

REIHE WISSENSCHAFTSKOMMUNIKATION UND LINGUISTIK

Forschungseinblicke • Projektberichte • Aufsätze • Studierendenarbeiten

Nr. 3

Klara Laue

Anthropomorphisierung in der Berichterstattung über KI

Inwiefern beeinflusst die Anthropomorphisierung von KI-Systemen in der Berichterstattung die öffentliche Wahrnehmung von Künstlicher Intelligenz?

Wissenschaftskommunikation
mit Schwerpunkt Linguistik
Institut für Technikzukünfte



Impressum

Redaktion

Monika Hanauska, Annette Leßmöllmann, Johanne Mayer

Herausgeberin

Prof. Dr. Annette Leßmöllmann

E-Mail: annette.lessmoellmann@kit.edu

Karlsruher Institut für Technologie (KIT)

Institut für Technikzukünfte

Department für Wissenschaftskommunikation

Lehrstuhl mit Schwerpunkt Linguistik

Adenauerring 12

D-76131 Karlsruhe

© 2025 Klara Laue

ISSN: 2942-1438

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt allein bei den Autorinnen und Autoren.



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Nicht kommerziell – Keine Bearbeitungen 4.0 International Lizenz (CC-BY-NC-ND 4.0 DE):
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Vorwort

Mit der Reihe Wissenschaftskommunikation und Linguistik verfolgen wir das Ziel, Einblicke in den ›work of progress‹ an unserem Lehrstuhl mit dem Schwerpunkt Linguistik zu liefern.

Naturgemäß zeichnen sich die Aufgaben innerhalb eines Lehrstuhls durch eine Zweiteilung aus: Einerseits betreiben die wissenschaftlichen Mitarbeitenden Forschung, deren Zwischenergebnisse oder laufende Prozesse selten zeitnah in einer Fachzeitschrift veröffentlicht werden. Andererseits nimmt die Lehre einen großen Raum im akademischen Alltag ein, deren Resultate in Form von exzellenten Hausarbeiten meistens nicht über den Teilnehmendenkreis hinaus zugänglich sind.

In beide Bereiche gewährt diese Schriftenreihe Einblicke, indem sowohl herausragende Abschlussarbeiten von Studierende als auch Forschungs- und Projektberichte, Vortragsbeiträge sowie unveröffentlichte Aufsätze publiziert werden.

Der Lehrstuhl vertritt einen weiten Linguistikbegriff, weshalb sich in dieser Publikation Beiträge aus der Sprach-, Kommunikations- und Kulturwissenschaft sowie Schnittstellenarbeiten zur Multimodal- und Rezeptionsforschung finden.

Welche Texte veröffentlicht werden, wird durch ein Redaktionsteam entschieden, das seine Wahl auf Basis eines Kriterienkatalogs trifft. Die Arbeit sollte thematisch aktuell, gesellschaftsrelevant sowie zukunftsfähig sein und gleichzeitig eine Forschungslücke bedienen. Studentische Arbeiten werden auf Empfehlung eines/r Erstgutachters/in berücksichtigt. Für diese Arbeiten gelten ebenfalls die genannten Kriterien.

Hintergründe zu der Person und der veröffentlichten Arbeit

Name:	Klara Laue
Höchster Abschluss:	Bachelor of Arts
Studiengang:	Wissenschaft - Medien – Kommunikation (M.A.)
Textart:	Hausarbeit
Seminar:	Wissenschaftskommunikation und Künstliche Intelligenz
Beschreibung:	Das Seminar befasste sich mit der Forschung über den Einsatz generativer KI in Wissenschaft und Wissenschaftskommunikation sowie die Auswirkungen der Nutzungen auf die Praktiken und das Wissensverständnis in Wissensberufen.

Inhalt

1 Einführung.....	1
2 Theoretischer Hintergrund.....	2
2.1 Begriffsbestimmungen	2
2.2 Theorie der Anthropomorphisierung	2
3 Stand der Forschung: Anthropomorphisierung von KI.....	3
3.1 Beweggründe für Anthropomorphisierung von KI	3
3.2 Herkunft der Anthropomorphisierung von KI.....	4
4 Diskussion der öffentlichen Wahrnehmung von Anthropomorphisierung in der KI- Berichterstattung	6
4.1 Darstellung von KI in den Medien	6
4.2 Positive Effekte der Anthropomorphisierung von KI	8
4.3 Negative Effekte der Anthropomorphisierung von KI.....	9
4.3.1 Unrealistische Erwartungen.....	9
4.3.2 Ängste und Bedenken	11
4.3.3 Reproduktion von Stereotypen.....	11
4.3.4 Rechtliche Fragen und Urheberrechtsverletzungen	12
4.3.5 Auswirkungen auf die KI-Forschung.....	13
5 Implikationen für die Berichterstattung über KI.....	14
6 Fazit.....	16
7 KI-Nachweis.....	17
8 Literaturverzeichnis.....	17

Abbildungsverzeichnis

Abb. 1: Symbolbilder: „Klassische“ Visualisierungen von KI (Quelle: Getty Images).....	5
Abb. 2: Symbolbilder: Neue Visualisierungen von KI (Quelle: betterimagesofai.org)	15

1 Einführung

KI als Tutor, Therapeut, Assistent? Beim Sprechen über Künstliche Intelligenz (KI) werden technische Systeme häufig vermenschlicht. In der Fachsprache wird dieses Phänomen als *Anthropomorphisierung* oder Vermenschlichung bezeichnet. Dabei schreiben wir nicht-menschlichen Entitäten, etwa Tieren, Maschinen oder dem Wetter menschliche Eigenschaften zu. Vor diesem Hintergrund widmet sich die Arbeit der Frage, inwiefern KI-Systeme in der Berichterstattung über KI anthropomorphisiert werden und welche potenzielle Wirkung die Anthropomorphisierung von KI auf die öffentliche Wahrnehmung hat. Im Rahmen einer narrativen Literaturübersicht werden kommunikationswissenschaftliche, sozialwissenschaftliche und medienwissenschaftliche Texte kritisch analysiert und verschiedene Forschungsperspektiven beleuchtet.

Zunächst werden im Zuge der theoretischen Verortung zentrale Begriffe erläutert (2.1). Auf Basis der Theorie des Verhaltenswissenschaftlers Nicholas Epley wird anschließend dargelegt, wieso Menschen zu einer Anthropomorphisierung nicht-menschlicher Entitäten neigen (2.2). Kapitel drei widmet sich dem Stand der Forschung zu Anthropomorphisierung von KI. Es wird beleuchtet, weshalb wir Metaphern und konkret Vermenschlichung nutzen, um das Konzept KI besser begreifen zu können (3.1). Ein Blick in die Geschichte zeigt, dass popkulturelle Narrative menschliche Darstellungen von Techniksystemen geprägt haben und wir schon immer zu Anthropomorphisierung nicht-menschlicher Entitäten neigen (3.2). Daraufhin wird erläutert, wie KI in den letzten Jahren medial thematisiert wurde (4.1). Dabei werden bestimmte Rahmungen dargelegt, die im Kontext von KI häufig genutzt werden. Anschließend widmet sich Kapitel 4.2 positiven Effekten der Anthropomorphisierung von KI. Insbesondere in der Anwendung von KI-Systemen können menschliche Eigenschaften die Benutzerfreundlichkeit sowie das Verständnis von Technologie erhöhen. Folgend werden negative Effekte der Anthropomorphisierung von KI beleuchtet (4.3). Das Unterkapitel ist in fünf weitere Unterkapitel gegliedert und behandelt unrealistische Erwartungen an KI (4.3.1), Ängste und Bedenken (4.3.2), Reproduktion von Stereotypen (4.3.3), rechtliche Fragen und Urheberrechtsverletzungen (4.3.4) sowie Auswirkungen auf die KI-Forschung (4.3.5). Abschließend werden konkrete Implikationen für die Berichterstattung über KI dargelegt (5) und zentrale Aspekte der Arbeit resümiert (6).

