

Deutschunterricht im Zeichen Künstlicher Intelligenz

Bereits 1673 denkt René Descartes in seiner *Abhandlung über die Methode, richtig zu denken und die Wahrheit in den Wissenschaften zu suchen* über Maschinen nach, die dem Menschen perfekt nachempfunden sind. Bei der Frage nach möglichen Unterschieden kommt er zum Urteil, dass

diese Maschinen nie sich der Worte oder Zeichen bedienen können, durch deren Verbindung wir unsere Gedanken einem Anderen ausdrücken. Man kann zwar sich eine Maschine in der Art denken, dass sie Worte äusserte, [...] aber niemals wird sie diese Worte so stellen können, dass sie auf das in ihrer Gegenwart Gesagte verständig antwortet, wie es doch selbst die stumpfsinnigsten Menschen vermögen. (Descartes 1870, S. 66)

Heute wissen wir: Descartes lag falsch. Während Maschinen zwar noch immer nicht alle unliebsamen Aufgaben des Menschen übernehmen können, sind die Fähigkeiten der durch maschinelles Lernen erzeugten großen Sprachmodelle wie GPT-4 so weit fortgeschritten, dass sie auf einen Input in natürlich menschlicher Sprache doch recht verständlich mit Worten und

Zeichen antworten können, mitunter eloquenter als so manch geistreicher Mensch. Das geht sogar so weit, dass der künstlich generierte Output dann nicht immer zweifelsfrei einem Menschen oder dem Programm zugeordnet werden kann.

Bevor wir uns der deutschdidaktischen Diskussion dieses Phänomens zuwenden, ist es wichtig zu klären, wovon eigentlich die Rede ist, wenn man von Künstlicher Intelligenz (KI) spricht, zumal unter den Begriff mehr fällt als die gegenwärtig diskutierten Sprachmodelle.

Das Konzept der Künstlichen Intelligenz und seine Benennung stammen ursprünglich vom englischen *Artificial Intelligence*, einer Wortschöpfung, die auf einen Antrag auf Forschungsförderung für die 1956 abgehaltene Dartmouth-Konferenz zurückgeht. Künstliche Intelligenz »diente als Sammelbegriff für alle Verhaltensweisen von Maschinen, die man als intelligent bezeichnen würde, wenn Menschen sie zeigen« (Deutscher Ethikrat 2023, S. 87). Seit damals entwickelte sich ein weitläufiges Forschungsfeld der Künstlichen Intelligenz, wobei auf viele Arten und Weisen versucht wurde, Maschinen Verhaltensweisen beizubringen, die als intelligent eingestuft werden. Unter den verschiedenen Verfahren des sogenannten Maschinellen Lernens, das sich als eigener Fachbereich herauskristallisierte, setzte man mit zunehmender Verfügbarkeit von großen Datenmengen und Rechenleistungen vor allem auf *deep learning*, eine Methode, die auf künstlichen neuronalen Netzwerken beruht. Als generative KI – zu der auch Sprachmodelle wie GPT oder LLaMA zählen

– bezeichnet man neuronale Netzwerke, die auf Basis eines enormen Datensatzes¹ lernen, zu generalisieren und damit zu generieren – also etwas hervorbringen, das im Datensatz in dieser Form nicht vorhanden war.² Das können Bilder, Videos, Musik, aber eben auch Texte sein. Im Fall von Wörtern und Text spricht man daher von *Large Language Models* (LLM) oder großen Sprachmodellen.

Wer bereits Erfahrungen mit großen Sprachmodellen gemacht hat, mag über die Qualität des generierten Outputs, die Angemessenheit der Antworten und die geschliffene Sprache erstaunt gewesen sein. Der Eindruck entsteht, als wäre mit der Entwicklung der sogenannten Transformer-Architektur für Sprachmodelle (Vaswani u. a. 2017) etwas gelungen, das lange Zeit für unmöglich gehalten wurde: Einen Computer in die Lage zu versetzen, die durch Komplexität, Polysemie und Idiosynkrasien gekennzeichnete menschliche Sprache so gut zu imitieren, dass es wirkt, als würde der Computer die Sprache beherrschen. Ob Sprachmodelle menschliche Sprache nun auch »verstehen« oder nur gut

