

Der Design-Justice-Ansatz mit einer Anwendung im Bereich der KI

Felicitas Macgilchrist

1 Einleitung

Wenn es um künstliche Intelligenz (KI) und Bildung geht, vor allem in Zeiten der *generativen KI*, werden oft Fragen gestellt nach den Vor- und Nachteilen der Nutzung, ob ChatGPT Einfluss auf die Noten der Lernenden haben könnte oder wie KI die Unterrichtsvorbereitung, die Organisation von Klassenfahrten oder die Verwaltung der Schule verändern bzw. erleichtern kann. Der Design-Justice-(Designgerechtigkeits-)Ansatz ist eine Perspektive, die nahelegt, dass diese Fragen nur einige Aspekte der aktuellen Entwicklungen betreffen und von weiteren wichtigen sozialen, bildungspraktischen und gesellschaftlichen Dimensionen ablenken. In diesem Beitrag wird der Design-Justice-Ansatz vorgestellt und anhand des Beispiels generativer KI wie *ChatGPT* exemplifiziert. Es geht also in diesem Beitrag nicht um *EdTech* im Sinne von explizit für Bildungszwecke entwickelten Medientechnologien, sondern um Technologien, die durch ihren Gebrauch in (formalen) Bildungseinrichtungen zu EdTech werden (Macgilchrist et al. 2023).

Um die zentrale Idee vorwegzunehmen: Mit diesem Ansatz wird nicht nach den allgemeinen Vor- und Nachteilen (Chancen und Herausforderungen, Potenziale und Risiken) von Bildungsmedien bzw. -technologien (EdTech) gefragt, sondern EdTech wird spezifischer danach befragt,

- für wen, für was, in welchen Situationen und wann diese EdTech *Vorteile* bringt;
- für wen, für was, in welchen Situationen und wann sie *Nachteile* bringt und
- wer daran *beteiligt* war oder ist, diese Medientechnologie zu entwickeln.

Design Justice bündelt diese drei Fragen in der Kurzform: ›Who participates, who benefits, and who is harmed?‹ (Wer ist beteiligt, wer hat einen Nutzen und wer trägt die Last?) (Costanza-Chock 2020: 134). Nachdem in diesem Beitrag einführend der Design-Justice-Ansatz in Abschnitt 2 vorgestellt wird, wird in Abschnitt 3 zur Veranschaulichung eine Analyse von ChatGPT durchgeführt. In Abschnitt 4 werden Alternativen reflektiert, die gestaltet werden könnten, bevor in Abschnitt 5 ein Fazit gezogen wird.

2 Der Design-Justice-Ansatz

Design entfaltet eine große Wirkkraft auf unser Leben, denn es geht um die Gestaltung der (großen und auch alltäglichsten) Dinge, Praktiken und Kontexte um uns herum. Im Kontext dieses Beitrags, der den Bereich Bildung fokussiert, bezieht sich Design vor allem auf Hardware, Software und Unterrichtssequenzen. Das Wort ›Design‹ kann sich sowohl auf die entstandenen Dinge wie z.B. eine Lernapp für Mathematik als auch auf den Prozess des Gestaltens wie z.B. die Entscheidungen, welche Farben, welche Grafiken oder welche dahinterliegenden Systeme einzubauen sind, beziehen. Somit eröffnet Design Möglichkeiten für Handeln, Sein, Beziehungen und Begehren und schließt andere Möglichkeiten aus (Escobar 2018). Wird z.B. eine Quizapp, die wie *Kahoot!* designed ist, im Unterricht eingesetzt, kann das motivierend und energetisierend sein, es eröffnet die Möglichkeit für Schüler*innen, schnell zu reagieren, viel zu lachen und sich als Gewinner*in auf der Siegetreppe sehen zu wollen, aber es schließt gleichzeitig ein kooperatives, entschleunigendes Nachdenken aus. Breiter betrachtet »formt [Design] unsere Fähigkeit, auf die Welt zuzugreifen, an ihr teilzuhaben und zu ihr beizutragen« (Holmes 2020). Design determiniert aber unsere Welt nicht (siehe auch den Beitrag ›Nicht determinierend, aber doch stark regulierend: Eine Studie zu Design und pädagogischen Wirkungen der ›Antolinx‹-Leserlernplattform‹ von Sigrid Hartong in diesem Band).

Für das Design-Justice-Netzwerk,¹ ein internationales, praxis- und forschungsorientiertes Netzwerk aus Designer*innen, Forscher*innen, Aktivist*innen, Lehrer*innen und weiteren Praktiker*innen, ist die zentrale Problematik beim Design, dass typischerweise in Gestaltungsprozessen sehr

1 Informationen vom Design Justice Network befinden sich hier: <https://designjustice.org/>.

wenige Menschen aktiv involviert sind. Vor allem die Menschen, die die negativsten Folgen der Designs erfahren, werden selten an der Entwicklung beteiligt. Den Hintergrund des Ansatzes bilden Theorien der (Un-)Gerechtigkeit und Ungleichheit (Collins 1990; Crenshaw 1991). Somit erfüllt laut Sasha Costanza-Chock in ihrem Buch *Design Justice: Community-Led Practices to Build the World We Need* Design Justice eine doppelte Funktion:

Design Justice ist ein *Rahmen für die Analyse*, wie Design Nutzen und Lasten zwischen verschiedenen Gruppen von Menschen verteilt. Design Justice konzentriert sich explizit auf die Art und Weise, wie Design die Matrix der Herrschaft (weiße Vorherrschaft, Heteropatriarchat, Kapitalismus, Behindertenfeindlichkeit, Siedler*innenkolonialismus und andere Formen der strukturellen Ungleichheit) reproduziert und/oder herausfordert. Design Justice ist auch eine wachsende *Praxisgemeinschaft*, die darauf abzielt, eine gerechtere Verteilung der Vorteile und Lasten von Design, eine sinnvolle Beteiligung an Designentscheidungen und die Anerkennung von gemeinschaftsbasierten, Indigenen und diasporischen Designtraditionen, -wissen und -praktiken sicherzustellen. (Costanza-Chock 2020: 23, Übersetzung und Hervorhebungen FM)

Design Justice ›überdenkt [also] Designprozesse, stellt Menschen in den Mittelpunkt, die normalerweise durch Design marginalisiert werden, und nutzt kollaborative, kreative Praktiken, um die größten Herausforderungen unserer Gemeinschaften anzugehen.² Es geht unter anderem darum, nicht nur – wie z.B. in einigen *Design-Thinking*-Ansätzen (Macgilchrist et al. 2023); (siehe auch den Beitrag ›Was steckt in Lernmanagementsystemen und welche Implikationen ergeben sich hieraus für eine digital mündige Gestaltung im Schulkontext? Eine Studie zu LMS.lernen.hamburg‹ von Nina Brandau in diesem Band) – an den ›prototypischen‹ Nutzenden zu denken, wenn Unterricht, Software, Plattformen usw. gestaltet werden. Stattdessen priorisiert das Gestaltungsteam die potenziellen Nutzungserfahrungen der am stärksten marginalisierten Nutzer*innen, das heißt die Erfahrungen der Menschen, die eine ›Design-Mikroaggression‹ oder andere ›Unangemessenheiten‹ des Designs erleben könnten bzw. deren positive Erfahrung durch das Design selbst irritiert oder blockiert wird (Costanza-Chock 2020: 39, 45).