2 Theoretischer Hintergrund

2.1 Begriffsbestimmungen

Um zu untersuchen, inwiefern die Anthropomorphisierung von KI-Systemen in der Berichterstattung die öffentliche Wahrnehmung von Künstlicher Intelligenz beeinflusst, müssen zunächst die zentralen Konzepte *Künstliche Intelligenz* und *Anthropomorphisierung* definiert werden. Künstliche Intelligenz (KI) bezeichnet die Fähigkeit eines technischen Systems, menschenähnliches Verhalten zu imitieren und Aufgaben auf intelligente Weise auszuführen (Li & Suh, 2022, S. 2246). Diese Intelligenz zeigt sich in kognitiven, emotionalen und sozialen Aspekten und basiert auf selbstlernenden Prozessen sowie der kontinuierlichen Erweiterung des Wissens. Die Definition greift auf McCarthy et al. (1955) zurück, wonach KI das Konzept beschreibt, eine Maschine so zu gestalten, dass ihr Verhalten als intelligent bezeichnet werden würde, wenn sich ein Mensch so verhalten würde. Da es sich um eine sozial- und kommunikationswissenschaftliche Arbeit handelt, liegt der Fokus auf der Wahrnehmung von KI und Interaktion mit KI durch Menschen. In der Informatik oder anderen technischen Disziplinen stehen in der Regel Problemlösungen durch Algorithmen und maschinelles Lernen im Vordergrund (vgl. Bundesamt für Sicherheit in der Informationstechnik (BSI), 2025).

Der Verhaltenswissenschaftler Nicholas Epley definiert Anthropomorphisierung als die Wahrnehmung menschlicher Eigenschaften in nichtmenschlichen Akteuren (2018, S. 591). Eine Erweiterung dieser Definition liefern Salles et al., die Anthropomorphisierung als Zuschreibung eindeutig menschlicher Gefühle, mentaler Zustände und Verhaltensweisen auf unbelebte Objekte, Tiere sowie natürliche Phänomene und übernatürliche Entitäten beschreiben (2020, S. 89). Tobias Kopp ergänzt explizit, dass sich diese Zuschreibung auch auf offensichtlich unbelebte Maschinen beziehen kann (2020, S. 42). In dieser Arbeit werden die Begriffe Anthropomorphisierung und Vermenschlichung synonym verwendet. Folgend wird die Theorie der Anthropomorphisierung nach Epley erläutert.

2.2 Theorie der Anthropomorphisierung

Nach Nicholas Epley ist die Anthropomorphisierung nicht-menschlicher Entitäten ein kognitiver Prozess, der auf induktiven Schlussfolgerungen unseres Wissens über Menschen basiert (2018, S. 591). Das bedeutet, dass wir aus unseren Erfahrungen mit menschlichem Verhalten Muster ableiten und diese unbewusst auf nicht-menschliche Objekte übertragen. Der Verhaltenswissenschaftler identifiziert drei zentrale Faktoren, die Anthropomorphisierung bedingen. Zum einen nennt Epley das Bedürfnis nach einer kompetenten Interaktion mit unserer Umwelt (2018, S. 592). So können wir unser Umfeld besser verstehen und kontrollieren. Ein weiterer zentraler Einflussfaktor ist das Verlangen nach sozialen Bindungen, das uns unter Abwesenheit echter Menschen menschliche Eigenschaften auf nicht-menschliche Entitäten projizieren lässt, um emotionale Bindungen zu schaffen (ebd., S. 592). Die

Motivation für Anthropomorphisierung basiert demnach auf ähnlichen Mechanismen wie die Fähigkeit, anderen Menschen mentale Zustände zuzuschreiben, etwa, um ihr Verhalten vorherzusagen oder zu erklären. Dieses Phänomen wird als *Theory of Mind* bezeichnet. Das Konzept der Anthropomorphisierung geht darüber hinaus, weil auch Emotionen, Verhaltensweisen und physische Merkmale auf nicht-menschliche Entitäten projiziert werden (ebd., S. 592f.).

Der dritte Einflussfaktor ist die wahrgenommene Ähnlichkeit oberflächlicher Eigenschaften mit Menschen. Während das Bedürfnis nach kompetenter Interaktion und das Verlangen nach sozialen Bindungen vom Menschen ausgehen, handelt es sich bei der wahrgenommenen Ähnlichkeit um einen Faktor, der primär vom anthropomorphisierten Objekt ausgeht. Menschliche Gesichtszüge, Bewegungen und eine menschliche Stimme können den Grad der Vermenschlichung maßgeblich beeinflussen, so Epley (ebd., S. 593). Die Tendenz zur Anthropomorphisierung hängt dabei stark vom eigenen Bedürfnis nach sozialen Beziehungen und davon ab, ob und in welchem Maß dieses erfüllt ist. Vor diesem Hintergrund wird im Folgenden erläutert, inwiefern wir KI-Systemen menschliche Eigenschaften zuweisen.

3 Stand der Forschung: Anthropomorphisierung von KI

3.1 Beweggründe für Anthropomorphisierung von KI

Das System KI ist hochkomplex und nur sehr wenige Menschen besitzen ein fundiertes Verständnis darüber. Expert*innen verschiedener Disziplinen haben zahlreiche Metaphern entwickelt, die uns helfen sollen, KI zu verstehen. Melanie Mitchell trägt diese in der Publikation „The metaphors of artificial intelligence“ in *Science* aus dem Jahr 2024 zusammen. KI-Systeme werden als Schauspieler beschrieben, die sehr viele Rollen imitieren können (Shanahan et al., 2023) oder als kulturelle Technologien, vergleichbar mit Enzyklopädien oder Bibliotheken, die Menschen effektiv Zugang zu Informationen schaffen (Yiu et al., 2024). Wir können uns KI-Technologien als Spiegel menschlicher Intelligenz vorstellen, die komplexe Reflexionen unserer eigenen Gedanken kreieren (Vallor, 2024) oder als verschwommene JPEGs des Internets, die komprimierte Versionen ihrer Trainingsdaten sind (Chiang, 2023). Die Linguistin Emily Bender beschreibt große Sprachmodelle wie *ChatGPT* als stochastische Papageien, die zwar beeindruckend klingende Texte erzeugen können, aber kein echtes Sprachverständnis besitzen. Sie fügen Texte anhand von Wahrscheinlichkeiten aus Trainingsdaten zusammen, ohne zu wissen, was die Worte bedeuten (Bender et al., 2021, S. 616f.).

Diese Vielfalt an Metaphern, welche die Funktionsweise von KI-Systemen erläutern sollen, verdeutlicht, dass es uns schwerfällt, über KI zu sprechen. Selbst Expert*innen sind sich uneinig darüber, wie KI am besten konzeptualisiert werden sollte, sodass wir sie verstehen. Forschende argumentieren, dass Metaphern in Anbetracht des mangelnden Verständnisses von KI derzeit unser bestes Mittel sind, um KI zu verstehen und darüber zu sprechen (Brokensha & Conrady, 2021; Mitchell, 2024). Wenn wir über

KI sprechen, dann begegnen uns Metaphern ohnehin häufig: Wir sprechen von *neuronalen* Netzen, von maschinellern *Lernen*, von künstlicher *Intelligenz*. Diese zentralen Begriffe wurden von menschlichen Eigenschaften abgeleitet, obwohl sie diese nur auf sehr abstrakte Weise nachbilden. Bereits in den 1970er Jahren kritisierte der KI-Forscher Drew McDermott, dass diese Vermenschlichung lediglich als *wishful mnemonics* dienen. Diese spiegeln allerdings nicht die Realität wider, sondern bringen eher den Wunsch zum Ausdruck, dass KI eines Tages tatsächlich wie ein Mensch funktionieren könnte (McDermott, 1976).

Wie bereits erläutert, tendieren Menschen dazu, nicht-menschlichen Entitäten, wie Tieren, dem Wetter oder Unternehmen, menschliche Eigenschaften zuzuschreiben. Melanie Mitchell legt dar, dass diese Tendenz in Bezug auf große Sprachmodelle wie ChatGPT besonders ausgeprägt ist, da diese in flüssiger Sprache sowie in der ersten Person Singular kommunizieren und ihre Gefühle zum Ausdruck bringen (2024). Die Beobachtung, dass es eine starke Tendenz gibt, KI zu anthropomorphisieren, ist in der Forschung weit verbreitet (Kopp, 2020, S. 41; Salles et al., 2020, S. 88; Brokensha & Conradie, 2021). Große Sprachmodelle können menschliche Kommunikation perfekt nachahmen, was es uns erschwert, sie nicht als bewusste Wesen wahrzunehmen. Wir werden daher auch immer wieder dazu verleitet, ihnen Gefühle oder Intentionen zuzuschreiben, obwohl sie in Wahrheit nur Daten berechnen und wiedergeben (Mitchell, 2024). Die Tendenz zur Anthropomorphisierung hängt folglich nicht unbedingt von den tatsächlichen Merkmalen oder Verhaltensweisen des Objekts ab. Es handelt sich weniger um eine Beschreibung objektiver Eigenschaften, sondern viel mehr um eine menschen-ähnliche Interpretation bestehender physischer Eigenschaften und Verhaltensweisen (Salles et al., 2020, S. 89).