täuschen, ist eine Diskussion, die weit zurückreicht: Bereits 1966 sorgte der erste Chatbot ELIZA für Aufsehen, der von Josef Weizenbaum (1966) entwickelt wurde, um den illusorischen Charakter von Computerintelligenz zu zeigen. Selbst das beschränkte sprachliche Repertoire löste bei Nutzer:innen den Eindruck von Intelligenz aus. Diese Illusion von Intelligenz nannte man daher ELIZA-Effekt (Natale 2021, S. 50–67). In der darauffolgenden Rezeptionsgeschichte wurde der Chatbot ELIZA als Beweis für zwei verschiedene konträre Sichtweisen interpretiert, die den Diskurs bis heute prägen: einerseits als Zeichen dafür, dass KI immer nur den Anschein von Intelligenz erzeugen kann, aber eben nicht intelligent ist, und andererseits, dass KI tatsächlich Intelligenz und Verstehen durch künstliche Mittel repliziert (ebd., S. 51). Während also KI-Skeptiker:innen auf die Beschränktheit und Tendenz zur Täuschung durch diese Systeme hinweisen und vor einem überzogenen Hype warnen (Marcus 2022; Bender u. a. 2021), erkennen Befürworter:innen gar Funken einer generellen Künstlichen Intelligenz (Bubeck u. a. 2023).

Nun stellt sich allerdings nicht nur die Frage, ob neuronale Netzwerke verstehen, sondern auch, ob Menschen überhaupt in der Lage sind, diese komplexen Strukturen zu verstehen³. Denn

1 Unternehmen wie OpenAI verraten wenig über den zugrunde liegenden Datensatz, es wird allerdings angenommen, dass der Datensatz, der für das Training von GPT-4 verwendet wurde, 40 Terabyte Text enthält. Das sind geschätzt 16–20 Billionen Wortteile (siehe Thompson 2024).

2 Obwohl es eine breite Diskussion darüber gibt, ob generative KI wirklich originell sein kann, ist festzuhalten, dass die Generierung nicht auf einem Kopieren von bereits vorhandenem Material beruht.

3 Wolfram spricht von »computationally irreducible processes« und folgert: »But at least as of now we don't have a way to ›give a narrative description‹ of what the network is doing« (Wolfram 2023).

der schier unvorstellbar große Datensatz und die daraus emergierende Komplexität lassen »wenig Aufschluss über ihre inneren Abläufe zu, sind weder durch äußere Beobachtung des Outputs noch selbst durch das Wissen um die Details des Modells in jene eindeutigen Ablaufpläne übersetzbar, die einen klassischen Programmcode auszeichnen« (Bajohr 2021, S. 177). Dieser Umstand und die rasante Geschwindigkeit der Entwicklung von Sprachmodellen (Roser 2024) machen es zum einen schwierig, definitive Aussagen über die emergierenden Fähigkeiten und deren Limitationen zu machen, und legen zum anderen nahe, die Irreduzibilität mit Visualisierungen (Bycroft 2024) oder Vergleichen einzuholen. Imaginiert werden Sprachmodelle oftmals als Raum, man spricht von der Architektur von Modellen: Die »Übersetzung« menschlicher Sprache in von Computern verarbeitbare Zahlen (Wörter werden zu *tokens* und schließlich durch Zahlen in Vektoren repräsentiert), das anschließende *embedding* – also die im Vektorraum repräsentierte Verteilung von Wörtern und ihre Nähe bzw. Verbindungen zueinander – erzeugen einen *meaning space* (Wolfram 2023), der sich eben als »Modell des menschlichen Sprachgebrauchs begreifen« (Müller/Fürstenberg 2023, S. 327) lässt.

Germanist:innen mögen bei dieser Beschreibung vielleicht an Roland Barthes' Aufsatz zum Tod des Autors denken. Tatsächlich trifft sein Konzept von Text die vermutete Funktionsweise von großen Sprachmodellen ganz gut: Text bestehe Barthes zufolge nämlich nicht aus einer Reihe von Wörtern, »sondern aus einem vieldimensio-

nalen Raum, in dem sich verschiedene Schreibweisen [*écritures*], von denen keine einzige originell ist, vereinigen und bekämpfen. Der Text ist ein Gewebe von Zitaten aus unzähligen Stätten der Kultur.« (Barthes 2000, S. 190) Der Autor ist für Barthes degradiert zum Schreiber [*scripteur*], der nicht registriert, konstatiert oder repräsentiert, sondern im Modus des Performativ schreibt, einer Verbalform also, die keinen anderen Inhalt hat »als eben den Akt, durch den sie sich hervorbringt« (ebd., S. 189). Tatsächlich lässt sich auch der nach mathematischen Prinzipien synthetisierte Text eines Sprachmodells als Performativ im Barthes'schen Sinne begreifen, eine Verbalform, die nicht repräsentiert, sondern selbstbezüglich generiert. Sprachmodelle können folglich als Phänomene der Algorithmizität gelten, also eine jener Formen, die nach Stalder konstitutiv für die Kultur der Digitalität sind: In algorithmisch vorsortierten Ordnungen wird die Welt »nicht mehr repräsentiert; sie wird für jeden User eigens generiert und anschließend präsentiert« (Stalder 2017, S. 189). Sprachmodelle sind daher – und das ist eine der wichtigsten Informationen – (noch) keine Wissensmodelle. Weder gibt es im Training eine Quelle für Wahrheit bzw. Faktizität noch kann sichergestellt werden, dass der generierte Output frei von Halluzinationen ist (siehe Hendler u. a. in diesem Heft).

Welche Vorstellungen von Sprachmodellen den Diskurs prägen und was man von generativer KI versteht, bestimmt in besonderer Weise den Umgang mit diesen Anwendungen: Ähnlich wie Barthes Folgerung, dass auf

den Tod des Autors die »Geburt des Lesers« (Barthes 2000, S. 193) folgt, ist vor allem im didaktischen Kontext zentral, was Menschen mit diesen Sprachmodellen *tun*. In den Blick geraten dann Praktiken des Umgangs mit dieser (Sprach-)Technologie und es zeigt sich eine immanent didaktische Dimension: Über welches Wissen sollten Menschen verfügen? Wie können sie die Technologie kompetent nutzen, aber auch bewusst nicht nutzen? Wie kann eine kritische, reflexive Auseinandersetzung mit den gegenwärtigen technologischen Trends angestoßen werden?

Eine zentrale Aufgabe von Bildungssystemen besteht darin, Menschen zu befähigen, verantwortungsbewusst mit Technik und Medien umgehen zu können. Wenn eine Technologie nun so fundamental mit Sprache verknüpft ist, folgt daraus, dass jene, die mit Vermittlung von Sprache in verschiedenen Domänen betraut sind, in besonderer Weise aufgefordert sind, über diese Technologie und ihren Einsatz nachzudenken.

Im Zusammenhang mit dem Deutschunterricht stellt sich daher eine Vielzahl von Fragen: Kann generative KI als disruptive Technologie gelten, die die Grundfesten des Deutschunterrichts erschüttert? Welche Domänen sind dabei besonders betroffen und welche Auswirkungen hat das auf sprachliche, literarische und mediale Bildung? Wie verändern derartige Sprachmodelle Textrezeption und -produktion gerade auch im schulischen Kontext? Welche Medienkompetenzen brauchen Lehrpersonen und welche müssen sie ihren Schüler*innen nunmehr vermitteln? Und

wie verändert sich die Lern- und Prüfungskultur, wenn nicht nachvollzogen werden kann, ob eine Aufgabe von Schüler*innen ganz selbstständig oder mithilfe von Künstlicher Intelligenz durchgeführt wurde? Im Lichte dieser Veränderungen gilt es nun auch fachdidaktisch zu reflektieren, welche Auswirkungen, Anwendungsfelder und Herausforderungen sich für den schulischen Deutschunterricht bzw. die Fachdidaktik selbst ergeben. Da derartige Programme auch eine schulpraktische Nutzung im Unterricht nahelegen bzw. in Schulen – wenn nicht verboten – schon verwendet werden, eröffnet sich auch ein anwendungsbezogenes Feld, das in diesem *ide*-Heft reflektiert werden soll.