Ein Beispiel für die Produktivität der Doppelfunktion von Design Justice als ›Rahmen für die Analyse‹ und ›Praxisgemeinschaft‹ wäre das Design eines

2 Text von hier übersetzt: <https://designjustice.org/principles-overview>.

klassischen Schulbuchs. Design beinhaltet auch die Entscheidungen, welche Aufgaben im Buch erscheinen sollen. In unterschiedlicher Ausprägung war eine Aufgabe bis vor kurzem in vielen Sozialkunde-, Politik- oder Geschichtsschulbüchern zu finden: ›Fragt ggf. eure ausländischen Mitschülerinnen und Mitschüler, was für sie Integration bedeutet.‹ (Beauftragte der Bundesregierung für Migration 2015: 43) Hier wird die Aufgabe nur an die Schüler*innen adressiert, die nicht als ›ausländisch‹ gelten. Die Aufgabe ›verändert‹ (macht zu etwas ›anderem‹ als der ›Norm‹) Schüler*innen, die durch die Aufgabe als ›ausländisch‹ markiert werden. Auch die Definition von Ausländer*in verteilt Lasten auf unter anderem diejenigen Schüler*innen mit deutschem Pass, die aber unter dem Begriff ›Ausländer*in‹ Diskriminierungs- bzw. Rassismuserfahrungen ausgesetzt sind. Die Design-Justice-Praxisgemeinschaft würde darauf abzielen, Bildungsmedien wie Schulbücher mit Blick auf die marginalisierten Positionen (das heißt, wer in der Aufgabe nicht adressiert wird) anders zu gestalten und somit den Nutzen und die Lasten anders zu verteilen.

Das Design Justice Network hat zehn (allgemeine) Prinzipien für die Gestaltung entwickelt, die für diese Neugestaltung der Schulbücher produktiv gemacht werden könnten.³ Um nur einige hier zu erwähnen:

- ›Im Fokus stehen die Stimmen derjenigen, die direkt von den Ergebnissen des Designprozesses betroffen sind.‹
- ›Veränderung entsteht für uns durch einen nachvollziehbaren, zugänglichen und kollaborativen Prozess, und nicht als ein Punkt am Ende des Prozesses.‹
- ›Wir glauben, dass jede Person aufgrund ihrer eigenen Lebenserfahrungen ein*e Expert*in ist, und dass wir alle einzigartige Beiträge in einen Designprozess einbringen können.‹

Als Analyserahmen wird der Design-Justice-Ansatz bereits in Schulen und weiteren Bildungseinrichtungen eingesetzt. Im ›Justice by Design‹-Projekt (JusticexDesign, JxD) haben z.B. Sara Sheya und ihr Team an der Harvard Universität gemeinsam mit lokalen Schulen in den USA eine Reihe von Arbeitsblättern entwickelt, um die Sensibilität junger Schüler*innen für Fragen der Design Justice zu schärfen (Sheya 2020). Die Arbeitsblätter leiten Lehrpersonen und

3 Die zehn Prinzipien sind auf Deutsch hier zugänglich: <https://designjustice.org/german>.

Schüler*innen an, über Macht, (Unterdrückungs-)Systeme und Möglichkeiten für gerechte Zukünfte nachzudenken. Sie laden dazu ein, darüber nachzudenken, für wen bestimmte Designs sind und was sie ermöglichen. Und zugleich, wen diese Designs blockieren und was sie verunmöglichen. In einem Arbeitsblatt sollen sich beispielsweise Schüler*innen die Inhalte eines Bildungsmediums vornehmen (z.B. einen Podcast, eine Schulbuchseite, eine App, eine Webseite oder das automatisch erstellte Output nach einem ChatGPT-Prompt, siehe unten) und danach fragen, *wessen Perspektiven* vertreten sind, wessen Stimmen am lautesten sind und woran sie das erkennen. Sie sollen auch fragen, wessen Perspektiven *nicht* vertreten sind und wessen Stimmen am leisesten sind. Weitere Arbeitsblätter (die bei JxD ›Denkroutinen‹ oder ›Protokolle‹ genannt werden) beinhalten Fragen wie: ›Welche Auswirkungen haben die Botschaften [des Bildungsmediums, Anm. FM] auf dich und/oder deine Gemeinschaft?‹, ›Was wissen wir über das Artefakt und das historische System, das es darstellt?‹ und ›Welche gestalterischen Entscheidungen haben die Entwickler*innen dieses Inhalts getroffen?‹⁴. Insgesamt zielen diese Arbeitsblätter für Schüler*innen – unterstützt durch weiterführende Hinweise für Lehrkräfte zum Einsatz im Unterricht – darauf ab, Schüler*innen anzuregen, eingespielte Machthierarchien in der Gesellschaft, die in schulischen Medientechnologien reproduziert werden, zu hinterfragen. Sie zielen auch darauf an, Schüler*innen dabei zu unterstützen, sich alternative Möglichkeiten – für Medien und Gesellschaft – vorzustellen.

Aber nicht nur Schüler*innen, auch Lehrkräfte, Wissenschaftler*innen und weitere Personen können Design Justice für die konkrete Analyse von EdTech-Produkten einsetzen. Um die Potenziale des Design-Justice-Ansatzes für die Analyse vertiefter zu illustrieren, wird im folgenden Abschnitt das Beispiel ChatGPT anhand der Dreifachfrage, die eingangs erwähnt wurde, analysiert. Auf Englisch: ›Who participates, who benefits, and who is harmed?‹ (Costanza-Chock 2020: 134). In meiner deutschen Übersetzung: Wer ist beteiligt, wer hat einen Nutzen und wer trägt die Last?

4 Übersetzungen im Fließtext alle von FM. JusticexDesign Materialien hier (auf Englisch) zugänglich: <https://www.justicexdesign.org/tools>.

3 Design Justice anwenden am Beispiel generative KI

Seit Ende 2022 wird *generative KI* in der Öffentlichkeit kontrovers diskutiert. Bei textgenerierenden Chatbots bezieht sich der Begriff auf große Sprachmodelle, die darauf zielen, Inhalte zu erstellen, die aussehen, als wären sie von Menschen kreiert. Es ist üblich geworden, Texte über generative KI mit der Antwort einer generativen KI beginnen zu lassen. In diesem Beitrag verwende ich die Selbstbeschreibung von ChatGPT als Basis für die Analyse der Beteiligung, Nutzen und Lasten der Technologie. Auf meiner eingetippten Frage (das sogenannte *Prompt*) ›Was ist ChatGPT?‹ generierte der Chatbot einen Text, der wie folgt anfängt:

ChatGPT ist ein AI-Modell, das auf der GPT-3.5-Architektur von OpenAI basiert. ›ChatGPT‹ bezieht sich auf die Fähigkeit dieses Modells, *menschenähnlichen Text* in natürlicher Sprache zu generieren und auf Anfragen oder Fragen zu antworten. Es wurde entwickelt, um in Textkommunikationsszenarien wie Chatbots, virtuellen Assistenten, Kundensupport-Anwendungen und mehr eingesetzt zu werden.

ChatGPT verwendet ein tiefes neuronales Netzwerk, das auf *umfangreichen Textdaten trainiert* wurde, um menschenähnlichen Text zu generieren und *natürliche Konversationen mit Benutzern* zu führen. Es kann in verschiedenen Bereichen wie Textgenerierung, Übersetzung, Zusammenfassung von Texten, Beantwortung von Fragen und mehr eingesetzt werden. (ChatGPT, Auszug aus der Antwort auf der Frage ›Was ist ChatGPT?‹ am 10.9.2023, Hervorhebungen FM)

Wer ist beteiligt?

Diese Antwort von ChatGPT deutet auf drei Beteiligte, die durch die folgende Analyse näher betrachtet werden können. Zum einen ist *OpenAI* involviert. Im Jahre 2015 hat *OpenAI*, ein US-amerikanisches Non-Profit-Unternehmen, Zusagen von Spenden in Höhe von 1 Milliarde Dollar erhalten, um künstliche Intelligenz (KI) auf gemeinnütziger Basis zu entwickeln und der Öffentlichkeit zur Verfügung zu stellen. Die Presse berichtete, dass Elon Musk, CEO von *Tesla Motors* und *SpaceX*, Peter Thiel, Mitbegründer von *Paypal*, der indische Tech-Riese *Infosys* und *Amazon Web Services* zu den damaligen Geldgebern gehörten.⁵

5 <https://www.bbc.com/news/technology-35082344>

Seit 2019 bezieht sich der Name *OpenAI* allerdings auf die gewinnorientierte *OpenAI LP*, eine Tochtergesellschaft des nun kleineren *OpenAI Nonprofit*.⁶

Im Alltag werden Designentscheidungen von den Beschäftigten bei *OpenAI* getroffen. Obwohl keine genauen Daten zu *OpenAI* vorliegen, zeigen Berichte zur Beschäftigung in der Region (San Francisco, USA), dass Frauen, Schwarze Menschen⁷, Latinx und weitere minorisierte Gruppen in der Tech-Branche unterrepräsentiert sind. Frauen machen beispielsweise 49 % der Erwerbsbevölkerung in San Francisco aus, aber nur 26 % der Tech-Beschäftigten. Schwarze Menschen machen 7 % der Gesamtbeschäftigten in der Region aus und 3 % der Beschäftigten in der Technologiebranche; Latinx haben 24 % aller Arbeitsplätze in San Francisco, aber nur 8 % der Tech-Beschäftigten (CompTIA (Computing Technology Industry Association) 2021: 132). Insgesamt ist die Tech-Branche in dieser Region also sehr durch die ›Dominanzgesellschaft‹ geprägt, mit der entsprechenden Wahrscheinlichkeit, dass Diskriminierungs- und Marginalisierungserfahrungen nicht zum Alltag der Entwickler*innen und Geschäftsführung gehören und somit nicht auf diese Weise ihre Entscheidungen prägen.⁸

Die oben zitierte Antwort von ChatGPT deutet auch auf ein zweites Set an beteiligten Personen hin: Diese stecken hinter dem ›tiefe[n] neuronale[n] Netzwerk, das auf *umfangreichen Textdaten trainiert* wurde‹. Woraus diese Trainingsdaten genau bestehen, macht *OpenAI* nicht transparent, aber sie basieren primär auf Texten aus dem Internet, das heißt Webseiten, Artikeln, Forenbeiträgen, sozialen Medien, Onlinebüchern und vielem mehr. Auch hier reproduziert sich die fehlende Diversität, diesmal in Bezug auf die Inhalte der Ant-

6 <https://openai.com/blog/openai-lp>

7 Als Selbstbezeichnung beschreibt ›Schwarze Menschen‹ eine von Rassismus betroffene gesellschaftliche Position. ›Schwarz‹ wird großgeschrieben, um zu verdeutlichen, dass es sich um ein konstruiertes Zuordnungsmuster handelt und keine reelle ›Eigenschaft‹, die auf die Farbe der Haut zurückzuführen ist (Shearer und Haruna 2013). Da weiß auch ein konstruiertes Zuordnungsmuster ist, wird hier der dekolonialen Nutzung gefolgt und das Wort klein und kursiv gesetzt (vgl. <https://www.amnesty.de/glossar-fuer-diskriminierungssensible-sprache>).

8 Eine Aufmerksamkeit für Marginalisierungserfahrungen soll hiermit nicht auf Identitätszuschreibungen reduziert werden, zudem es vielfältige weitere Gründe für Diskriminierungserfahrungen gibt. Eine fehlende Sensibilisierung oder Aktivismus in der Techbranche in Bezug auf Diversität, Rassismuskritik usw. ist allerdings vielfach kritisiert worden (Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press).

worten des Chatbots. Die Mehrheit der Texte, die online zur Verfügung stehen, wird von relativ jungen, *weißen* Männern der Mittelschicht des Globalen Nordens verfasst. Sie geben somit ein bestimmtes Set an Normen, Werten und Biases (Verzerrung) der Dominanzgesellschaft wieder (Bender et al. 2021). Die Antworten von ChatGPT reproduzieren unhinterfragt und unkommentiert dieses Dominanzwissen. Da die Trainingsdaten wenig minoritäre Perspektiven enthalten, kann der Chatbot wenig diverse Antworten geben. Ein Beispiel ist in dem Auszug von ChatGPT oben zu sehen: Auf Deutsch gendert ChatGPT kaum. Es geht um ›Benutzer‹, nicht um Benutzer*innen, Benutzer_innen oder weitere plurale Formen. Es gibt in den Ausgaben des Chatbots generell ›den‹ Arzt, ›den‹ Professor und ›die‹ Krankenschwester.⁹ Die Antworten enthalten in erster Linie auch nur zwei Geschlechter. Je mehr Texte von ChatGPT oder weiteren generativen KI-Maschinen produziert werden und in der Öffentlichkeit zirkulieren (z.B. als Werbetexte, Unterrichtsentwürfe, Einführungstexte oder auch Bilder), desto häufiger werden unhinterfragte und marginalisierende Normen in Alltag und Schule auftauchen.