3.2 Herkunft der Anthropomorphisierung von KI

Unsere Vorstellung von Künstlicher Intelligenz wird maßgeblich durch verbreitete Narrative geprägt, die KI-Systemen menschliche Eigenschaften zuschreiben. Ein Bericht der Royal Society aus dem Jahr 2018 legt dar, dass diese Erzählmuster in der Regel bestimmte charakteristische Merkmale aufweisen. Häufig zu beobachten sind Verkörperungen und damit bis zu einem gewissen Grad Vermenschlichung und eine Tendenz zu utopischen oder dystopischen Extremen (The Royal Society, 2018, S. 4). Die Geschichte von Narrativen über intelligente Maschinen reicht bis in die Mythologie der Antike zurück. Dort erschuf Hephaistos, der Gott der Schmiedekunst, Maschinen aus Gold, denen menschliche Eigenschaften wie Intelligenz und Treue zugewiesen wurden.

Im Mittelalter wurden Erzählungen intelligenter Maschinen vom Narrativ einer menschlichen *Hybris* (deutsch: Maßlosigkeit, Hochmut) begleitet. Demnach sollte die Macht, eine intelligente Kreatur zu schaffen, nicht den Sterblichen obliegen (The Royal Society, 2018, S. 7). Im Jahr 1738 entwarf der französische Ingenieur Jacques de Vaucanson seinen „Flötenspieler“. Dabei handelt es sich um einen erstaunlich realistisch wirkenden Automaten, der rein mechanisch betrieben wurde, aber Eigenschaften

eines modernen Roboters aufwies. Die Vorstellung, dass menschenähnliche Maschinen bald existieren könnten, wurde so realistischer. Damit einher ging die Angst vor Maschinen, die Macht über ihre Erschaffer gewinnen. Diese wurde in verschiedenen Erzählungen aufgegriffen. Nach der Einführung des Begriffs *Künstliche Intelligenz* im Jahr 1955 wurden KI-Narrative immer populärer und behandelten eine Vielzahl an Themen wie Kontrolle, Unsterblichkeit und Bewusstsein (The Royal Society, 2018, S. 8).

Der Blick in die Geschichte fiktionaler Narrative über intelligente Maschinen zeigt, dass diese von einer starken Tendenz beherrscht werden, Maschinen menschliche Eigenschaften zuzuweisen. Die Gründe hierfür sind vielfältig. Zum einen gilt der Mensch als Maßstab für Intelligenz, da er andere Lebewesen nach der westlichen Kultur darin übertrifft. Intelligente Wesen – etwa Götter, Engel oder Dämonen – werden in der Regel menschenähnlich dargestellt. Darüber hinaus ist das Narrativ der Maschinen, die für Menschen arbeiten, weit verbreitet. Oft werden diese in metallischer Form visualisiert und ersetzen menschliche Diener*innen. Außerdem ist es für viele Geschichten notwendig, Maschinen einen Körper zu verleihen. In menschlicher Form ist es in der Regel einfacher, Maschinen sinnvoll in eine Erzählung einzubinden und es erlaubt den Rezipient*innen, sich mit den Figuren zu identifizieren. In Filmen geht die Anthropomorphisierung von technischen Systemen oder Robotern häufig mit einer Übertonung sekundärer Geschlechtsmerkmale einher (The Royal Society, 2018, S. 8).

Neben den Einflüssen aus der Popkultur wird unser Bild von KI auch durch die mediale Berichterstattung geprägt. In einem Bericht des Projektes „Better Images of AI“ aus dem Jahr 2023 identifizieren die Autorinnen zentrale Elemente, die Visualisierungen von KI häufig enthalten. Anthropomorphisierung ist häufig ein zentraler Aspekt von KI-Bildern. Außerdem sehen wir im Kontext von KI häufig Roboter, die wie Menschen aussehen, blaue, mechanische Gehirne oder Variationen des Gemäldes „Die Erschaffung des Adam“ von Michelangelo (Dihal & Duarte, 2023, S. 5). Häufig reproduzieren KI-Visualisierungen geschlechterspezifische Stereotype. Während Männer häufig als Entwickler porträtiert werden, werden KI-Assistenten häufig weibliche Attribute zugewiesen (Costa & Ribas, 2019; Cave et al., 2023).

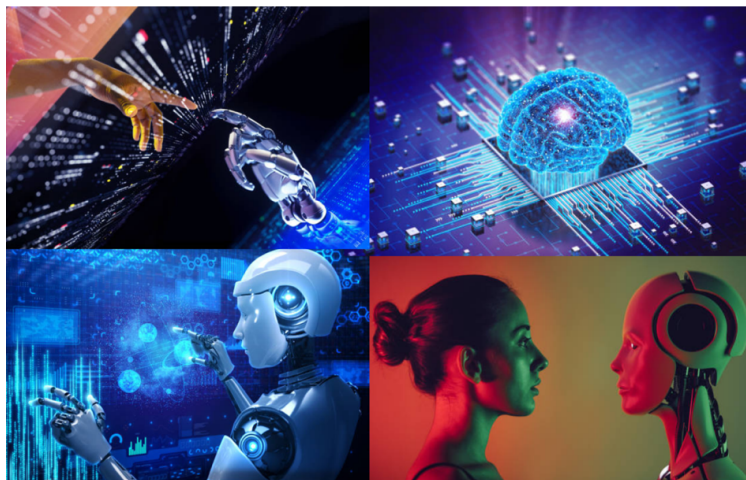


Abb. 1: Symbolbilder: „Klassische“ Visualisierungen von KI (Quelle: Getty Images)

4 Diskussion der öffentlichen Wahrnehmung von Anthropomorphisierung in der KI-Berichterstattung

Die mediale Berichterstattung spielt eine wichtige Rolle hinsichtlich der Meinungsbildung über KI. Häufig stellen Massenmedien die einzige Quelle für Informationen über Wissenschaft und Technologie dar (Fucker, 2022, S. 81). Journalist*innen und Wissenschaftskommunikator*innen tragen folglich eine große Verantwortung, da sie die öffentliche Wahrnehmung von KI maßgeblich prägen. Das ist insbesondere in Anbetracht der Tatsache relevant, dass die KI-Entwicklung in Zukunft gesellschaftliche Entscheidungen politischer, moralischer und wirtschaftlicher Natur fordern wird. Vor diesem Hintergrund widmet sich dieses Kapitel der Frage, wie KI in der medialen Berichterstattung dargestellt wird. Anschließend wird diskutiert, welche potenziellen positiven und negativen Effekte die Vermenschlichung von KI auf die öffentliche Wahrnehmung haben.

4.1 Darstellung von KI in den Medien

Die Berichterstattung über KI bedient sich häufig an einem bestimmten *Framing*. Darunter versteht man eine Einrahmung von Informationen in der medialen Berichterstattung, durch die ausgewählte Aspekte hervorgehoben und andere ausgeblendet werden, um eine bestimmte Deutung oder Wirkung beim Publikum zu erzeugen (Nelson et al., 1997, S. 221f.; Fucker, 2022, S. 83). In einer Studie aus dem Jahr 2019 untersuchten die Autor*innen Chuan et al. im Rahmen einer qualitativen Inhaltsanalyse das Framing von KI in ausgewählten amerikanischen Zeitungen. Die Ergebnisse zeigen, dass die Vorteile von KI häufiger thematisiert, aber die Risiken spezifischer diskutiert wurden. Über ethische Bedenken wurde berichtet, allerdings häufig nur oberflächlich. Ethische Probleme wurden selten spezifisch analysiert. Die Forscher*innen resümieren, dass eine detailliertere und ausgewogenere Berichterstattung über die Risiken und Vorteile von KI nötig wäre, um eine kritische Bewertung zu ermöglichen (Chuan et al., 2019).

Eine andere Studie aus dem Jahr 2020, die untersuchte, inwiefern ethische Fragen rund um KI in den Medien thematisiert werden, kommt zu ähnlichen Ergebnissen. Die Berichterstattung über ethische Fragen erfolgte meist realistisch und praktisch, häufig aber oberflächlich und wenig spezifisch (Ouchchy et al., 2020). Häufig beziehen sich Artikel über KI auf die Chancen und Risiken der Technologie. Eine Studie aus dem Jahr 2017 legt dar, dass die Berichterstattung über KI seit den späten 2000er-Jahren deutlich zugenommen hat und die Darstellung der Technologie seitdem weitgehend optimistisch war. Die Möglichkeit, KI-Systeme im Gesundheitswesen und in der Bildung einzusetzen, machte Hoffnung. Aber auch Bedenken im Umgang mit KI wurden häufiger thematisiert, darunter die Sorge vor einem Kontrollverlust der KI-Entwicklung, ethische Bedenken im Umgang mit KI und der negative Einfluss auf die Arbeitswelt (Fast & Horvitz, 2017). Zudem spielen regulatorische Fragen und Verantwortlichkeiten

eine zentrale Rolle in der medialen Darstellung von KI (Chuan et al., 2019; Ouchchy et al., 2020). Insgesamt zeigt sich, dass die Berichterstattung über KI anfangs eher positiv oder neutral war, mit zunehmender Berichterstattung jedoch negativer wurde (Ouchchy et al., 2020). Ein Experiment aus dem Jahr 2022 zeigte, dass die Darstellung von Chancen und Risiken künstlicher Intelligenz emotionale Reaktionen bei den Rezipient*innen hervorrief (Fucker, 2022).

Im Rahmen einer quantitativ-qualitativen Inhaltsanalyse untersuchten Bunz und Braghieri das Framing von KI in der englischsprachigen Berichterstattung insbesondere im Hinblick auf eine mögliche *Agency* im Sinne von Handlungsfähigkeit. Die Forscher*innen stellten fest, dass einige Artikel KI-Systeme als innovative Technologien framt, die menschliche Fähigkeiten in bestimmten Bereichen übertreffen. Vermenschlichung von KI ließ sich insbesondere gegen Ende des Untersuchungszeitraums feststellen. In den untersuchten englischsprachigen Artikeln über den Einsatz von KI im Gesundheitssystem wurden KI-Systeme oft als soziale Akteure beschrieben, etwa als „Arzt“ oder als „Chef“ (Bunz & Braghieri, 2022).

Die eben erläuterten Studienergebnisse beziehen sich auf die Zeit vor November 2022. Da zu diesem Zeitpunkt das große Sprachmodell ChatGPT des Unternehmens OpenAI veröffentlicht wurde, stellt dieser Zeitpunkt gewissermaßen einen Wendepunkt in der medialen Berichterstattung über KI dar (Tandoc et al., 2025, S. 7). Eine jüngst veröffentlichte Studie untersuchte die mediale Berichterstattung über KI in Singapur in Hinblick auf Framing und Anthropomorphisierung von KI. Die meisten Artikel berichteten positiv oder neutral über KI. Die Autoren konnten in den untersuchten Artikeln aus Singapur kaum Anthropomorphisierung von KI feststellen. KI wurde überwiegend als Werkzeug oder Maschine dargestellt, nicht als sozialer Akteur (Tandoc et al., 2025, S. 12). Nach der Veröffentlichung von ChatGPT im November 2022 veränderte sich die Berichterstattung über KI teilweise. Während technologische Fortschritte weiterhin das dominierende Frame blieben, stieg der Anteil an Artikeln, die Risiken, Herausforderungen sowie Fragen der Regulierung und Verantwortung thematisierten – insbesondere im Zusammenhang mit Plagiaten und Urheberrecht. Anthropomorphisierung von KI war selten zu beobachten, stattdessen wurde KI weiterhin eher als Werkzeug beschrieben. Die Autoren merken an, dass nur Artikel aus Singapur analysiert wurden, wodurch die Repräsentativität eingeschränkt sein könnte (Tandoc et al., 2025, S. 12f.).

Eine qualitative Inhaltsanalyse von südafrikanischen Online-Artikeln der Forscher*innen Brokensha und Conradie aus dem Jahr 2021 identifizierte Anthropomorphisierung von KI als allgegenwärtiges Motiv in der Berichterstattung. Etwa die Hälfte der untersuchten Artikel enthielten Darstellungen von KI als menschliche Akteur*innen. In einigen Artikeln wurde maschinelle Intelligenz mit menschlicher Intelligenz gleichgesetzt oder übertraf diese sogar. Unabhängig von der Anthropomorphisierung wurde KI überwiegend als positiv und für den Menschen nützlich dargestellt. Anthropomorphisierung führte

jedoch zu einer unrealistischen Darstellung dessen, was KI derzeit leisten kann und verdeckte tatsächliche ethische und regulatorische Probleme der KI-Entwicklung (Brokensha & Conradie, 2021).

4.2 Positive Effekte der Anthropomorphisierung von KI

Die Anthropomorphisierung von KI kann sich im Kontext von Technologie-Design positiv auf die Benutzerfreundlichkeit von KI-Systemen sowie auf das Verständnis der Technologie auswirken. Sie kann die Interaktion zwischen Anwender*innen und KI erleichtern und das Vertrauen in KI-Systeme steigern. Besonders im Kontext von Sprachmodellen, Chatbots und Robotik hat sich gezeigt, dass gezielte anthropomorphe Gestaltungselemente einen positiven Einfluss auf die Nutzererfahrung haben (Darling, 2015; Złotowski et al., 2015). Roboter, die zu therapeutischen Zwecken eingesetzt werden, wie die robotische Robbe *Paro*, sollen Patient*innen als emotionale Stütze dienen und so Pflegepersonal gezielt entlasten. Die Vermenschlichung wird hier zum zentralen Design-Element (Darling, 2015, S. 6).

Ein zentraler Vorteil der Anthropomorphisierung liegt in der besseren Steuerung von Nutzererwartungen. Go und Sundar (2019) empfehlen, dass Website-Designer und PR-Experten vermenschlichte Elemente gezielt einsetzen sollten, um die Interaktion mit Chatbots zu optimieren und negative Nutzererfahrungen zu vermeiden. Für die Gestaltung von KI-Systemen folgt daraus, dass Nutzer*innen diese intuitiver verstehen und effektiver nutzen, wenn bestimmte Elemente vermenschlicht werden (Go & Sundar, 2019, S. 314).

Empirische Studien haben zudem gezeigt, dass ein maßvoller Einsatz von Anthropomorphisierung Vertrauen in KI-Systeme schaffen und Berührungängste abbauen kann. Dies gilt insbesondere für den Einsatz von Sprache. Ein gezieltes sprachliches Framing – also die bewusste Nutzung anthropomorphisierter Sprache, etwa die Verwendung der dritten Person Singular („er“ oder „sie“) anstelle der neutralen Formulierungen „das System“ oder „der Roboter“ – kann die Akzeptanz von KI positiv beeinflussen (Nijssen et al., 2019; Onnasch & Roesler, 2019). Zu ähnlichen Erkenntnissen kommen ältere Studien, die feststellten, dass die Integration menschlicher Eigenschaften in das Design von Robotern die Interaktion erleichtert und die allgemeine Benutzerfreundlichkeit erhöht (Złotowski et al., 2015). „Beziehungen“ zwischen Menschen und anthropomorphisierten Robotern sind auf der einen Seite kritisch zu betrachten, da sie echte zwischenmenschliche Beziehungen ersetzen, unerwünschtes Verhalten fördern und emotionale Manipulation ermöglichen könnten. Da es klare empirische Belege für ein besseres Verständnis der Funktionsweise von Technologien durch Anthropomorphisierung gibt, warnen Forscher*innen aber davor, diese grundsätzlich abzulehnen (Darling, 2015, S. 14).

Aktuelle Forschungen unterstützen diese Erkenntnisse und liefern weitere empirische Belege für die positiven Effekte von Anthropomorphisierung im Design von KI-Systemen. Cohn et al. (2024) führten ein Online-Experiment durch, in dem sie untersuchten, wie verschiedene anthropomorphe Elemente die Nutzerwahrnehmung eines großen Sprachmodells beeinflussen. Dabei wurden insbesondere zwei

Formen der Anthropomorphisierung getestet: die Nutzung einer computergenerierten Stimme sowie die Verwendung der 1. Person Singular („Ich“). Die Studienergebnisse zeigen, dass die computergenerierte Stimme die wahrgenommene Genauigkeit der Informationen erhöhte. Dies bestätigt, dass gesprochene Sprache als anthropomorphisierendes Element die Nutzerwahrnehmung positiv beeinflussen kann. Die Forscher*innen stellten fest, dass die Verwendung der 1. Person Singular je nach Kontext unterschiedlich bewertet wurde. Während sie in Bezug auf medizinische Informationen das Vertrauen in die Aussagen des Sprachmodells stärkte, führte sie bei Karrieretipps eher zu einer skeptischen Haltung (Cohn et al., 2024, S. 8). Es folgt, dass ein gezielter und kontextabhängiger Einsatz von Anthropomorphisierung dazu beitragen kann, das Vertrauen der Nutzer*innen zu erhöhen und die allgemeine Benutzerfreundlichkeit eines KI-Systems zu verbessern.