Das Themenheft beschäftigt sich mit den Folgen der breiten Zugänglichkeit von Künstlicher Intelligenz für die Deutschdidaktik und den schulischen Deutschunterricht. Die Beiträge ergründen Implikationen für Literatur-, Medien- und Sprachdidaktik und fragen nach Konzeptionen und Good-Practice-Beispielen aus den verschiedenen Bereichen des Deutschunterrichts. Zur Diskussion stehen didaktische Konzepte und unterrichtspraktische Modelle, die angesichts von Sprachmodellen wie ChatGPT neu entwickelt oder anders perspektiviert werden müssen. Dabei werden zum einen Fragen der produktiven Nutzung bzw. anwendungsbezogene Szenarien erläutert. Zum anderen wird auch Raum für kritische Einschätzungen und Herausforderungen im Umgang mit Künstlicher Intelligenz gegeben.

Der erste Teil des Heftes gibt mit zwei einleitenden Artikeln einen Überblick

über bildungswissenschaftliche und deutschdidaktische Zugänge. *Claudia de Witt* fragt nach den pädagogischen Implikationen, die generative KI-Systeme im schulischen Kontext mit sich bringen, und erläutert verschiedene didaktische Szenarien von der Primarstufe bis zur Universität. *Stefan Kramer* und *Matthias Leichtfried* loten dann das transformative wie auch disruptive Potential aus, das Sprachmodelle wie ChatGPT für den Deutschunterricht haben. Ausgehend von einem Problemaufriss werden aktuelle deutschdidaktische Beiträge zu Fragen der Künstlichen Intelligenz systematisiert und kommentiert.

Der zweite Teil des Heftes versammelt Beiträge zu den unterschiedlichen Lernbereichen des Deutschunterrichts. Schreibdidaktischen Fragen widmen sich *Kirstin Schindler* sowie *Philipp Wampfler*. Während Erstere Prinzipien zum Schreiben mit, durch und von Künstlicher Intelligenz darlegt, mit Fokus auf Schreib- und Textkompetenzen, wie sie in der Sekundarstufe vermittelt werden, widmet sich Letzterer dem Verhältnis von Schriftlichkeit und Mündlichkeit im Zeichen der Digitalisierung, um daraus neue Möglichkeiten der Textproduktion abzuleiten. Dass textgenerierende Programme auch für die Lesedidaktik und das damit verbundene Einüben eines epistemisch wachsamem Lesens wichtig sind, verdeutlicht *Maik Philipp* in seinem Beitrag. *Carolin Führer* und *Daniel Nix* fragen hingegen danach, welche Rolle ChatGPT als Lektürebegleitung von literarischen Texten spielen kann, und zeigen in ihrem Beitrag auf, inwieweit die Antworten des

Programms Schüler:innen potentiell dazu befähigen, konstruktiv mit Literatur umzugehen. Sie greifen dabei auf Daten einer qualitativen Rezeptionsstudie zurück, in der Schüler*innen zu Franz Kafkas *Vor dem Gesetz* gearbeitet haben. Dass Motive von künstlichen Menschen und damit Künstlicher Intelligenz in der Literaturgeschichte bis in die Antike zurückgehen und seit damals eine Vielfalt von damit verbundenen Geschichten und Diskursen hervorgebracht haben, zeigt *Sebastian Kugler* auf. Die literarischen Beispiele, die er anführt, dienen auch als Anregung für die Auseinandersetzung im Literaturunterricht. Schließlich ergründet *Matthias Ballod* in seinem Beitrag die Rolle von textgenerierenden Sprachmodellen in Hinblick auf Sprachlichkeit und Sprachreflexion.