Diese Trainingsdaten weisen wiederum auf ein drittes Set an beteiligten Personen hin: Personen, die ihre – oft unsichtbare – manuelle Arbeitskraft in die Entwicklung von automatisierten Systemen einfließen lassen. Ein wesentlicher Unterschied zwischen ChatGPT und früheren Chatbots, die mit Textdaten aus dem Internet trainiert wurden, ist, dass ChatGPT kaum unangemessene (das heißt explizit rassistische, sexualisierte, gewaltvolle) Aussagen macht. Der Grund dafür ist, dass OpenAI Verfahren zur ›Filterung‹ und ›Reinigung‹ der Daten entwickelt bzw. eingebaut hat.¹⁰ Diese Verfahren beinhalten nicht nur automatische, sondern auch manuelle Schritte von sogenannten ›Annotators‹ (Markov et al. 2022). Im Januar 2023 berichtete das *Time Magazine*, dass outgesourcete Arbeiter*innen in Kenia ca. 1,50 bis 2 EUR/Stunde verdienen,

9 Im Englischen war dies mit dem folgenden Prompt besonders sichtbar: ›Analyse the following sentence: The doctor apologized to the nurse because she was late.‹ Jede Ausgabe von ChatGPT ist anders, aber in jeder gab es monatelang eine Variante davon, dass ›der Arzt‹ sich entschuldigt, weil ›die Krankenschwester‹ verspätet war. Es analysierte den Satz nie mit ›der Ärztin‹ und ›der Pfleger‹, z.B. ›In this sentence, the doctor apologized to the nurse for being late. Let's break it down: [...] So, in summary, the sentence is saying that the doctor said sorry to the nurse because the nurse was late.‹ (ChatGPT, Mai 2023)

10 Hier beschrieben: <https://openai.com/blog/new-and-improved-content-moderation-tooling> (letzter Zugang: 14.09.2023)

um gewalttätige, sexualisierte und weitere toxische Inhalte als solche zu labeln (Perrigo 2023). Josh Dzeiza (2023) beschrieb ihre Arbeit im Detail. Das bedeutet: Die Nutzer*innen von ChatGPT sehen keine toxischen Texte, weil diese Annotator*innen jeden Tag mit solchen Texten gearbeitet haben.

In diesem Abschnitt wurde gefragt: Wer ist an ChatGPT beteiligt? Sofern die Leser*innen dieses Beitrags den Design-Justice-Ansatz selbst zur Analyse einsetzen, können sie an dieser Stelle die Analyse fortführen und weitere Beteiligte identifizieren. Ziel ist es unter anderem, ihren Beitrag zur Reproduktion oder Irritation dominanter gesellschaftlicher Diskurse zu reflektieren und zu überlegen, wessen Perspektiven nicht oder nur selten einbezogen werden.

Wer hat einen Nutzen?

Auf meinem Prompt antwortet der Chatbot, dass ChatGPT unter anderem ›in verschiedenen Bereichen wie Textgenerierung, Übersetzung, Zusammenfassung von Texten, Beantwortung von Fragen und mehr eingesetzt werden‹ kann (siehe oben). Einen Nutzen hätten somit diejenigen Personen, die Texte generieren, übersetzen und zusammenfassen, und Personen, die Fragen beantwortet haben möchten. Wer mit ChatGPT herumexperimentiert, hat zum Zeitpunkt des Schreibens dieses Beitrags erfahren, dass es sehr gute, einfache, generische Antworten auf Fragen erstellen kann, aber auch, dass diese Antworten oft fehlerhaft oder lückenhaft sind und dass die Antworten auf Gebieten, in denen man selbst Expertise hat, selten zufriedenstellend sind. Nachrichtensender oder Rechtsanwälte, die ChatGPT eingesetzt haben, sind beispielsweise in Schwierigkeiten geraten, weil die Informationen und Quellen in ihren Texten nicht korrekt waren (Bohannon 2023; Farhi 2023).

Textsorten, die sehr gut erstellt werden können, sind Spam-E-Mails, generische Werbetexte und weitere nicht faktenbasierte Texte. Personen, zu dessen Arbeit die Erstellung solcher Textsorten gehört, sehen durch die Nutzung von ChatGPT eine Effizienzsteigerung in ihrem Aufgabengebiet. Auch Unternehmen, die diese Art von Text generieren, sehen eine Effizienzsteigerung, aufgrund dessen sie zukünftig weniger Menschen beschäftigen müssen. Chatbots übernehmen die Kommunikation mit Kund*innen, was die Last der Kommunikation von dem Unternehmen mitunter auf die Kund*innen verschiebt. Eine besonders große Sorge ist, dass Organisationen oder Individuen, die Propaganda und Desinformation verbreiten, von Chatbots profitieren. Experimentiert wurde auch mit Chatbots in der Psychotherapie, mit kontrovers diskutierten Auswirkungen auf Behandlung und Vertrauen (Reardon 2023).

Beispiele für weitere nicht faktenbasierte Texte sind aber auch z.B. Briefe. Für Personen, die Barrieren zum Schreiben haben, sich aber schriftlich für etwas einsetzen möchten, kann der Nutzen beispielsweise sehr hoch sein. Auf den folgenden Prompt erstellte ChatGPT einen Brief, der (trotz fehlenden Genders) eine gute Wirkung bei einer Schulleitung erzielen könnte: ›Formuliere eine E-Mail, in der ich darum bitte, dass sich die Schule meines Sohnes um einen inklusiveren Unterricht kümmert.«

Betreff: Bitte um einen inklusiveren Unterricht für meinen Sohn

Sehr geehrte Damen und Herren,

ich wende mich heute an Sie, da ich mir für meinen Sohn einen inklusiveren Unterricht an Ihrer Schule wünsche. Ich glaube, dass jeder Schüler das Recht auf eine faire und gleichberechtigte Bildung hat, unabhängig von individuellen Unterschieden oder Behinderungen.

Leider habe ich in den letzten Wochen und Monaten festgestellt, dass mein Sohn nicht immer die Unterstützung und Anpassungen erhält, die er benötigt, um erfolgreich am Unterricht teilzunehmen. Ich denke, dass die Schule noch viel mehr tun könnte, um sicherzustellen, dass alle Schülerinnen und Schüler die gleichen Chancen haben, ihre Fähigkeiten und Talente zu entwickeln.

Ich würde es sehr begrüßen, wenn die Schule in dieser Angelegenheit proaktiv tätig werden könnte. Es gibt viele Ansätze und Maßnahmen, die dazu beitragen können, einen inklusiveren Unterricht zu gestalten, wie zum Beispiel:

- Ermöglichung von individuellen Lernplänen für Schülerinnen und Schüler mit speziellen Bedürfnissen
- Einsatz von geeigneten Lehrmaterialien und –methoden
- Schulung des Lehrpersonals im Umgang mit inklusiven Lernumgebungen
- Zusammenarbeit mit externen Fachleuten wie Therapeutinnen und Therapeuten

Ich bin davon überzeugt, dass eine stärkere Einbindung von inklusiven Ansätzen nicht nur meinem Sohn, sondern allen Schülerinnen und Schülern zugutekommen würde. Ich würde mich sehr freuen, wenn Sie diese Ideen in Be-

tracht ziehen und mit mir in Kontakt treten könnten, um weitere Möglichkeiten zu besprechen.

Vielen Dank im Voraus für Ihre Aufmerksamkeit und Ihre Bemühungen um einen inklusiveren Unterricht.

Mit freundlichen Grüßen,
[Name des Absenders]
(ChatGPT 3.5)

Aber auch Personen, die ChatGPT spielerisch herausfordern im Versuch, interessantere Antworten auf immer komplexer werdenden Prompts zu erhalten, haben positiv über dieses Experimentieren berichtet. Künstler*innen und Kreativarbeiter*innen haben mit der generativen KI beispielsweise auf diese Weise neue Formen der Kunst erschaffen. Studierende mit Deutsch als Zweitsprache können sich mit den schlichten, generischen Antworten von ChatGPT einen ersten Zugang zu einem Themengebiet erarbeiten, um dann in die Fachliteratur einzusteigen. Personen, die ›Angst vor dem weißen Blatt‹ spüren, lassen sich erste Sätze generieren, um dann eigenständig besser schreiben zu können. Lehrkräfte berichten in den sozialen Medien, dass sie Unterrichtspläne oder Aufgaben erstellen lassen. Bei diesen Beispielen kann ChatGPT unterstützen, wenngleich es problematisch sein kann, wenn dabei keine kritische Reflexion stattfindet. Weitere Lehrkräfte bieten Lernenden die Möglichkeit an, mit ChatGPT zu arbeiten. Wenn Schüler*innen angeleitet werden, ChatGPT eine fachliche oder analytische Frage zu stellen, die Antwort dann mithilfe der Fachliteratur zu überprüfen und die vom Chatbot generierte Antwort zu korrigieren, zu verbessern oder zu differenzieren, setzen sie sich gleichzeitig mit der Fachliteratur zum Thema auseinander und können eine kritische Reflexion des aktuellen Dominanzwissens und dieser Technologie unternehmen.

In diesem Abschnitt wurde gefragt: Wer hat einen Nutzen von der Anwendung von ChatGPT? Sofern die Leser*innen dieses Beitrags die Analyse selbst fortführen möchten, können sie weitere Personen identifizieren, die einen Nutzen davon haben könnten. Wenn der Design-Justice-Ansatz für die Analyse verwendet wird, gibt es zwei Ziele:

- zu reflektieren, wie diese Prozesse der Inklusion, der gesellschaftlichen Teilhabe, der Fürsorge, des Lernens, der Effizienzsteigerung, der Profit-

- maximierung, der Ausbeutung, der Reproduktion von Dominanzwissen o. Ä. von ChatGPT profitieren; und
- eine Bandbreite an Beispielen zu erarbeiten, wie die Nutzung so gestaltet werden kann, dass sich ein gerechtigkeitsorientierter Einsatz von generativer KI entfalten kann (siehe Abschnitt 4).

Wer trägt die Last?

Die Frage nach der Last kann auf individueller Ebene und auf gesellschaftlicher/ökologischer Ebene analysiert werden. Individuen wie die Annotator*innen (siehe oben) tragen eine Last, indem sie stunden-, tage- und monatelang zum Teil toxischen Inhalten ausgesetzt sind. Wie oft in der Geschichte der Automatisierung verlieren mit generativer KI aber auch Personen ihre bisherige Arbeit. In diesem Fall, weil bestimmte Aspekte der textgenerierenden Arbeit durch Roboter (Chatbots) ersetzt werden. Stattdessen entstehen andere Berufe, die den neuen technischen Gegebenheiten entsprechen (wenn, z.B. ein*e Autor*in zum/zur Lektor*in wird). Studien im englischsprachigen Raum zeigen des Weiteren, dass Personen mit Englisch als Zweitsprache eine Last tragen, sobald GPT-Detektoren – also Technologie, die ChatGPT-generierte Texte identifizieren soll – eingesetzt werden. Denn Texte von Fremdsprachenlernenden weisen offenbar Überschneidungen zu den Merkmalen von KI-generierten Texten auf. Diese Detektoren stufen die Texte von Fremdsprachenlernenden häufiger fälschlicherweise als KI-generiert ein (>false positives<) als die Texte ihrer Kommiliton*innen, deren Erstsprache Englisch ist (Liang et al. 2023).

Auf gesellschaftlicher oder ökologischer Ebene können verschiedene Lasten identifiziert werden. Zum Beispiel werden die finanziellen und Umweltkosten diskutiert. Der Geschäftsführer von *OpenAI* sagte, die Rechenkosten (die Kosten für die statistischen Berechnungen, die hinter jedem generierten Text von ChatGPT stehen) seien schmerzhaft teuer (>eye-watering<). Diese Rechenkosten führen auch durch ihren massiven CO₂-Ausstoß zu Kosten für die Umwelt, die die gegenwärtige ökologische Krise stärken. Die einkommensbasierte Schere in den Nutzungsmöglichkeiten der Technologie ist eine weitere Last: Laut Presseberichten hat *OpenAI* den oben erwähnten Investor*innen bis 2024 eine Milliarde Dollar Umsatz versprochen. Aktuell wird entsprechend eine kostenlose (ChatGPT 3.5) und eine kostenpflichtige (ChatGPT 4) Version angeboten mit je unterschiedlichen Funktionalitäten. ChatGPT 4 (nutzbar mit der Plus-Lizenz) weist weniger Fehler auf, kann mehr Daten schneller verar-

beiten und soll einen kreativeren Output erzeugen. Für noch höhere Kosten verspricht OpenAI die Prompts und Eingaben der Kund*innen nicht für das Training des Systems zu nutzen, das heißt, mit dieser ›ChatGPT Enterprise‹-Lizenz wird Datenschutz angeboten.

Dem Bereich der Lasten ist schließlich auch das grundlegende Menschenbild zuzuordnen und somit die Haltung gegenüber Menschen, der Welt und der Zukunft, das in generative KI-Systeme eingeschrieben wird und somit zirkuliert. Hierzu entsteht eine breite Diskussion, die hier nur angedeutet werden kann und in die es sich lohnt einzusteigen. Diese KI-Systeme haben keinen Sinn für Sinn und können Sprache nicht verstehen (Bender et al. 2021). Sie operieren mit statistischen Wahrscheinlichkeiten, um Wörter zusammenzunähen. So werden Texte kreiert, deren ›Bedeutung‹ irrelevant ist. Nichtsdestotrotz zeigte schon der erste KI-basierte Chatbot ›ELIZA‹ in den 1960er Jahren, dass einige Personen schwer zu überzeugen waren, dass sie nicht mit einem Menschen kommunizierten (Weizenbaum 1966). »A certain danger lurks there«, schrieb damals der Erfinder, Joseph Weizenbaum (1966: 43). Dies schreibt aber zugleich, so einige Beobachter*innen, eine Bedeutungslosigkeit in die Welt hinein, die aus dem Nihilismus stamme (Golumba 2022). Wenn Texte verbreitet werden, die Sinn vortäuschen, aber mit Sinnstiftung nichts zu tun haben, ist das eigentliche gesellschaftliche Produkt von ChatGPT nicht Text, sondern Verzweiflung (›despair‹). Für andere Beobachter*innen ist die reduktive und bürokratisierte Ordnung, die solchen KI-Systemen unterliegt, ein Ausdruck eines Faschismus, der mit der zunehmenden Verbreitung der Systeme einen subtilen, aber weitgreifenden Einfluss auf den gesellschaftlichen Zusammenhalt (bzw. deren Zusammenbruch) hat (McQuillan 2022). Dan McQuillan verortet die Geschichte der heutigen KI in der Eugenik des 19. Jahrhunderts und sagt, dass KI in den Händen von bürokratischen Institutionen ›zu einem Instrument [wird], das einige der schädlichen und grausamen Dinge verstärkt, die innerhalb dieser Institutionen bereits mit vulnerablen Menschen geschehen‹ (McQuillan et al. 2023; Übersetzung FM). Zwischenmenschliches Vertrauen wird fraglich, wenn wir uns bei jeder Online-Kommunikation fragen: Kommuniziere ich mit einem Menschen oder einem Bot? Ben Williamson (2023) nennt diese Systeme in der Bildung ›degenerative KI‹, um auf die problematischen Implikationen für Bildung und Gesellschaft hinzuweisen.

In diesem Abschnitt wurde gefragt: Wer trägt die Last von ChatGPT? Auch diese Analyse kann fortgeführt werden, um weitere Aspekte zu identifizieren, die auf eine Belastung hindeuten. Ziel ist,

- herauszuarbeiten, welche Individuen wo, wann und in welchen spezifischen Situationen in Kommunikation mit wem bei der Nutzung der Technologie eine Last tragen; und
- die breitere soziale, kulturelle, politische, ökonomische und ökologische Wirkkraft der generativen KI bzw. der jeweils zu analysierenden EdTech zu reflektieren.

4 Welche Alternativen können gestaltet werden?

Neben dem Analyserahmen ist Design Justice, wie oben beschrieben, auch eine Praxisgemeinschaft, mit dem Ziel, alternative Designs zu entwickeln. Diese können technische Alternativen sein, aber auch alternative Einsatzszenarien für den Unterricht oder den Nachmittagsbereich, für Projektstage oder für nonformale Lerngelegenheiten. Während einige Wissenschaftler*innen dafür plädieren, generative KI gar nicht zu nutzen (z.B. Golumba 2022; McQuillan 2022), bleibt die Frage in der Schule bzw. im Bildungsbereich: Wenn wir nicht auf neu entstehende Technologien eingehen, wer würde die (individuelle und gesellschaftliche) Wirkkraft für junge Menschen erfahrbar und reflektierbar machen?

Auf einer lokalen Ebene können kritisch-gestaltende Einsatzszenarien entwickelt werden. In der Schule z.B. gibt es noch kein eigenes Fach für die praktische, informationstechnische, politisch-ethische, medienpädagogische Auseinandersetzung mit digitalen Technologien. Dennoch bieten die Schulfächer Gestaltungsraum, um auf fachspezifische Dimensionen einzugehen. In Wirtschaft beispielsweise können Schüler*innen die Frage der menschlichen Arbeit bei der Erstellung und Begleitung von Systemen, die sich als ›automatisiert‹ vermarkten, besprechen. In Erdkunde können sie den ökologischen Impact der jeweiligen Technologie untersuchen. In Ethik können sie die Werte und Normen, die in KI-Systeme eingeschrieben werden, diskutieren. In Deutsch können KI-generierte Texte mit Literatur ins Verhältnis gesetzt werden. In Kunst können Schüler*innen die Künstler*innen besprechen, deren Arbeiten ohne ihre Erlaubnis benutzt wurden, um *DALL-E*, *Stable Diffusion* oder andere KI-Kunstgeneratoren zu trainieren. Auch der Aktivismus der Künstler*innen, die das Problem, wenn ihre Arbeit ohne ihre Erlaubnis benutzt wird, bekannter zu machen versuchen, kann Eingang in den Unterricht finden (Seemel 2022).

Auf rechtlicher Ebene wird mit dem neuen KI-Gesetz in der Europäischen Union eine Aufmerksamkeit für Unrechtssituationen und KI geschaffen (siehe auch den Beitrag ›Die KI-Verordnung. Der zukünftige Rechtsrahmen für Ed-Tech an Schulen‹ von Nina Galla in diesem Band). Weitere Anregungen zielen nicht nur auf eine Einschätzung und Regulierung des Risikos von KI für Individuen und Gruppen, sondern auch eine Regulierung des Risikos für die Gesellschaft. Nathalie Smuha zum Beispiel nimmt in dieser Hinsicht Inspiration vom Umweltrecht (Smuha 2021).

Auf gesellschaftlicher Ebene gibt es Bestrebungen, eine kollektive Stimme zu heben für eine Änderung der Entwicklungsprozesse von KI-Technologien (Bender et al. 2021). Emily Bender, Timnit Gebru und Kolleg*innen fragen u.a., warum die Bots so entwickelt werden, dass sie ›menschenähnliche‹ Sprache generieren. Wem nutzt das? Wie würde die Botkommunikation aussehen, wenn die Bothaftigkeit transparent gemacht werden würde? So würde den Nutzer*innen klarer vermittelt werden, dass es sich dabei um nicht intentionale Agenten handelt. Warum werden solche großen Datensätze für das Training genutzt? Die Größe bedeutet, dass sie nicht kuratiert sein können. Kuratierte Datensätze würden aber diverser sein können und stärker die unterschiedlichen gesellschaftlichen Perspektiven und Haltungen – auch minorisiertes Wissen – generieren können. Auch gibt es Vorschläge, dass, wenn KI-Systeme zur notwendigen Infrastruktur der gegenwärtigen Gesellschaft unerlässlich sind, diese Infrastruktur vergesellschaftet werden und in öffentlicher Hand liegen sollte (Viljoen 2020).

Bei rechtlichen und gesellschaftlichen Fragen liegt eine potenzielle Rolle von alltäglichen Nutzer*innen, die weder Jurist*innen noch Entwickler*innen sind, darin, Teil dieser kollektiven Stimme zu sein, wenn auch nur durch soziale Medien, Online-Unterschriften oder ähnliche Aktivitäten. In Schulen und weiteren Bildungsstätten kann die Sensibilisierung der jungen Menschen für die Möglichkeit ›anderer‹ Entwicklungsprozesse und Rechtslagen geschärft werden. Inspiration hierfür liegt bei den vielfältigen Überlegungen, Technologie ›anders‹ zu gestalten, z.B. in einer Art und Weise, die vielfältige Stimmen, Positionen und Wissensformen einbezieht und unterschiedliche, kritische und gestaltende Nutzungsweisen möglich macht (z.B. netzforma 2020; SUPERRR Lab n. d.; Swist et al. 2023).

5 Fazit

Dieser Beitrag zielte darauf ab, mit dem Design-Justice-Ansatz eine kritische und gestaltende Perspektive vorzustellen, die für die Analyse von Technologien eingesetzt werden kann. Anhand des Beispiels von generativer KI wurden erste Antworten auf eine Dreifachfrage (›Wer hat einen Nutzen, wer trägt die Last und wer nimmt teil?‹) gegeben. Diese Frage kann auf weitere Hard- und Software sowie Infrastrukturen und Einsatzszenarien in der Bildung übertragen werden. Sie kann alleine bearbeitet oder gemeinsamen im Kollegium bzw. mit Lernenden besprochen werden.

Der Ansatz weist auch – wie jeder Ansatz – eine Reihe von Grenzen auf. Die Frage nach Nutzen, Lasten und Teilnahme, wie in diesem Beitrag priorisiert, bildet nur einen Ausschnitt des gesamten Ansatzes. Es gibt komplexere Ansätze, EdTech kritisch zu reflektieren (vgl. z.B. Krutka et al. 2019; siehe auch die Beiträge ›Der EdTechReflektor – Bericht über die Entwicklung eines Instruments für den reflektierten Einsatz digitaler Bildungsmedien‹ von Sigrid Hartong sowie ›Medienreflexion mit Methode – Medienkonstellationsanalyse als Verbindung von Medienwissenschaft und Bildungspraxis‹ von Philipp Deny und Andreas Weich in diesem Band). Auch kann der Ansatz implizieren, eine ›bessere‹ Welt sei harmonisch, statt, wie es andere politische/theoretische Ansätze sehen würde, dass die Welt grundsätzlich agonistisch und konfliktreich ist und dass wir Zukünfte imaginieren sollten, die das Agonistische produktiv aufgreifen, statt es zu negieren. Ruha Benjamin hinterfragt auch die Nutzung von ›Design‹ und ›Justice‹ zusammen: Wollen wir nicht schlicht ›justice‹?, fragt sie (Benjamin 2019). Unterminiert die Nutzung des Designbegriffs nicht den egalitären Anspruch, weil es den Designer*innen (wenn auch unbeabsichtigt) einen höheren Status gibt? Wie jede (gesellschafts)kritische Perspektive muss auch diese sich immer wieder selbst hinterfragen.

In seiner aktuellen Form dient Design Justice zum einen als Analyseraster. Die Perspektive regt aber zum anderen an, selbst über alternative ›Designs‹ (Technologieentwicklung, Rechtslage, Einsatzszenarien usw.) nachzudenken. Auch wenn die Alternativen unrealistisch aussehen, verschiebt sich durch das Sprechen über Unmögliches doch oft das, was möglich erscheint (Seeman et al. 2022: 78). Der Design-Justice-Ansatz erinnert uns auch daran, nicht nur über die individuellen, sozialen und pädagogischen Auswirkungen nachzudenken, sondern auch die gesellschaftliche Wirkkraft kritisch zu beleuchten und bei der Gestaltung von Alternativen, die wir besser finden (wie prekär und kontingent dieses ›Besser‹ auch sein mag), mitzudenken.

Zentrale Take-Aways:

- Der Design-Justice-Ansatz wurde vom internationalen Design-Justice-Netzwerk entwickelt und hat zum Ziel, Designs (u.a. im Bildungsbereich) kritisch mit Blick auf Fragen der (Un-)Gerechtigkeit zu befragen und Ideen für zukünftig fairere Designs zu entwickeln.
- Ein beispielhaftes Analysemodell für eine Anwendung des Ansatzes im Bereich EdTech ist die Dreifachfrage: 1. Für wen, für was, in welchen Situationen, wann bringt die EdTech Vorteile? 2. Für wen, für was, in welchen Situationen, wann bringt sie Nachteile? 3. Wer war oder ist an der Entwicklung der EdTech beteiligt?
- Für die Beantwortung der Frage gibt es keine Vorgabe, sondern die Idee ist, frei in diverse Richtungen zu denken, aber bewusst sowohl individuelle als auch kollektiv-gesellschaftliche Positionen einzunehmen (z.B. welche Gruppen in der Gesellschaft und in der Welt werden benachteiligt?).
- Im Falle von ChatGPT können auf diese Weise unterschiedliche Dinge sichtbar(er) gemacht werden, darunter z.B. die Bedeutung der Technologie für unterschiedliche Personengruppen, die Hintergründe der KI-Antwortproduktion (Wer sind die Entwickler*innen? Wer sind die Annotatoren?) oder auch Umwelt- und politische Dimensionen.
- Gemeinsam mit oder im Anschluss an die Reflexion können mögliche Ansätze einer alternativen EdTech-Produktion oder -Nutzung diskutiert werden, die gerechter erscheinen. Auch hierbei kann in unterschiedliche Richtungen überlegt werden (z.B. weitere Unterstützungsmaßnahmen bei der Nutzung von EdTech oder Produktion alternativer Aufgaben für Plattformen).

Acknowledgements

Teile dieses Beitrags sind zum ersten Mal in früheren Publikationen erschienen (Macgilchrist 2023, 2024). Das hier vorgestellte Verständnis des Design-Justice-Ansatzes ist im Rahmen des vom BMBF geförderten Verbundforschungsprojekts FaBuLoUS (FabLabs als Bildungs- und Lernorte zur Unterstützung von Schulen; Förderkennzeichen: 01JD1902C) entstanden. Die Verantwortung für die Inhalte des Textes bleibt bei der Autorin.

Autorinneninformation

Felicitas Macgilchrist ist Professorin für Digitale Bildung in der Schule an der Carl von Ossietzky Universität Oldenburg.