4.3 Negative Effekte der Anthropomorphisierung von KI

In diesem Kapitel wird erläutert, inwiefern Anthropomorphisierung von KI-Systemen sich negativ auf unsere generelle Wahrnehmung von KI auswirkt. Die Gründe dafür sind vielfältig. Es werden fünf zentrale Gründe dargelegt, die aus der Forschungsliteratur hervorgehen. Demnach fördert die Anthropomorphisierung von KI unrealistische Erwartungen an die Technologie sowie Ängste und Bedenken im Zusammenhang mit KI. Weiter werden durch die Anthropomorphisierung von KI häufig Stereotype reproduziert. Außerdem kann sie die Regulierung und Gesetzgebung von KI-Technologien sowie die KI-Forschung beeinflussen.

4.3.1 Unrealistische Erwartungen

Ein wesentlicher Kritikpunkt an der Anthropomorphisierung von KI-Systemen ist die Gefahr, dass sie unrealistische Erwartungen an die Technologie fördert. Die Informatikerin Melanie Mitchell warnt in ihrem Text „The metaphors of artificial intelligence“ davor, das Training von großen Sprachmodellen mit menschlichen Lernprozessen gleichzusetzen. Die menschliche Kognition würde dadurch stark vereinfacht und damit die Bedeutung des Menschseins unangemessen reduziert. Dies führt zu einer verzerrten Vorstellung davon, wie KI tatsächlich funktioniert und welche Grenzen sie hat (Mitchell, 2024). Salles et al. legen in einer Studie aus dem Jahr 2020 dar, dass Vermenschlichung in der Kommunikation über KI oft nicht bewusst oder intendiert geschieht, aber dennoch Missverständnisse über die Fähigkeiten dieser Systeme reproduziert und verbreitet. Dies kann entweder zu überhöhten Erwartungen an die Leistungsfähigkeit von KI oder zu unbegründeten Ängsten führen (Salles et al., 2020, S. 90). Darüber hinaus werden in der Forschung mögliche ethische Konsequenzen der Anthropomorphisierung von KI diskutiert. Es besteht das Risiko, dass Nutzer*innen durch die Zuschreibung menschlicher Eigenschaften an KI-Systeme in ihren Entscheidungen manipuliert werden. Besonders problematisch wäre dies,

wenn emotionale Bindungen entstehen, die zeitaufwendig und emotional involvierend sind, aber echte menschliche Interaktionen nicht adäquat ersetzen können (Salles et al., 2020, S. 91).

Die KI-Forschung wird seit langer Zeit stark durch die Neurowissenschaften und das Wissen inspiriert, das uns über die Funktionsweise des menschlichen Gehirns vorliegt. Seit den 1950er-Jahren, als sich das Feld der KI-Forschung etablierte, dient das menschliche Nervensystem als Vorbild für die KI-Entwicklung. Mittlerweile profitieren KI-Forscher*innen und Neurowissenschaftler*innen wechselseitig voneinander: Während KI-Systeme als Modell und Werkzeug in der biologischen Forschung eingesetzt werden, profitiert die KI-Forschung von den Erkenntnissen der Neurowissenschaft und ahmt Gehirnaktivitäten nach (Salles et al., 2020, S. 92).

Da die KI-Forschung stark von anthropomorphisierter Sprache geprägt ist, liegt die Annahme nahe, dass Analogien zwischen menschlichen Eigenschaften und Fähigkeiten von KI-Systemen essenziell für das wechselseitige Verständnis sind. Insbesondere in der Forschung zu sogenannten tiefen neuronalen Netzen („Deep Neuronal Network“, kurz DNN) scheint diese Annahme naheliegend. Es handelt sich dabei um eine zentrale Technologie im KI-Bereich, die bereits vielfach eingesetzt wird – etwa in Übersetzungsprogrammen, in der Gesichtserkennung oder in der Sprachassistenten. Der Erfolg dieser Technologie hat unser Verständnis von KI als eine Erweiterung beziehungsweise Optimierung menschlicher Fähigkeiten geprägt (ebd., S. 92).

Obwohl sich zentrale Aspekte von KI durch Anthropomorphisierungen besser begreifen lassen, warnen Salles et al. vor einer übermäßigen Verwendung. Denn trotz ihrer Ähnlichkeiten sind biologische neuronale Netze um ein Vielfaches komplexer als ihre künstlichen Äquivalente. Insbesondere die Abwesenheit eines echten Bewusstseins ist ein zentraler Unterschied zwischen menschlichen und computerbasierten neuronalen Netzen (ebd., S. 93). Darüber hinaus gelten DNNs als anfällig für Täuschungen, kurzsichtig im Vergleich zum menschlichen Gehirn und ineffizient, da ihr Training große Datenmengen fordert. Tatsächlich können viele menschliche Fähigkeiten und Eigenschaften bereits künstlich nachgeahmt oder repliziert werden. Die Frage, ob man infolgedessen die menschliche Einzigartigkeit in Frage stellen kann, wird von der Philosophie und anderen wissenschaftlichen Disziplinen allerdings kritisch debattiert. Die Autor*innen betonen die Notwendigkeit einer Differenzierung zwischen der Frage, ob KI-Systeme menschliche Eigenschaften haben können und der Annahme, dass Menschen moralisch gesehen einzigartig sind (ebd., S. 93).

Auch Tobias Kopp stellt fest, dass ein überzogenes Maß an Anthropomorphisierung zu Fehleinschätzungen führen und Frustration im Umgang mit KI-Technologien fördern kann. Häufig werden KI-Algorithmen als intelligente Entscheider beschrieben, wodurch die Grenzen zwischen Subjekt und Objekt verschwimmen können (Kopp, 2020, S. 46). Insbesondere die Tatsache, dass öffentliche Debatten über KI sich häufig auf starke KI beziehen, also solche KI, der man eine eigene Kreativität zuschreibt, ist kritisch zu betrachten. Obwohl sich dahinter großes Potenzial verbirgt, bietet sie (bis jetzt) noch keine

Anwendungsmöglichkeiten außerhalb der Forschung. Kopp betont, dass eine unreflektierte Sprachverwendung die Wahrnehmung von KI-Technologien verzerren kann und plädiert für bewusste, sachliche und interdisziplinär verständliche Sprache im Zusammenhang mit KI, um unrealistische Darstellungen zu vermeiden (ebd., S. 47).

4.3.2 Ängste und Bedenken

Eng mit der Etablierung unrealistischer Erwartungen an KI sind Ängste und Bedenken verknüpft, dass Menschen die Kontrolle über KI verlieren könnten. Die dystopischen Vorstellungen reichen bis zu einer absoluten Machtübernahme von KI-gesteuerten Maschinen, die Menschen versklaven. Die Royal Society hebt hervor, dass insbesondere die Überbetonung humanoider Darstellungen in der medialen und populärwissenschaftlichen Berichterstattung die öffentliche Wahrnehmung von KI verzerren kann. Eine solche Darstellung kann Misinformation begünstigen und dazu führen, dass die realen Möglichkeiten und Grenzen von KI falsch verstanden werden (The Royal Society, 2018, S. 4). Zudem kann dadurch die Illusion verstärkt werden, dass KI-Systeme bereits über eine eigenständige Intentionalität oder Entscheidungsfähigkeit verfügen.

Auch Salles et al. stellen fest, dass Anthropomorphisierung von KI Missverständnisse über die tatsächlichen Fähigkeiten von KI-Systemen befeuern kann. Dies kann nicht nur unrealistische Hoffnungen, sondern auch unbegründete Ängste in der Gesellschaft schüren (Salles et al., 2020, S. 90). Des Weiteren fördert die Anthropomorphisierung von KI einerseits die Vorstellung, dass sie Menschen in ihrer Leistungsfähigkeit um ein Vielfaches übertreffen könnten, andererseits aber auch die Angst, dass sie Menschen beherrschen könnten (ebd., S. 93).

Mitchell legt in ähnlicher Weise dar, dass die Vermenschlichung von KI-Systemen im öffentlichen Diskurs dystopische Narrative fördert. Die Vorstellung, dass KI eine eigene Intentionalität oder ein Bewusstsein entwickeln könnte, verstärkt Befürchtungen, dass sie sich der menschlichen Kontrolle entziehen und zur Bedrohung werden könnte. Solche Ängste werden durch anthropomorphisierende Darstellungen legitimiert, obwohl es derzeit keine wissenschaftliche Grundlage für diese Annahmen gibt (Mitchell, 2024).

