Das Sprechen über KI und AI Literacy

Didaktische Konsequenzen aus einer Korpusanalyse[1]

Sophia Hercher[2]

Versteht generative KI, was sie da sagt? Eine intuitive Betrachtung des Sprachgebrauchs im Zusammenhang mit sprachgenerativer KI legt ein hohes Maß an Anthropomorphisierung nahe, das ist die Zuschreibung menschlicher Qualitäten auf nichtmenschliche Dinge. Diese spiegelt sich in unserer Nutzung der Systeme wider. Durch die Zuschreibung, dass generative KI die Bedeutung des Outputs kennt, schreiben wir beispielsweise den Texten einen Informations- oder sogar Wahrheitsgehalt zu. Dies führt jedoch zu einem falschen Verständnis der KI-Funktionsweise und zieht potenziell eine ungünstige Entwicklung der KI-Kompetenz nach sich.

In diesem Beitrag wird die linguistische Repräsentation des Verständnisses text-generativer KI mit Methoden der Korpuslinguistik untersucht, um die intuitive Betrachtung empirisch zu überprüfen. Hieraus werden Konsequenzen für die didaktische und sprachliche Gestaltung von Aufgaben in der Hochschullehre, bei denen text-generative KI zum Einsatz kommt, abgeleitet. Im Zentrum steht dabei die originäre Nutzungsweise und der konstruktive Umgang mit dem Output sowie die Betonung der aktiven Rolle Lernender.

Talking about AI and AI literacy – didactic design consequences from a corpus analysis

Does text-generative AI understand its own output? Looking at how we talk about generative AI suggests a high degree of anthropomorphism, that is the attribution of human characteristics onto non-human things. The way we speak about the AI systems can manifest in the way we use them. Assuming that generative AI knows the meaning of its output, we ascribe the generated texts informative value or substance. This, however, may lead to a wrong understanding of its functionality and a resulting adverse development of AI literacy.

This contribution studies the linguistic representation of our understanding of text-generative AI utilizing methods from corpus linguistics in order to verify intuitive assumptions about general language use empirically. From the results, consequences for higher education didactics and university teaching are drawn, specifgically for the design of tasks that include the

1 Basiert auf einem Impulsbeitrag im Rahmen der Tagung.

2 ORCID: 0009–0008-6018-1161

usage of text-generative systems. At the center of the recommendations are the focus on the original design of the technology, the critical interaction with output as well as the emphasis of learners' active roles in the process.

Die Darstellung eines Chatbots als Mensch

Die Nähe zu menschlicher Sprache, die im Output textgenerativer KI zu erkennen ist, wirkt gleichermaßen eindrucksvoll wie potenziell besorgniserregend. In einer Studie zur linguistischen Repräsentation konzeptueller Metaphern gehe ich der Frage nach, ob ChatGPT generell anthropomorphisiert[3] wird und in welchen Kontexten dies geschieht (Hercher, 2026, angenommen). Im vorliegenden Beitrag umreiße ich die Kernannahmen und -ergebnisse dieser Arbeit, um Ableitungen für die Hochschuldidaktik und Hochschullehre vorzunehmen.

Ausgangspunkt für die Untersuchung war die Vermutung zahlreicher Autor*innen, dass die Anthropomorphisierung von KI in einem Missverständnis der Technologie (vgl. Bender/Koller 2020: 5186; Bender et al. 2021: 617) und entsprechend ungeeigneter Nutzungsweisen resultieren könnte (vgl. Hunger 2023: 4; Shanahan 2024: 79). Stellt man dies der Forderung nach einer reflektierten und konstruktiven Nutzungsweise von KI-Systemen in der Bildung gegenüber, ergibt sich die Frage, ob zur KI-Kompetenz auch eine bestimmte Weise des Sprechens über KI gehört, z.B. in einer nicht-anthropomorphisierenden Weise. Und wie könnte Kommunikation in der Bildung gestaltet sein, um diese möglichen negativen Konsequenzen zu vermeiden?

Die Forderung nach der Reduzierung von Anthropomorphismen (Hunger 2023) erinnert an die Metapherntheorie, welche ihren Ursprung in den 1980ern mit dem Werk Metaphors We Live By hat (vgl. Lakoff/Johnson 1980). Die Autoren zeigen, dass die Weise, in der wir sprechen, Aufschluss über unser Verstehen geben kann. Metaphern sind demnach nicht lediglich rhetorische Mittel, sondern essenzieller Teil unserer kognitiven Strukturen. So würde zum Beispiel der Satz »Solide Annahmen bilden das Fundament meiner Theorie.« einen Hinweis auf die konzeptuelle Metapher THEORIEN SIND GEBÄUDE[4] geben. Die Weise, in der wir Sprache nutzen, kann also Aufschluss darüber geben, wie wir Dinge verstehen (vgl. Lakoff/Johnson 1980: 5). Im Beispiel werden für Theorien ähnliche oder identische Konzepte benutzt wie für Gebäude, ausgedrückt durch die Worte Fundament und solide. Die Autoren erklären, dass die Übertragung von etwas uns Vertrautem (Menschsein) auf etwas weniger Vertrautes oder Abstraktes (Computer) stattfindet (vgl. Lakoff/Johnson 1980:

3 Ein Anthropomorphismus ist die Zuschreibung menschlicher Eigenschaften auf nichtmenschliche Dinge.

4 Konzeptuelle Metaphern werden konventionell in Kapitälchen geschrieben.

105). Bereits 1991 unternahmen Lakoff et al. den Versuch, eine umfassende Liste solcher konzeptuellen Metaphern aufzustellen. Hier findet sich auch die konzeptuelle Metapher COMPUTERS/PROGRAMMS ARE INTELLIGENT PEOPLE WITH INDEPENDENT WILLS. Dies deckt sich mit der Beobachtung, dass Anthropomorphisierung eine wichtige Rolle bei der Definition von neuen Technologien spielt (vgl. Carbonell/Sánchez-Esguevillas/Carro 2016: 148).

Anthropomorphismen von KI sind in vielen Bereichen prävalent: Systemseitig werden z.B. »Denkprozesse« in animierter Weise dargestellt (vgl. Abercrombie et al. 2023: 4778) oder der Bot »spricht« in der ersten Person singular. Skulmowski (2024) erwähnt in seiner Übersicht über die Effekte von KI-Anthropomorphisierung die Reduktion psychologischer Distanz zum System (vgl. Li/Sung 2021: 6) und die Steigerung der Kundenzufriedenheit (vgl. Janson 2023: 11). Demnach hebt Anthropomorphisierung tendenziell eher positive Eigenschaften eines Systems hervor.

Mit der Annahme, dass Sprache Aufschluss über die Denkstrukturen gibt, geht auch einher, dass linguistische Repräsentation sich auf das Verständnis eines Konzepts auswirken kann (vgl. Lakoff/Johnson 1980: 34). Hieran knüpfen die eingangs genannten Befürchtungen an: Durch übermäßige Anthropomorphisierung könnte ein falsches Verständnis der Funktionalität der Technologie entstehen. Dies reicht von unangemessenem Vertrauen in die »Fähigkeiten« von KI-gestützten Chatbots, über die daraus resultierende Nutzung der Systeme für Zwecke, für die sie nicht entwickelt wurden[5], bis hin zu einer möglichen Verantwortungsdiffusion durch die Attribution von Willenskraft oder Eigenständigkeit des Chatbots.

Um Konsequenzen für die Hochschullehre aus diesem Sprachgebrauch abzuleiten, ist zunächst zu identifizieren, wie tiefgreifend die Anthropomorphisierung von KI ist und in welchem Zusammenhang sie speziell auftaucht.

Methodischer Zugang und Datenbeispiele

In der Linguistik ist es üblich, Intuitionen über Sprachgebrauch mit großen Mengen an Sprachdaten zu überprüfen, so auch in der Metaphernforschung (z.B. Deignan 2005). Hierfür werden Sprachkorpora genutzt, also große Sammlungen an authentischem, repräsentativem Sprachmaterial. Für die hier referenzierte Untersuchung wurde der *News on the Web Corpus* (NOW, Davies 2016) genutzt, der umfassende und tagesaktuelle Sprachdaten in Form von Nachrichtentexten aus dem Internet enthält. Er ist für das Vorhaben gut geeignet, weil die enthaltenen Nachrichtentexte in der Regel eine große Reichweite haben und als wichtige Quelle für Informatio-

5 vgl. Bender und Kollers Ausführungen zur »Bedeutung« von Output textgenerativer KI, S. 5186f.

nen und die öffentliche Meinungsbildung angesehen werden (vgl. Krennmayr 2015: 531).

Für die Untersuchung wurde nach Konstrukten der Form »ChatGPT VERB« im Zeitraum November 2022 bis November 2024 gesucht, um Personifikationen ausfindig zu machen. Nach der Datenbereinigung wurden aus den 100 häufigsten Verbkonstruktionen 4052 Konkordanzen (Treffer inkl. Kurzkontext) extrahiert. Typische Beispiele aus dem Datenset sind:

1. »A lawyer is currently facing harsh consequences in court after ChatGPT **made up** six cases that the lawyer cited without first verifying case details …« (2023–06–09, US, arstechnica.com)
2. »… ChatGPT **knows** what fact checking means.« (2023–08–16, NG, staradvertiser.com)
3. »… ChatGPT **refused** to write a poem about Donald Trump, referring to the president as a [a model of hate speech]« (2023–02–15, US, nypost.com)
4. »… ChatGPT **recommended** Etsy, Amazon and Zales – again, nothing I couldn't have thought [of myself].« (2024–11–12, AU, cnet.com)

Anhand der Beispiele lässt sich die Vielseitigkeit der Anthropomorphismen in den Nachrichtentexten ablesen. In Beispiel 1 wird es so dargestellt, als ob sich das System etwas ausdenken könne. Nicht nur die Annahme, dass ChatGPT derart kreativ sei, auch die Information dazu, dass ein Anwalt solche Inhalte bereits aktiv in einer Gerichtsverhandlung genutzt hat, zeigen auf, welche Konsequenzen unangemessenes Vertrauen in die Richtigkeit des KI-generierten Outputs haben kann. In Beispiel 2 wird nicht nur behauptet, dass ChatGPT irgendetwas wisse, sondern gleichzeitig, dass der Wissensinhalt sich auf das Überprüfen von Fakten beziehe. Als System, welches Worte auf Basis zuvor berechneter Wahrscheinlichkeiten aneinanderreiht, ist es ChatGPT nicht möglich, Fakten zu überprüfen, da es kein Konzept von »Bedeutung« besitzt (vgl. Shanahan 2024: 72; Bender/Koller 2021). Die Beschreibung des Outputs als Weigerung in Beispiel 3 unterstellt dem System, dass es nicht nur Kenntnis der Bedeutung des generierten Outputs hat, sondern auch eigene Willensstärke, die es nutzen kann, um einer Aufforderung nicht nachzukommen. Hier wird die Programmierung des Systems mit dessen Intention gleichgesetzt. Das Beispiel verdeutlicht zudem den systeminhärenten Bias, der durch die Vorauswahl von Datensätzen entsteht. Man kann es gut oder schlecht finden, dass das System bestimmte Outputs nicht generiert und es gibt gewiss auch Texte, die nicht produziert werden sollten. Doch besonders das dritte Beispiel zeigt, welche Kontrolle über die Ausgabe ausgeübt werden kann und schon wird (siehe auch Morgan/Shepardson über den *regulation ban* der US-Regierung). Dass die Ausgaben von KI-Chatsystemen zum Beispiel in Form von Namensnennungen bestimmter Firmen kommerzialisiert werden

(vgl. Bsp. 4), stellt eine weitere Ausprägung der Informationskontrolle dar, in der Anthropomorphismen in Form persönlicher Empfehlungen genutzt werden.

Dies sind nur vier typische Beispiele, in denen einem KI-System Wissen über die Bedeutung von Text-Output, Verständnis des Inhalts, eigene Willenskraft oder Meinung zugeschrieben werden. Insgesamt konnten etwa die Hälfte der analysierten Konkordanzen als Personifikationen klassifiziert werden. Zu etwa einem Drittel waren dies sehr starke Zuschreibungen in der eben beschriebenen Form. Weiterhin bezogen sich die Personifikationen überwiegend auf den Output bzw. die Interaktion mit dem Chatbot, die i.d.R. das Verstehen des Inhalts seitens des Bots implizierten. Somit kann gefolgert werden, dass Anthropomorphismen prävalent sind und es insbesondere das Beschreiben der Funktionalität ist, für das sie genutzt werden.

Diskussion und Anwendung auf die Hochschullehre

Nun könnte man entgegenhalten, dass diese Beschreibungen eher harmlos sind. Wie auch Shanahan (2024) hierzu hervorhebt, verschwimmen bei textgenerativen KI-Systemen jedoch die Grenzen zwischen Sprachgebrauch und resultierender Nutzung, weil die Sprache in beiden Fällen das Medium ist. Es ist wichtig und teilweise schwierig, zu verstehen, dass es nicht ChatGPT ist, das Empfehlungen für das Shopping ausspricht (Beispiel 4) oder sich weigert, ein Gedicht über Donald Trump zu schreiben (Beispiel 3), sondern die Personen und Firmen, welche die Entwicklung des Chatbots verantworten. Die Annahme eines neutralen und wohlwollenden Gesprächspartners ist nicht nur falsch, sondern potenziell gefährlich, weil Menschen zu Entscheidungen auf der Grundlage falscher Annahmen bewogen werden können.

Doch auch die Deanthropomorphisierung ist keineswegs trivial, denn oft kennen wir keine nicht-anthropomorphisierenden Begriffe in diesem Kontext. Sachangemessene Beschreibungen können ferner als umständlich empfunden werden (Shanahan/McDonell/Reynolds, 2023: 493). Zusätzlich ist zu beobachten, dass Prompt-Kurse oder Hilfeseiten nahelegen, emotionale Sprache zu nutzen (Mok, 2023) oder Rollen zuzuweisen, um bessere Outputs zu generieren (Schulhoff, 2024). Die menschenähnliche Kommunikationsweise erleichtert die Interaktion für die Nutzenden. Auch wenn die Ratschläge zur Nutzung sich fast noch schneller ändern als neue Modelle veröffentlicht werden, liegt hierin der zentrale Konflikt. Die Verbesserung des Outputs durch Anthropomorphisierung seitens der Nutzenden steht im direkten Gegensatz zur akkuraten Beschreibung und daraus resultierendem Verständnis des Systems.

Es wird deutlich, dass KI-Kompetenz auch in Bezug auf das Sprechen über KI mehrere Facetten beinhaltet. Lehrende sollten deshalb ein gutes Gefühl für die Angemessenheit bestimmter Ausdrucksweisen und die Etablierung akkurater Sprach-

nutzung entwickeln, um den Aufbau von KI-Kompetenz in diesem Bereich gut unterstützen zu können. Hierfür werden im Folgenden ein paar Vorschläge gemacht.

Ein wichtiger Bereich der Kommunikation, in dem Lehrende die Funktionsweise von KI-Systemen adressieren, ist die Aufgabenstellung. Wenn wir vermeiden wollen, dass aus der Anthropomorphisierung der Systeme eine unsachgemäße Nutzung resultiert, können wir in diesem Bereich besondere Sorgfalt aufwenden. Im Folgenden werden drei einfache Prinzipien des Aufgabendesigns vorgeschlagen, die bei der Entwicklung von KI-Kompetenz helfen sollen. Diese sind:

1. Betonung der aktiven Rolle der Lernenden in der Interaktion mit dem Chatbot
2. Die Hervorhebung des eigentlichen Zwecks des Systems (Texterstellung)
3. Die Förderung der kritisch-konstruktiven Interaktion mit Output

Um die aktive Rolle der Lernenden hervorzuheben, reicht es zum Beispiel schon, statt »Fragen Sie die KI nach Vorschlägen für eine Werbekampagne«, zu sagen »Ermitteln Sie unter Nutzung eines textgenerativen KI-Systems, Vorschläge für die Werbekampagne«. Hier wird der Werkzeug-Charakter hervorgehoben und die Nutzenden bleiben die Akteure.

Um die eigentliche Funktion zu nutzen in den Vordergrund zu stellen, wäre etwa denkbar, das System zur Herstellung von Textformen zu nutzen. Ein Beispiel hierfür könnte sein: »Erstellen Sie mit generativer KI ein Gedicht. Erörtern Sie, welche Weltansichten in der Ausführung erkennbar werden. Welche Rückschlüsse lässt das Ergebnis auf Trainingsdaten zu?« Eine solche Aufgabe hebt den textgenerativen Charakter des Systems hervor und leitet zur Reflexion der Trainingsdatensätze und deren Auswirkung auf die Textgenese ein. Besonders geeignet für die kritisch-konstruktive Interaktion mit dem Output sind auch Vergleichsaufgaben, in denen der gleiche Prompt in unterschiedliche KI-Systeme eingegeben wird, um die Qualität der Outputs begründet zu bewerten. Grundsätzlich sollten Lehrende Aufgaben vermeiden, bei denen eine Kenntnis oder ein Verständnis des Inhalts seitens des KI-Chatbots vorausgesetzt ist, z. B. Rechercheaufgaben, denn die Gleichsetzung generativer KI-Chatbots mit Suchmaschinen oder gar Datenbanken scheint eine sehr häufige zu sein, obgleich sie irreführend ist.

Fazit

Zusammenfassend lässt sich sagen, dass Anthropomorphismen ein fester Bestandteil des Sprechens über KI sind. Jedoch ist eine vollständige Deanthropomorphisierung des Sprechens über generative KI-Systeme wahrscheinlich sehr schwierig und vielleicht nicht einmal wünschenswert oder möglich, wenn es um die Gestaltung von Prompts geht. Die Prägung unserer Wahrnehmung und unseres Verstehens der

Systeme durch Anthropomorphismen hin zu unangemessenem Vertrauen in deren Outputs steht im Kontrast zur Steigerung der Qualität durch die Nutzung menschenähnlicher Sprache bei der Interaktion mit den Systemen. Jedoch haben wir als Lehrende die Möglichkeit, unsere Kommunikation so zu gestalten, dass die aktive Rolle Lernender in den Fokus gerückt wird und wir uns durch die bewusste Verwendung unserer Sprache regelmäßig die tatsächliche Funktionsweise der KI-Systeme in unser und das Bewusstsein der Lernenden rufen.

Literatur

Abercrombie, Gavin/Cercas Curry, Amanda/Dinkar, Tanvi/Rieser, Verena/Talat, Zeerak (2023): »Mirages. On Anthropomorphism in Dialogue Systems« in: Houda Bouamor/Juan Pino/Kalika Bali (Hg.), Proceedings of the 2023 Conference on Empirical Methods. in: Natural Language Processing Singapore: Association for Computational Linguistics, S. 4776–90.

Bender, Emily M./Gebru, Timnit/McMillan-Major, Angelina/Shmitchell, Shmargaret (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, FAccT ›21. New York, NY, USA: Association for Computing Machinery, S. 610–23. https://doi.org/10.1145/3442188.3445922.

Bender, Emily M./Koller, Alexander (2020): »Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data« in: Dan Jurafsky/Joyce Chai/Natalie Schluter/Joel Tetreault (Hg.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, USA: Association for Computational Linguistics, S. 5185–98. https://doi.org/10.18653/v1/2020.acl-main.463.

Carbonell, Javier/Sánchez-Esguevillas, Antonio/Carro, Belén (2016): »The Role of Metaphors in the Development of Technologies. The Case of the Artificial Intelligence.«, in: Futures 84, S. 145–53. https://doi.org/10.1016/j.futures.2016.03.019.

Davies, Mark (2016): »Corpus of News on the Web (NOW).« https://english-corpora.org/now

Deignan, Alice (2005): Metaphor and Corpus Linguistics, Converging Evidence in Language and Communication Research. Amsterdam [u.a.]: Benjamins.

Hercher, Sophia (2026, angenommen): »What's ChatGPT Doing? – Towards a Metaphor Profile of Generative AI (Working Title).« Unveröffentlichter Beitrag, in Planung, in: Callies, Marcus und Peters, Arne. Cognitive Sociolinguistic Studies on Conceptualisation: Contexts, Data, and Methods (Working Title). Amsterdam.

Hunger, Francis (2023): »Unhype Artificial ›Intelligence‹! A Proposal to Replace the Deceiving Terminology of AI«, in: Zenodo, 1.

Janson, Andreas (2023): »How to Leverage Anthropomorphism for Chatbot Service Interfaces: The Interplay of Communication Style and Personification«, in: Computers in Human Behavior 149, S. 107954.

Krennmayr, Tina (2015): »What Corpus Linguistics Can Tell Us about Metaphor Use in Newspaper Texts«, in: Journalism Studies 16, 4, S. 530–46.

Lakoff, George/Espenson, Jane/Goldberg, Adele (1991): »Master Metaphor List.« http://araw.mede.uic.edu/alansz/metaphor/METAPHORLIST.pdf

Lakoff, George/Johnson, Mark (1980): Metaphors We Live by, Chicago, Ill. [u.a.]: Univ. of Chicago Press.

Li, Xinge/Sung, Yongjun (2021): »Anthropomorphism Brings Us Closer: The Mediating Role of Psychological Distance in User–AI Assistant Interactions«, in: Computers in Human Behavior, 118, S. 106680.

Morgan, David/Shepardson, David (2025): »US Senate Strikes AI Regulation Ban from Trump Megabill.«, in: https://www.reuters.com/legal/government/us-senate-strikes-ai-regulation-ban-trump-megabill-2025-07-01/ vom 01.06.2025.

Schulhoff, Sander (2024): »Rollen Prompting.«, in: https://learnprompting.org/de/docs/basics/roles

Shanahan, Murray (2024): »Talking about Large Language Models«, in: Commun. ACM, 67 (2), S. 68–79.

Shanahan, Murray/McDonell, Kyle/Reynolds, Laria (2023): »Role Play with Large Language Models.«, in: Nature 623, S. 493–98.

Skulmowski, Alexander (2024): »Placebo or Assistant? Generative AI Between Externalization and Anthropomorphization.«, in: Educational Psychology Review 36: 2, S. 58.