nature.

Krainer, Larissa (2018): *Gerechtigkeit* (Teil 13). In: Communicatio Socialis 51 (3), S. 319–324. Online verfügbar unter doi: 10.5771/0010-3497-2018-3-319

Lehuedé, Sebastián (2024): *An alternative planetary future? Digital sovereignty frameworks and the decolonial option.* In: Big Data & Society 11 (1). Online verfügbar unter doi: 10.1177/20539517231221778

Li, Pengfei/Yang, Jianyi/Islam, Mohammad A./Ren, Shaolei (2023): *Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models.* Online verfügbar unter doi: 10.48550/ARXIV.2304.03271

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Luccioni, Alexandra S./Viguier, Sylvain/Ligozat, Anne-Laure (2022): *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model.* Online verfügbar unter doi: 10.48550/arXiv.2211.02001

Luo, Queenie/Puett, Michael (2024): *Anglo-American bias could make generative AI an invisible intellectual cage.* In: Nature 629 (8014), S. 1004–1004. Online verfügbar unter doi: 10.1038/d41586-024-01573-9

Miceli, Milagros/Posada, Julian/Yang, Tianling (2022a): *Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?* Proceedings of the ACM on Human-Computer Interaction 6 (GROUP), S. 1–14. Online verfügbar unter doi: 10.1145/3492853

Miceli, Milagros/Yang, Tianling/Garcia, Adriana Alvarado/Posada, Julian/Wang, Sonja Mei/ Pohl, Marc/Hanna, Alex (2022b): *Documenting Data Production Processes: A Participatory Approach for Data Work.* Proceedings of the ACM on Human-Computer Interaction, 6 (CSCW2). Online verfügbar unter: 10.48550/arXiv.2207.04958

Mittelstadt, Brent Daniel/Allo, Patrick/Taddeo, Mariarosaria/Wachter, Sandra/Floridi, Luciano (2016): *The ethics of algorithms: Mapping the debate.* In: Big Data & Society 3 (2). Online verfügbar unter doi: 10.1177/2053951716679679

Nussbaum, Martha C. (2006). *Education and democratic citizenship: Capabilities and quality education.* In: Journal of human development 7 (3), S. 385-395. Online verfügbar unter doi: 10.1080/14649880600815974

Mollen, Anne/Kannengießer, Sigrid (*im Druck*): *Generative KI und Nachhaltigkeit als Generationengerechtigkeit.* In: Hepp, Andreas/Kannengießer, Sigrid/Pfadenhauer, Michaela/Wimmer, Jeffrey (Hg.) Zukunft der Medien – Medien der Zukunft. Wiesbaden: Springer VS.

Müller, Jane/Tischer, Moritz/Thumel, Mareike/Petschner, Paul (2022): *Unboxing digitale Souveränität.* In: Medienimpulse, 60 (4), 37 Seiten. Online verfügbar unter doi: 10.21243/MI-04-22

Pohle, Julia/Thiel, Thorsten (2020): *Digital sovereignty*. In: Internet Policy Review 9 (4). Online verfügbar unter doi: 10.14763/2020.4.1532

Prinzing, Marlis/Rath, Matthias/Schicha, Christian/Stapf, Ingrid (Hg.) (2015): *Neuvermessung der Medienethik: Bilanz, Themen und Herausforderungen seit 2000*. Weinheim: Beltz Juventa.

Rath, Matthias (2018): *Freiheit* (Teil 12). In: Communicatio Socialis 51 (2), S. 192–198. Online verfügbar unter doi: 10.5771/0010-3497-2018-2-192

Rohde, Friederike/Wagner, Josephin/Meyer, Andreas/Reinhard, Philipp/Voss, Marcus/ Petschow, Ulrich/Mollen, Anne (2024): *Broadening the perspective for sustainable artificial intelligence: sustainability criteria and indicators for Artificial Intelligence systems.* In: Current Opinion in Environmental Sustainability 66. Online verfügbar unter doi: 10.1016/j.cosust.2023.101411

Sen, Amartya (2000): *Social exclusion: Concept, application, and scrutiny.* Social Development Papers No. 1. Manila: Office of Environment and Social Development Asian Development Bank

Sætra, Henrik Skaug (2022): *AI for the sustainable development goals*. Boca Raton, Abingdon: CRC Press. Online verfügbar unter doi: 10.1201/9781003193180

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Sætra, Henrik Skaug (2023): Technology and Sustainable Development: The Promise and Pitfalls of Techno-Solutionism. New York: Routledge. Online verfügbar unter doi: 10.1201/9781003325086

Sahebi, Siavosh/Formosa, Paul (2022): Social Media and its Negative Impacts on Autonomy. In: Philosophy & Technology 35, 70. Online verfügbar unter doi: 10.1007/s13347-022-00567-7

Solaiman, Irene (2023): *The Gradient of Generative AI Release: Methods and Considerations.* 2023 ACM Conference on Fairness, Accountability, and Transparency, S. 111–122. Online verfügbar unter doi: 10.1145/3593013.3593981

United Nations (2015). *Transforming our World: The 2030 Agenda for Sustainable Development.* Online verfügbar unter: https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf (Abruf am 18.11.2024)

UN WCED (1987). *Our common future. World Commission on Environment and Development.* Online verfügbar unter: https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf (Abruf am 18.11.2024)

von Hauff, Michael/Claus, Katja (2018): *Fair Trade: Ein Konzept nachhaltigen Handels.* 3. vollständig überarbeitete Auflage; Studienausgabe. Konstanz, München: UVK/Lucius

Vinuesa, Ricardo/Azizpour, Hossein/Leite, Iolanda/Balaam, Madeline/Dignum, Virginia/Domisch, Sami/Felländer, Anna/Langhans, Simone Daniela/Tegmark, Max/Fuso Nerini, Francesco (2020): *The role of artificial intelligence in achieving the Sustainable Development Goals.* In: Nature Communications 11 (1), S. 1–10. Online verfügbar unter doi: 10.1038/s41467-019-14108-y

Vipra, Jai/Korinek, Anton (2023): *Market concentration implications of foundation models: THE INVISIBLE HAND OF CHATGPT.* Center on Regulation and Markets Working Paper #9). Washington: Brookings. Online verfügbar unter: https://www.brookings.edu/wp-content/uploads/2023/09/market-concentration-implications-of-foundation-models-final-1.pdf (Abfrage am 18.11.2024)

Weder, Franziska/Eriksen, Marte (2023): *CSR Communication and Cultures of Sustainability.* St Lucia: The University of Queensland Pressbook. Online verfügbar unter doi: 10.14264/428d81b.

Wessler, Hartmut/Haffner, Patrik/Rinke, Eike Mark (2018): *Selbstbestimmung in der digitalen Welt. Über die Vorteile eines ebenenübergreifenden normativen Basiskonzepts für die empirische Erforschung der digitalen Kommunikation*. In: Medien & Kommunikationswissenschaft 66 (4), S. 395–406. Online verfügbar unter doi: 10.5771/1615-634X-2018-4-395

Zuiderveen Borgesius, Frederik (2018): *Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy*. Straßburg: Directorate General of Democracy. Online verfügbar unter: https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73 (Abfrage am 18.11.2024)

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

https://doi.org/10.5771/9783748940616-289 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY