

Generative Sprachmodelle als Herausforderung für erziehungswissenschaftliche Vorstellungen von Lernen und Sozialisation?

Christoph Richter & Heidrun Allert

Abstract

Generative Language Models as a Challenge for Pedagogical Concepts of Learning and Socialization?

Generative language models have opened up a wide range of previously unimaginable possibilities for text production and use in both everyday life and in educational contexts. Addressing generative language models as tools to be harnessed, recent educational discourse tends to overlook the ways in which these models reshape individual and collective means of experience and articulation. This article argues that, due to their underlying operational logic, generative language models are necessarily geared to reproduce existing social and cultural orders. Their epistemic conformism aligns them with affirmative practice regimes and makes it difficult to develop idiosyncratic and marginalized perspectives. Rather than being regarded as tools, they should instead be understood as objects of collective learning themselves.

Jenseits einer Kritik von außen

Auch wenn komplexe algorithmische Systeme schon seit geraumer Zeit unseren Alltag durchdringen, eröffnen generative Sprachmodelle eine Vielzahl neuer, bisher kaum vorstellbarer Anwendungen. Diese ermöglichen es, sowohl mit Chatbots, virtuellen Assistenten und Suchmaschinen in natürlicher Sprache zu interagieren als auch Texte etwa in Form von Essays, Geschichten, Tutorials oder Softwarecode zu erzeugen, zu übersetzen, zu korrigieren, zu analysieren und vieles mehr. Im Bildungskontext wird bereits intensiv an Anwendungen gearbeitet, die diagnostische und tutorielle Funktionen übernehmen, Arbeitsergebnisse bewerten, Lernenden formatives Feedback geben und Lernmaterialien erstellen oder kuratieren.

Mit ihrer Fähigkeit, natürlichsprachliche Texte zu verarbeiten wie auch zu erzeugen, haben generative Sprachmodelle die Möglichkeiten der Produktion und des Umgangs mit Texten innerhalb kürzester Zeit grundlegend verändert und sind zu einem neuen „digitalen Leitmedium“ (Döbeli Honegger 2024) geworden. Als Teil jener „Kulturgestalt“ (Loch 1968), in die Menschen hineinwachsen und in der sie agieren, wirkt sich die Verbreitung generativer Sprachmodelle notwendiger-

weise auch auf Lern- und Sozialisationsprozesse aus. Ob sich vor diesem Hintergrund aber auch die erziehungswissenschaftlichen Vorstellungen von Lernen und Sozialisation ändern müssen, hängt davon ab, wie wir das Verhältnis von generativen Sprachmodellen und ihrer Verwendung verstehen.

In der noch jungen Auseinandersetzung mit generativen Sprachmodellen sind im erziehungswissenschaftlichen Diskurs bislang vor allem Stimmen laut geworden, die diese Technologien als Werkzeuge begreifen, die es kompetent zu nutzen bzw. in sozial verantwortlicher Weise zu kontrollieren und auszugestalten gelte (z. B. Kasneci et al. 2023). Auch wenn sich an einigen Stellen andeutet, dass sich mit der Verfügbarkeit generativer Sprachmodelle auch Fragen hinsichtlich tradierter Prüfungsformen (z. B. Ständige wissenschaftliche Kommission der Kultusministerkonferenz 2024) oder des Erwerbs bisher als unabdingbar erachteter Kompetenzen (z. B. Döbeli Honegger 2024) stellen, so liegt auch hier ein instrumentelles Technikverständnis zugrunde. In der Herstellung und Nutzung der verschiedenen Anwendungen manifestieren sich zwar individuelle und kollektive Werthaltungen und Interessen, die Technik als solche wird diesen gegenüber aber als neutral konzipiert.

In einer derartigen Betrachtung bleibt jedoch „die Quelle der Kritik der Technik rein äußerlich“ (Matzner 2024, S. 140). Sie übersieht sowohl die mediale Wirkkraft der Technik, „welche die Modalitäten unseres Denkens, Wahrnehmens, Erfahrens, Erinnerns und Kommunizierens prägt“ (Krämer 1998, S. 14), als auch den Eigensinn der technischen Verfahren und die Leistungen, die erbracht werden müssen, um die Technik in praktische Tätigkeitszusammenhänge zu integrieren (vgl. Richter/Allert 2024). Vor diesem Hintergrund lenken wir im Folgenden zunächst den Blick auf die technische Logik generativer Sprachmodelle und die für ihre Nutzung notwendigen „Regime der Praxis“ (Introna 2016), um dann die Implikationen für Lern- und Sozialisationsprozesse zu skizzieren.