4.3.3 Reproduktion von Stereotypen

Anthropomorphisierung von KI-Systemen führt nicht nur zu unrealistischen Erwartungen oder Ängsten, sondern verstärkt auch gesellschaftliche Stereotype. Besonders deutlich wird dies in der geschlechtsspezifischen Darstellung von KI, sowohl in realen Anwendungen als auch in der medialen Repräsentation. Im Rahmen eines Experiments zur Interaktion zwischen Menschen und Robotern argumentierte Kate Darling bereits 2015, dass anthropomorphisierte KI-Technologien häufig von sexistischen und rassistischen Stereotypen geprägt sind (Darling, 2015, S. 10).

In vielen Fällen werden Männer als schöpferische Genies inszeniert, während KI-Systeme, insbesondere Sprachassistenten, weiblich konnotiert sind. Nahezu alle erfolgreichen Sprachassistenten – „Siri“ von Apple, „Alexa“ von Amazon oder „Cortana“ von Microsoft – setzen auf weibliche Stimmen und Persönlichkeiten. Diese Designentscheidungen folgen dem Stereotyp der hilfreichen, geduldigen, humorvollen und attraktiven Sekretärin, die unterstützend agiert, aber keine eigenen Entscheidungen trifft. Traditionelle Rollenmuster spielen hier eine zentrale Rolle: Frauen agieren als Unterstützerinnen, Männer als Entscheider. Diese Rollenverteilung spiegelt sich in der Gestaltung von KI-Systemen wider (Bedford-Strohm, 2017, S. 492).

Auch die visuelle Darstellung von KI-Technologien folgt oft stark überzeichneten Geschlechterbildern. KI wird in der Popkultur häufig entweder als hypermaskuline oder hyperfeminine Figur dargestellt. Beispiele dafür sind der muskulöse und aggressive T-800 aus „Terminator“ oder die makellose Ava aus „Ex Machina“ (The Royal Society, 2018, S. 8). Solche Darstellungen zementieren nicht nur stereotype Vorstellungen, sondern verstärken auch die Wahrnehmung von KI als entweder bedrohlich und überlegen oder als anmutig und unterwürfig.

Eine Studie aus dem Jahr 2023 zeigt zudem, dass Frauen im Kontext von KI in Filmen unterrepräsentiert sind. Einer der Gründe dafür ist das Phänomen des sogenannten *Brilliance Bias*: Dies beschreibt das Phänomen, dass Menschen über verschiedene Altersgruppen hinweg außergewöhnliche Intelligenz und Genialität eher mit Männern assoziieren. In den untersuchten Filmen waren 37 von 38 als Genies dargestellte Figuren männlich. Die Darstellung von KI-Wissenschaftlern als außergewöhnlich brillant trägt somit dazu bei, das Bild zu verstärken, dass Frauen weniger für eine Karriere in diesem Feld geeignet sind (Cave et al., 2023). Männer treten in Filmen über KI häufig in der Rolle des Schöpfers oder Erschaffers auf. Sie tun dies häufig, um verlorene Angehörige zu ersetzen, ideale Partnerinnen zu gestalten oder sogar Kopien ihrer selbst zu erschaffen. Diese Darstellung verstärkt die Assoziation zwischen KI und Männlichkeit sowie das Bild des männlichen Wissenschaftlers als zentralen Akteur im Bereich künstlicher Intelligenz (ebd.).

4.3.4 Rechtliche Fragen und Urheberrechtsverletzungen

Die Art und Weise, wie KI-Systeme im öffentlichen Diskurs dargestellt werden, kann sich auch auf die Regulierung und Gesetzgebung von KI-Technologien auswirken. Insbesondere die Anthropomorphisierung von KI in der Berichterstattung trägt dazu bei, dass rechtliche Fragestellungen auf eine Weise diskutiert werden, welche die tatsächlichen technischen und ethischen Herausforderungen verzerren kann. Melanie Mitchell argumentiert, dass Anthropomorphisierungen von KI zunehmend auch in juristischen Debatten genutzt werden – insbesondere im Zusammenhang mit Urheberrechtsverletzungen durch große Sprachmodelle. KI-Anwendungen wie *ChatGPT* werden mit großen Mengen urheberrechtlich geschützter Daten trainiert, darunter Texte, Bilder und Artikel. Dies geschieht ohne die Zustimmung

der ursprünglichen Urheber*innen. Zahlreiche Autor*innen, Künstler*innen und Verlage legen daraufhin Klage gegen große Unternehmen ein, die KI-Systeme entwickeln. Die Unternehmen selbst weisen die Vorwürfe zurück und argumentieren, dass das Training ihrer Modelle unter das Prinzip des fairen Gebrauchs falle (Mitchell, 2024). Besonders deutlich wird diese Argumentation in einem Zitat von Microsoft-CEO Satya Nadella, der das Training großer Sprachmodelle mit menschlichem Lernen vergleicht: Wenn er ein Set an Lehrbüchern lese und daraus neues Wissen ableite, sei dies auch keine Urheberrechtsverletzung (Mitchell, 2024).

Kritiker*innen halten dagegen, dass dieser Vergleich irreführend ist: Während menschliches Lernen auf Verständnis basiert, analysieren KI-Systeme lediglich statistische Muster in großen Datenmengen und erzeugen daraus neue Texte, ohne den ursprünglichen Inhalt wirklich zu „verstehen“. Damit handele es sich nicht um eine kreative Leistung, sondern um eine unzulässige Nutzung geschützter Inhalte (ebd.). Rechtsexpert*innen warnen davor, dass eine zu starke Vermenschlichung von KI-Systemen dazu führen könnte, dass ethische und juristische Fragen verschleiert werden. Wird KI als autonomes, lernendes Wesen dargestellt, könnte dies die Verantwortung der Entwickler*innen und Unternehmen relativieren, so Mitchell (2024).

Auch in anderen Bereichen der KI-Regulierung zeigt sich, dass die verwendeten Metaphern die rechtliche Einordnung dieser Technologien beeinflussen können. Richards und Smart (2016) analysieren, wie die Darstellung von Robotern die Art und Weise prägt, wie Gesetze und Regulierungsvorschriften verfasst werden. Sie sprechen von der sogenannten *Android Fallacy* – dem Phänomen, dass die Anthropomorphisierung und das Framing von Robotern als menschliche Entitäten dazu führen, dass wir sie als eigenständige Akteure wahrnehmen, anstatt als Werkzeuge. Diese Fehlwahrnehmung kann zur Folge haben, dass Gesetze auf Roboter und KI-Systeme angewendet werden, die für menschliche Akteur*innen gedacht sind, obwohl grundlegende Unterschiede bestehen (ebd.).

4.3.5 Auswirkungen auf die KI-Forschung

Wie wir KI wahrnehmen, kann auch die wissenschaftliche Forschung beeinflussen. Anthropomorphisierung von KI kann wissenschaftliche Fragestellungen in eine bestimmte Richtung lenken und beeinflussen, welche Methoden und Konzepte als relevant erachtet werden (Mitchell, 2024). Die Vorstellung von KI als denkende Entität kann etwa dazu führen, dass Forschende Tests anwenden, die ursprünglich für Menschen entwickelt wurden – etwa IQ-Tests oder Persönlichkeitstests. Diese Herangehensweise impliziert wiederum ethische Fragen: Sollten für KI-Systeme ähnliche moralische Maßstäbe gelten wie für menschliche Proband*innen? Wenn Sprachmodelle als Akteure betrachtet werden, erscheinen Fragen nach einem Maschinenbewusstsein oder der moralischen Verantwortung von KI-Systemen plausibel (ebd.). Demgegenüber würde eine alternative Sichtweise, die KI als hochkomplexe Datenbanken oder benutzerfreundliche Enzyklopädien versteht, andere Forschungsschwerpunkte setzen. In diesem

Fall stünde die Untersuchung technischer Aspekte wie Effizienz, Datensicherheit oder algorithmische Verzerrungen im Vordergrund.

Nach Salles et al. könnte die KI-Forschung durch die Anthropomorphisierung von KI eingeschränkt werden. Wie bereits erläutert, besteht eine enge beidseitige Beziehung zwischen der KI-Forschung und der Neurowissenschaft. Wenn das Ziel der KI-Forschung darin besteht, möglichst realistisch menschliche Intelligenz nachzunahmen, dann könnten Bestrebungen, neue, nicht-menschliche Formen von Intelligenz zu schaffen, in den Hintergrund geraten (Salles et al., 2020, S. 91). Alternative Forschungsfragen, die nicht auf eine algorithmische Abbildung neuronaler Netze abzielen, könnten vernachlässigt werden. Die Entwicklung von KI-Technologien könnte folgend eingeschränkt werden. Die Orientierung an menschlicher Intelligenz sollte der Forschung eher als Leitlinie dienen, nicht als Vorgabe, so Salles et al (ebd., S. 94).