Literatur

- Beauftragte der Bundesregierung für Migration (2015): »Schulbuchstudie: Migration und Integration«, verfügbar unter: https://www.bundesregierung.de/Content/Infomaterial/BPA/IB/Schulbuchstudie_Migration_und_Integration_09_03_2015.pdf%3F__blob%3DpublicationFile%26v%3D3.
- Bender, Emily M. et al. (2021): »On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?«, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, S. 610–623.
- Benjamin, Ruha (2019): *Race After Technology*, Cambridge: Polity.
- Bohannon, Molly (2023): »Lawyer Used ChatGPT In Court – And Cited Fake Cases«, in: Forbes vom 08.06.2023, online verfügbar unter: <https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions/>
- Collins, Patricia Hill (1990): *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, Boston: Unwin Hyman.
- CompTIA (Computing Technology Industry Association). (2021). *Cyberstates 2021*. CompTIA, online verfügbar unter: https://www.cyberstates.org/pdf/CompTIA_Cyberstates_2021.pdf
- Costanza-Chock, Sasha (2020): *Design Justice: Community-Led Practices to Build the World We Need*, Cambridge: MA: MIT Press.
- Crenshaw, Kimberle (1991): »Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color«, in: *Stanford Law Review*, 43(6), S. 1241–1299.
- Dzeiza, Josh (2023): »AI is a lot of work«, in: *Intelligencer* vom 20. Juni 2023, online verfügbar unter: <https://nymag.com/intelligencer/article/ai-artificial-intelligence-humans-technology-business-factory.html>
- Escobar, Arturo (2018): *Designs for the Pluriverse*, Durham: Duke University Press.
- Farhi, Paul (2023): »A news site used AI to write articles: It was a journalistic disaster«, in: Washington Post vom 17. Januar 2023, online verfügbar unter: <https://www.washingtonpost.com/technology/2023/01/17/ai-news-articles-disaster/>

- ter <https://www.washingtonpost.com/media/2023/01/17/cnet-ai-articles-journalism-corrections/>
- Golumba, David (2022): ChatGPT Should Not Exist, online verfügbar unter: <https://davidgolumba.medium.com/chatgpt-should-not-exist-aabo867abace>
- Holmes, Kat (2020): *Mismatch: How Inclusion Shapes Design*, Cambridge, MA: The MIT Press.
- Krutka, Daniel G.; Heath, Marie K.; Willet, K. Bret Staudt (2019): Foregrounding Technoethics: Toward Critical Perspectives in Technology and Teacher Education, in: *Journal of Technology and Teacher Education* 4, S. 555–574.
- Liang, Weixin et al. (2023): »GPT detectors are biased against non-native English writers«, in: *Patterns* 4, online verfügbar unter: <https://www.sciencedirect.com/science/article/pii/S2666389923001307>.
- Macgilchrist, Felicitas (2024). »Design justice and educational technology: Designing in the fissures«, in Ben Williamson, Janja Komljenovic, Kalervo Gulson (Hg.), *World Yearbook of Education 2024. Digitalization of Education in the Era of Algorithms, Automation, and Artificial Intelligence*, London: Routledge, S. 294–310.
- Macgilchrist, Felicitas et al. (2023): »»Google_Suche«: Suche als soziale Praxis in Unterricht und Schule«, in Matthias Proske et al. (Hg.) *Schule und Unterricht im digitalen Wandel. Ansätze und Erträge rekonstruktiver Forschung*, S. 67–89.
- Macgilchrist, Felicitas et al. (2023): »Designing postdigital futures: Which designs? Whose futures?«, in: *Postdigital Science and Education*, online verfügbar unter: <https://doi.org/10.1007/s42438-022-00389-y>.
- Macgilchrist, Felicitas. (2023). *KI und Schule: Sichtweisen, Anwendungen und Gestaltungsmöglichkeiten. SCHÜLER-Heft* (Friedrich-Verlag), Thema »Hybride Lebenswelten«, S. 82–84.
- Markov, Todor et al. (2022). A Holistic Approach to Undesired Content Detection in the Real World, Oral presentation at AAAI-23, online verfügbar unter: <https://doi.org/10.48550/arXiv.2208.03274>.
- McQuillan, Dan (2022): *Resisting AI: An Anti-fascist Approach to Artificial Intelligence*, Bristol: Bristol University Press.
- McQuillan, Dan; Jarke, Julian; Cerratto Pargman, Teresa (2023): »We are at an extreme point where we have to go all in on what we really believe education should be about«, in: *Postdigital Science and Education 2023*, S. unbekannt.

- netzforma (Hg.) (2020): Wenn KI, dann feministisch: Impulse aus Wissenschaft und Aktivismus, online verfügbar unter: <https://netzforma.org/publikation-wenn-ki-dann-feministisch-impulse-aus-wissenschaft-und-aktivismus>.
- Noble, Safiya Umoja (2018): *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press.
- Perrigo, Billy (2023): »OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic«, in: Time Magazine vom 18. Januar 2023, online verfügbar unter: <https://time.com/6247678/openai-chatgpt-kenya-workers/>
- Reardon, Sara (2023): »AI Chatbots Could Help Provide Therapy, but Caution Is Needed«, in: Scientific American vom 14. Juni 2023, online verfügbar unter: <https://www.scientificamerican.com/article/ai-chatbots-could-help-provide-therapy-but-caution-is-needed/>
- Scheerer, Jamie; Haruna, Hadija (2013): Über Schwarze Menschen in Deutschland berichten. Initiative Schwarze Menschen in Deutschland (ISD), online verfügbar unter: <https://isdonline.de/uber-schwarze-menschen-in-deutschland-berichten/>
- Seemann, Michael et al. (2022): Konzeptstudie. Werte und Strukturen der Nationalen Bildungsplattform, online verfügbar unter: Wikimedia. <https://www.wikimedia.de/wp-content/uploads/2022/11/Konzeptstudie-Werte-und-Strukturen-der-Nationalen-Bildungsplattform.pdf>
- Seemel, Gween (2022): Welcome to AI's Disneyland: Eric Bourdages' Fight for Artists, online verfügbar unter: <https://gweenseemel.com/blog/2022/12/19-ai-eric-bourdages/> vom 19. Dezember 2022.
- Sheya, Sarah (2020). JusticexDesign: Developing a Sensitivity to Designed Injustices, online verfügbar unter: <https://www.agencybydesign.org/node/466>
- Smuha, Nathalie A. (2021): »Beyond the individual: governing AI's societal harm«, in: Internet Policy Review 10(3), online verfügbar unter: <https://doi.org/10.14763/2021.3.1574>.
- SUPERRR Lab (o.A.). Feminist Tech Principles, online verfügbar unter: <https://superrr.net/feministtech/principles/>.
- Swist, Teresa; Humphry, Justine; Gulson, Kalervo N. (2023): »Pedagogic encounters with algorithmic system controversies: a toolkit for democratising technology«, in: Learning, Media and Technology 48(2), S. 226–239.

- Viljoen, Salomé (2020): »Data as Property?«, in: Phenomenal World vom 16.10.2020, online verfügbar unter: <https://phenomenalworld.org/analyses/data-as-property>.
- Weizenbaum, Joseph (1966): »ELIZA – A computer program for the study of natural language communication between man and machine«, in: Communications of the ACM 9(1), S. 36–45.
- Williamson, Ben (2023): Degenerative AI in Education, in: code acts in education vom 30.06.2023, online verfügbar unter: <https://codeactsineducation.wordpress.com/2023/06/30/degenerative-ai-in-education/>.