Der dritte Teil des Heftes gibt Einblicke in die unterrichtliche Praxis im Umgang mit Künstlicher Intelligenz. *Alexander Sigmund* diskutiert allgemeine Implikationen für den Deutschunterricht aus der Perspektive eines HAK-Lehrers und weist auf eine Vielzahl von Anwendungsmöglichkeiten hin. Dabei skizziert er notwendige Spielregeln im schulischen Umgang mit ChatGPT, die u.a. Recherche, Textproduktion und Beurteilung betreffen. *Melanie Hendler*, *Jennifer-Carmen Frey*, *Stephan Schicker* und *Sabine Schmölzer-Eibinger* beleuchten in ihrem Beitrag die von Philipp problematisierte Eigenart von Sprachmodellen, Unwahrheiten zu verbreiten, aus unterrichtspraktischer Perspektive und stellen Materialien vor, die Schüler:innen einen kompetenten Umgang mit Fake-

News und Falschnachrichten vermitteln sollen. *Martin Schastak* und *Simon Jung* nehmen schließlich Sprachbewusstsein und Mehrsprachigkeit in den Blick und liefern konkrete Unterrichtsideen im Zusammenhang mit KI-basierten Übersetzungstools. Zusätzliches Material dazu findet sich digital auf unserer Homepage: *ide.aau.at*.

Weiterführende Informationen liefert die Bibliographie und Link-Sammlung, die *Matthias Leichtfried* für dieses Heft zusammengestellt hat. Da viele Materialien und Beiträge zum Thema zunächst nur online erschienen sind, wird in der Bibliographie nicht nur relevante Literatur aus dem Forschungsfeld empfohlen, sondern es werden auch Links und Portale vorgestellt, die sich dem Thema der Künstlichen Intelligenz widmen. Ein Glossar bietet Erklärungen zu den im Heft verwendeten Hauptbegriffen zur Künstlichen Intelligenz.

Im Kommentar setzen sich *Ursula Esterl*, *Luca Melchior*, *Nicola Mitterer*, *Hajnalka Nagy*, *Johannes Odendahl* und *Jürgen Struger* kritisch mit der Anfang des Jahres verordneten Neustrukturierung des Lehramtsstudiums auseinander. In den Rezensionen stellen *Hans Lösener*, *Jürgen Struger* und *Ursula Esterl* aktuelle deutschdidaktische Publikationen vor.

Das Thema der Künstlichen Intelligenz ist – wie dieses Themenheft veranschaulicht – ein Anlass, um den Deutschunterricht und die damit verbundenen Praktiken und Deutungen noch einmal aus einer anderen Perspektive zu beleuchten und zu reflektieren. Wir möchten (zukünftige) Lehrpersonen, Fachdidaktiker*innen und

Germanist*innen dazu einladen, sich auf diesen Prozess einzulassen und die Implementierung wie auch Verwendung von KI-Systemen in Lehr- und Lernkontexten selbst produktiv mitzugestalten. Um hier professionell, umsichtig, mit gebotener kritischer Distanz, aber auch mit Offenheit und Neugierde handeln zu können, braucht es fachliches Wissen und Kompetenz im Umgang mit derartigen Technologien. Es ist aus der gegenwärtigen Perspektive sehr wahrscheinlich, dass die technologische Weiterentwicklung der kommenden Jahre mit hoher Geschwindigkeit und starkem Einfluss auf unseren Alltag verlaufen wird. Eine kritische und ausführliche Beschäftigung damit kann also gar nicht früh genug beginnen.

Das Heft möchte dazu einen Beitrag leisten, indem es Fehlvorstellungen und Mythen rund um Künstliche Intelligenz entgegenwirkt, didaktische Möglichkeiten und Limitationen aufzeigt, Denkprozesse anstößt und mit Vorschlägen zur Mitgestaltung anregt.

Wir wünschen eine informative Lektüre.

STEFAN KRAMMER
MATTHIAS LEICHTFRIED

Literatur

- BAJOHR, HANNES (2021): Künstliche Intelligenz und digitale Literatur. Theorie und Praxis konnektionistischen Schreibens. In: Bajohr, Hannes; Gilbert, Annette (Hg.): *Digitale Literatur II*. München: edition text + kritik (Sonderband), S. 174–185.
- BARTHES, ROLAND (2000): Der Tod des Autors. In: Jannidis, Fotis; Lauer, Gerhard; Martínez, Matías; Winko, Simone (Hg.): *Texte zur Theorie der Autorschaft*. Stuttgart: Reclam, S. 185–193.
- BENDER, EMILY M.; GEBRU, TIMNIT; McMILLAN-MAJOR, ANGELINA; SHMITCHELL, SHMARGARET (2021): On the Dangers of Stochastic Parrots. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency. Virtual Event Canada, 03 03 2021–10 03 2021. New York, NY, United States: Association for Computing Machinery (ACM Digital Library), S. 610–623.
- BUBECK, SÉBASTIEN; CHANDRASEKARAN, VARUN; ELKAN, RONEN; GEHRKE, JOHANNES; HORVITZ, ERIC; KAMAR, ECE u. a. (2023): *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. Online: <http://arxiv.org/pdf/2303.12712.pdf> [Zugriff: 21.5.2024].
- BYCROFT, BRENDAN (2024): *LLM Visualization*. Online: https://bbycroft.net/llm?mj_campaign=nl_ref&mj_content=zeitde_text_link_x&mj_medium=nl&mj_source=int_zonaudev_Nat%C3%BCrlich%20intelligent [Zugriff: 21.5.2024].
- DESCARTES, RENÉ (1870): *Abhandlung über die Methode, richtig zu denken und Wahrheit in den Wissenschaften zu suchen*. Berlin: Heimann. Online: <https://www.digitale-sammlungen.de/de/details/bsb11163821> [Zugriff: 21.5.2024].
- DEUTSCHER ETHIKRAT (2023): *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Stellungnahme*. Berlin.
- MARCUS, GARY (2022): *Deep Learning Is Hitting a Wall*. Online: <https://nautilus.us/deep-learning-is-hitting-a-wall-238440> [Zugriff: 21.5.2024].
- MÜLLER, HANS-GEORG; FÜRSTENBERG, MAURICE (2023): Der Sprachgebrauchsautomat. Die Funktionsweise von GPT und ihre Folgen für Germanistik und Deutschdidaktik. In: *Mitteilungen des Deutschen Germanistenverbandes* 70, H. 4, S. 327–345.
- NATALE, SIMONE (2021): *Deceitful media. Artificial intelligence and social life after the turing test*. New York: Oxford University Press.
- ROSER, MAX (2024): The brief history of artificial intelligence: The world has changed fast – what might be next? In: *Our World in Data*. Online: <https://ourworldindata.org/brief-history-of-ai> [Zugriff: 21.5.2024].
- STALDER, FELIX (³2017): *Kultur der Digitalität*. Berlin: Suhrkamp (= Edition Suhrkamp, 2679).
- THOMPSON, ALAN D. (2024): GPT-4. Online: <https://lifearchitected.ai/gpt-4/#summary>, [Zugriff: 21.5.2024].
- VASWANI, ASHISH; SHAZEER, NOAM; PARMAR, NIKI; USZKOREIT, JAKOB; JONES, LLION; GOMEZ, AIDAN N. u. a. (2017): *Attention Is All You Need*, S. 1–15. Online: <https://arxiv.org/pdf/1706.03762> [Zugriff: 21.5.2024].
- WEIZENBAUM, JOSEPH (1966): ELIZA – a computer program for the study of natural language communication between man and machine. In: *Commun. ACM* 9 (1), S. 36–45. DOI: 10.1145/365153.365168.
- WOLFRAM, STEPHEN (2023): *What Is ChatGPT Doing ... and Why Does It Work?* Online: <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> [Zugriff: 21.5.2024].
-
- STEFAN KRAMMER ist Professor für Neuere deutsche Literatur und ihre Didaktik an der Universität Wien. Seine Arbeitsschwerpunkte liegen im Bereich der Literatur- und Medien-didaktik. E-Mail: stefan.krammer@univie.ac.at
- MATTHIAS LEICHTFRIED ist Postdoc-Universitätsassistent am Institut für Germanistik in Wien und lehrt bzw. forscht im Bereich der Fachdidaktik Deutsch. Seine Forschungsinteressen liegen in der Literatur- und Medien-didaktik, Forschungsschwerpunkte sind der Deutschunterricht in einer Kultur der Digitalität, Ästhetische Bildung sowie Künstliche Intelligenz.
E-Mail: matthias.leichtfried@univie.ac.at