5 Implikationen für die Berichterstattung über KI

Aus den eben erläuterten Forschungserkenntnissen gehen Implikationen für die Berichterstattung über KI hervor, die in diesem Kapitel dargelegt werden. Die KI-Experten Kapoor und Narayanan identifizieren in einem Blogpost aus dem Jahr 2022 insgesamt 18 Fallstricke, die es im Journalismus über KI zu vermeiden gilt. Vier dieser Fallstricke lassen sich unter dem Überbegriff "Flawed Human-AI Comparison" zusammenfassen. Darin warnen die Autoren davor, KI eine eigene Handlungsfähigkeit zuzuschreiben. Darstellungen von KI, die vermuten lassen, dass KI-Systeme sich der Überwachung durch Menschen bereits entziehen oder dies bald tun werden, sollten vermieden werden. Visualisierungen von KI als Roboter sind ebenfalls kontraproduktiv. Sie erwecken fälschlicherweise den Eindruck einer Verkörperung der KI, obwohl es sich lediglich um Software handelt, die Muster aus Trainingsdaten erkennt (Kapoor & Narayanan, 2022). Die Autoren raten außerdem von Vergleichen mit menschlichen Lernprozessen ab. Insbesondere DNNs werden häufig mit menschlicher Intelligenz gleichgesetzt. Dies könnte fälschlicherweise zu dem Schluss führen, dass KI „empfindungsfähig“ ist. Schließlich warnen die Autoren davor, KI-Leistung mit menschlichen Fähigkeiten zu vergleichen. Das ließe vermuten, dass KI und Menschen auf Augenhöhe agieren, obwohl die Anwendungsbereiche von KI (noch) begrenzt sind (ebd.).

Das Projekt "Better Images of AI" beleuchtet in einem Bericht aus dem Jahr 2023, warum bestehende Visualisierungen von KI oft irreführend sind. Es wird dargelegt, wie unser Verständnis von KI durch Bilder geprägt wird und warum alternative Darstellungen notwendig sind. Anthropomorphisierung ist ein zentraler Aspekt der verbreiteten KI-Visualisierungen. Häufig gezeigte Motive sind humanoide Roboter, blaue, mechanische Gehirne oder Adaptionen von Michelangelos „Die Erschaffung des Adam“ (Dihal & Duarte, 2023, S. 5). Im Rahmen des Projekts wurden in Zusammenarbeit mit über 100 Expert*innen eine Bilddatenbank mit realistischeren KI-Darstellungen erstellt. Abbildung 2 zeigt vier Beispiele der

neuen KI-Visualisierungen. Auf der Webseite *betterimagesofai.org* werden Stock-Bilder zur Verwendung für Journalist*innen und Wissenschaftskommunikator*innen bereitgestellt.



Abb. 2: Symbolbilder: Neue Visualisierungen von KI (Quelle: *betterimagesofai.org*)

Viele Expert*innen plädieren für neue Begriffe und Sinnbilder, um KI zu erklären (Kopp, 2020; Mitchell, 2024; Salles et al., 2020; The Royal Society, 2018). Die Royal Society schlägt beispielsweise vor, eindimensionale Narrative über humanoide Roboter oder dystopische Zukunftsvisionen zu vermeiden. Stattdessen sollten neue Metaphern unter dem Stichwort „verteilte Intelligenz“ genutzt werden. Diese könnten durch Bienen oder Ameisen oder auch als kollektive Intelligenz in sozialen Systemen (z.B. Unternehmen oder Netzwerke) visualisiert werden (The Royal Society, 2018, S. 21). Eine bewusste, sachliche und interdisziplinär verständliche Sprache ist essenziell, um fundiert über KI zu kommunizieren. Studierende sollten für die Wirkung von Begriffen sensibilisiert werden. Zudem sollte sichergestellt werden, dass ihr Technikverständnis nicht von populären, irreführenden Begriffen geprägt wird (Kopp, 2020, S. 47). Die Informatikerin Katharina Zweig schlägt vor, Klicklaute vor Verben zu verwenden, um zu signalisieren, dass wir über KI sprechen und diese in Schriftform mit einer Tilde zu kennzeichnen: ~nachdenken oder ~analysieren (Zweig, 2025).

Schließlich brauchen Journalist*innen fundierte Kenntnisse über KI, um komplexe Sachverhalte verständlich zu vermitteln. Eine enge Zusammenarbeit zwischen Wissenschaft und Journalismus kann dazu beitragen, irreführende Darstellungen wie die Vermenschlichung von KI zu vermeiden. Institutionen wie das *Science Media Center* können als Vermittlungsstellen zwischen Forschung und Journalismus dienen, doch die Zusammenarbeit gestaltet sich in der Praxis oft schwierig (Henschel, 2023). Angesichts der wachsenden Bedeutung von KI in vielen Lebensbereichen ist eine präzise und verantwortungsbewusste Kommunikation über KI für Journalismus und Forschung gleichermaßen von hoher Relevanz.

6 Fazit

Das Konzept KI ist für uns schwierig zu verstehen. Wir brauchen daher Metaphern, um es annähernd zu erfassen. Dass wir KI vermenschlichen, liegt nahe. Einerseits, weil wir Eigenschaften, die wir an uns selbst und unseren Mitmenschen beobachten, auf technische Systeme projizieren. Andererseits ist die KI-Forschung von so vielen Analogien zu menschlichen Eigenschaften – Künstliche *Intelligenz*, maschinelles *Lernen*, *neuronale* Netze – geprägt, dass es naheliegt, Parallelen zu ziehen. Zumal die Interaktion mit KI für uns auch immer alltäglicher wird.

Die Forschung zeigt, dass die Berichterstattung über KI in den letzten Jahren überwiegend positiv war, wenn auch eher oberflächlich. Seit der Veröffentlichung von ChatGPT im November 2022 spielen KI-Themen nicht nur in den Medien eine immer größer werdende Rolle. In der Bildung, in der Medizin, in der Forschung und im Journalismus gilt es, schnelle Antworten auf die Fragen zu finden: Wie gehen wir mit der rasanten Entwicklung dieser Technologie um? Wie können wir sie sinnvoll für uns nutzen? Und: Wie gehen wir mit Falschinformationen um?

Die Implementierung menschlicher Eigenschaften in KI-Technologien kann die Benutzerfreundlichkeit sowie das Verständnis der Technologie erhöhen. Wenngleich positive Effekte in der Anwendung nicht von der Hand zu weisen sind, überwiegen negative Effekte der Anthropomorphisierung von KI in der öffentlichen Wahrnehmung der Technologie. Wenn die Interaktion mit KI-Systemen so nah an Begegnungen mit Menschen herankommt, könnten wir echte zwischenmenschliche Beziehungen vernachlässigen. Außerdem fördert die Anthropomorphisierung unrealistische Erwartungen an KI. Informationen zu prüfen, wird vor dem Hintergrund fundierter (visueller) Beweise vermeintlich obsolet. Dystopische Narrative, in denen Menschen die Kontrolle über KI verlieren, erscheinen realistischer. Darüber hinaus reproduzieren Darstellungen von KI in menschlicher Form häufig gesellschaftliche Stereotype, etwa in Bezug auf das Geschlecht oder die Herkunft. Debatten rund um rechtliche Fragen werden verzerrt, wenn KI-Systemen eine eigene Handlungsfähigkeit zugeschrieben wird. Relevante Fragen nach der Verantwortlichkeit oder Urheberrechtsfragen könnten dadurch aus dem Fokus geraten. Schließlich wird auch die wissenschaftliche Forschung dadurch beeinflusst, welche Fähigkeiten und Eigenschaften wir KI zuschreiben und welche nicht.

Die Art und Weise, wie wir über KI berichten, formt unser Bild davon, wie KI funktioniert, was sie kann und wie wir sie nutzen können. Eine zentrale Erkenntnis für die Kommunikation über KI ist die Tatsache, dass das Konzept KI schwer zu erfassen ist und nur wenige Expert*innen ein umfassendes Verständnis davon haben. Es wurde gezeigt, dass die Berichterstattung über KI, wie sie jetzt stattfindet, die Wahrnehmung von KI verzerren kann und zentrale Herausforderungen im Umgang mit KI aus dem Fokus geraten. Daher brauchen wir dringend neue Bilder, um KI zu beschreiben - sowohl visuell als auch metaphorisch.

7 KI-Nachweis

In der vorliegenden Hausarbeit wurde das Sprachmodell ChatGPT (aktuelle Version, Stand März 2025) unterstützend eingesetzt. Der Einsatz erfolgte ausschließlich in folgenden Bereichen:

- **Zusammenfassung bereits gelesener Texte:** ChatGPT wurde genutzt, um Inhalte aus wissenschaftlicher Literatur in komprimierter Form darzustellen. Inhalte aus diesen Zusammenfassungen wurden teilweise in den Text übertragen, jedoch stets auf Korrektheit überprüft und nicht wortwörtlich übernommen.
- **Umformulierung eigener Notizen:** Eigene, auf Forschungsliteratur basierende Notizen wurden mit Hilfe von ChatGPT in Fließtext überführt. Auch hier erfolgte eine sorgfältige Prüfung und Überarbeitung, bevor die Texte in die finale Fassung der Arbeit integriert wurden.

Die inhaltliche und sprachliche Gestaltung der Arbeit unterliegt der Verantwortung der Verfasserin. Alle inhaltlichen Aussagen und Ergebnisse der Literaturanalyse wurden eigenständig überprüft, reflektiert und strukturiert. Somit verbleibt die wissenschaftliche Verantwortung vollständig bei der Verfasserin.

8 Literaturverzeichnis

- Bedford-Strohm, J. (2017). Voice first? Eine Analyse des Potentials von intelligenten Sprachassistenten am Beispiel Amazon Alexa. *Communicatio Socialis*, 50(4), 485–494.
<https://doi.org/10.5771/0010-3497-2017-4-485>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Brokensha, S., & Conradie, T. (2021). Killer robots, humanoid companions, and super-intelligent machines: The anthropomorphism of AI in South African news articles. *Journal for Cultural Studies*. <http://ensovoort.co.za/index.php/2021/06/07/killer-robots-humanoid-companions-and-super-intelligent-machines-the-anthropomorphism-of-ai-in-south-african-news-articles/>
- Bundesamt für Sicherheit in der Informationstechnik (BSI). (2025). *Künstliche Intelligenz*. Abgerufen am 14.07.2025 unter: <https://www.bsi.bund.de/dok/13394406>
- Bunz, M., & Braghieri, M. (2022). The AI doctor will see you now: Assessing the framing of AI in news coverage. *AI & Society*, 37(1), 9–22. <https://doi.org/10.1007/s00146-021-01145-9>
- Cave, S., Dihal, K., Drage, E., & McInerney, K. (2023). Who makes AI? Gender and portrayals of AI scientists in popular film, 1920–2020. *Public Understanding of Science*, 32(6), 745–760.
<https://doi.org/10.1177/09636625231153985>
- Chiang, T. (2023, February 9). ChatGPT is a blurry JPEG of the web. *The New Yorker*.
<https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web>
- Chuan, C.-H., Tsai, W.-H. S., & Cho, S. Y. (2019). Framing artificial intelligence in American newspapers. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 339–344.
<https://doi.org/10.1145/3306618.3314285>

- Cohn, M., Pushkarna, M., Olanubi, G. O., Moran, J. M., Padgett, D., Mengesha, Z., & Heldreth, C. (2024). Believing anthropomorphism: Examining the role of anthropomorphic cues on trust in large language models (Version 1). *arXiv*. <https://doi.org/10.48550/arXiv.2405.06079>
- Costa, P., & Ribas, L. (2019). AI becomes her: Discussing gender and artificial intelligence. *Technoetic Arts*, 17(1), 171–193. https://doi.org/10.1386/tear_00014_1
- Darling, K. (2015). „Who’s Johnny?“ Anthropomorphic framing in human-robot interaction, integration, and policy. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2588669>
- Dihal, K., & Duarte, T. (2023). *Better images of AI: A guide for users and creators*. The Leverhulme Centre for the Future of Intelligence & We and AI. <https://blog.betterimagesofai.org/wp-content/uploads/2023/02/Better-Images-of-AI-Guide-Feb-23.pdf>
- Epley, N. (2018). A mind like mine: The exceptionally ordinary underpinnings of anthropomorphism. *Journal of the Association for Consumer Research*, 3(4), 591–598. <https://doi.org/10.1086/699516>
- Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1). <https://doi.org/10.1609/aaai.v31i1.10635>
- Fucker, S. (2022). Medien zwischen Angstmachern und Hoffnungsstiftern: Zur emotionalen Wirkung der medialen Berichterstattung über künstliche Intelligenz. In F. van Oorschot & S. Fucker (Hrsg.), *Framing KI: Narrative, Metaphern und Frames in Debatten über Künstliche Intelligenz*. Heidelberg: heiBOOKS (FEST Forschung, Band 2). <https://doi.org/10.11588/HEI-BOOKS.1106.C15194>
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity, and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Henschel, A. (2023, November 7). Kommunizieren wir an der KI-Forschung vorbei? *Wissenschaftskommunikation.de*. <https://www.wissenschaftskommunikation.de/kommunizieren-wir-an-der-ki-forschung-vorbei-72093/>
- Kapoor, S., & Narayanan, A. (2022, September 30). Eighteen pitfalls to beware of in AI journalism. *AI Snake Oil*. <https://www.aisnakeoil.com/p/eighteen-pitfalls-to-beware-of-in>
- Kopp, T. (2020). Die Wirkkraft der Sprache auf die Wahrnehmung neuartiger Technologien wie kollaborationsfähiger Roboter (Cobots) oder künstlicher Intelligenz (KI). *Tagungsband zur 33. AKWI-Jahrestagung*, 41–52. http://akwi.de/documents/AKWI_Tagungsband2020.pdf
- Li, M., & Suh, A. (2022). Anthropomorphism in AI-enabled technology: A literature review. *Electronic Markets*, 32(4), 2245–2275. <https://doi.org/10.1007/s12525-022-00591-7>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12. <https://doi.org/10.1609/aimag.v27i4.1904>
- McDermott, D. (1976). Artificial intelligence meets natural stupidity. *ACM SIGART Bulletin*, 57, 4–9. <https://doi.org/10.1145/1045339.1045340>
- Mitchell, M. (2024). The metaphors of artificial intelligence. *Science*, 386(6723), eadt6140. <https://doi.org/10.1126/science.adt6140>
- Nelson, T. E., Oxley, Z. M., & Clawson, R. A. (1997). Toward a Psychology of Framing Effects. *Political Behavior*, 19(3), 221–246. <https://doi.org/10.1023/A:1024834831093>

- Nijssen, S. R. R., Müller, B. C. N., van Baaren, R. B., & Paulus, M. (2019). Saving the Robot or the Human? Robots Who Feel Deserve Moral Care. *Social Cognition*, 37(1), 41-S2. <https://doi.org/10.1521/soco.2019.37.1.41>
- Onnasch, L., & Roesler, E. (2019). Anthropomorphizing Robots: The Effect of Framing in Human-Robot Collaboration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 1311–1315. <https://doi.org/10.1177/1071181319631209>
- Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, 35(4), 927–936. <https://doi.org/10.1007/s00146-020-00965-5>
- Richards, N. M., & Smart, W. D. (2016). How should the law think about robots? In R. Calo, A. M. Froomkin, & I. Kerr (Hrsg.), *Robot Law*. Edward Elgar Publishing. <https://doi.org/10.4337/9781783476732.00007>
- Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987), 493–498. <https://doi.org/10.1038/s41586-023-06647-8>
- Tandoc, E. C., Seet, S., Chan, V. X., & Wong, P. J. O. (2025). Exploring AI identity: The media framing of communicative artificial intelligence in Singapore’s news sites. *Public Understanding of Science*. <https://doi.org/10.1177/09636625251317970>
- The Royal Society. (2018). Portrayals and perceptions of AI and why they matter. Abgerufen am 14.07.2025 unter: <https://royalsociety.org/news-resources/projects/ai-narratives/>
- Vallor, S. (2024). *The AI mirror: How to reclaim our humanity in an age of machine thinking*. New York: Oxford University Press.
- Yiu, E., Kosoy, E., & Gopnik, A. (2024). Transmission Versus Truth, Imitation Versus Innovation: What Children Can Do That Large Language and Language-and-Vision Models Cannot (Yet). *Perspectives on Psychological Science*, 19(5), 874–883. <https://doi.org/10.1177/17456916231201401>
- Złotowski, J., Proudfoot, D., Yogeewaran, K., & Bartneck, C. (2015). Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction. *International Journal of Social Robotics*, 7(3), 347–360. <https://doi.org/10.1007/s12369-014-0267-6>
- Zweig, K. (2025, February 27-28). *Plappernde Papageien oder Superhumans - Was können Sprachmodelle?* [Conference presentation]. GAL Research School: Methoden der linguistischen Analyse von und mit KI, Karlsruhe.