Die technische Logik und die Regime der Praxis

Ungeachtet ihrer hohen technischen Komplexität, die letztlich auf der geschickten Kombination verschiedener, zum Teil sehr ressourcenintensiver Algorithmen gründet (einführend z. B. Linde 2023), und losgelöst von ihrer möglichen Integration in andere technische Systeme wie etwa Suchmaschinen oder regelbasierten KI-Systeme basieren die derzeit verfügbaren generativen Sprachmodelle auf einer einheitlichen und vergleichsweise einfachen technischen Logik. Sie beruhen im Kern auf der Idee, in extrem große Mengen an Texten mithilfe statistischer Verfahren nach sprachlichen Mustern und Beziehungen zu suchen und eben diese sowohl zur Analyse als auch zur Erstellung neuer Texte zu verwenden. Für die Produktion von Texten analysieren generative Sprachmodelle mittels statischer Verfahren den „Kontext“ des jeweiligen Satzes und bestimmen dann anhand von Wahrscheinlichkeiten das nächste Wort (vgl. Simon/Spie-

cker genannt Döhmman/von Luxburg 2024). Worin genau die jeweiligen Muster und Beziehungen bestehen und was entsprechende Modelle im Rahmen ihres „Trainings“ tatsächlich „lernen“, ist dabei aus technischer Sicht unerheblich und kann aufgrund der eingesetzten Verfahren nicht unmittelbar nachvollzogen werden (vgl. Burrell 2016; Boge 2022). Das aus technischer Sicht relevante Kriterium ist die Erwartungskonformität der Ergebnisse.

Ohne weiter in die technischen Details einsteigen zu müssen, hat diese „Logik“ generativer Sprachmodelle zwei wesentliche Implikationen. Zum einen sind generative Sprachmodelle – wie andere Verfahren des Maschinellen Lernens auch – darauf ausgelegt, jene Informationen als relevant zu behandeln, die im Trainingsmaterial statistisch dominant sind. Der hiermit einhergehende „erkenntnistheoretische Konformismus“ (Miragoli 2024) bezieht sich dabei nicht nur auf die syntaktische Struktur der Sprache, sondern auch auf die im Trainingsmaterial eingelagerten Bedeutungsstrukturen in Form sozial geteilter Klischees (Durt/Froese/Fuchs 2023; Bajohr 2023). Zum anderen verfügen generative Sprachmodelle über keinen Mechanismus, der es ihnen erlauben würde, sich in reflexiver Weise mit den von ihnen produzierten Texten auseinanderzusetzen, und sind daher nicht in der Lage, den Wahrheitsgehalt ihrer Aussagen zu prüfen (Krämer 2024).

Wenn nun aber die Produkte generativer Sprachmodelle in einem praktischen Tätigkeitszusammenhang als sinnhaft und bedeutsam erachtet werden, so bedarf es hierfür eines entsprechenden Regimes der Praxis. Ein solches Regime umfasst neben der Erschließung notwendiger Kompetenzen und Ressourcen nicht zuletzt auch einen Abgleich der kollektiven Erwartungshorizonte und Werthaltungen. Im Rahmen des unmittelbaren Miteinandertuns wie auch in Form expliziter Aushandlungen ist unter anderem zu bestimmen, worauf es bei der Produktion und im Umgang mit Texten eigentlich ankommt, was die Qualität eines Textes eines bestimmten Genres ausmacht, wie verlässlich die enthaltenen Informationen sein müssen und welche Verantwortlichkeiten hiermit einhergehen.

In Anbetracht des in generativen Sprachmodellen angelegten epistemischen Konformismus und ihrer kommunikativen Verständnislosigkeit sind die jeweils zu erbringenden Leistungen jedoch ungleich verteilt. Aufgrund ihrer technischen Logik reduzieren sie den Aufwand für jene, die an den jeweils statistisch dominanten Bedeutungsstrukturen und Interpretationsmustern anknüpfen. Umgekehrt erhöhen sie den Aufwand, wenn eben jene Bedeutungsstrukturen und Interpretationsmuster zur Disposition stehen. Dies gilt sowohl, wenn Praktiker*innen nach Informationen, Sinnzusammenhängen oder Perspektiven suchen, die sich den statistisch dominanten Strukturen entziehen, als auch dann, wenn sie generative Sprachmodelle verwenden wollen, um spezifische, singuläre, aber statistisch marginale Erfahrungen und Ideen zu artikulieren (vgl. Miragoli 2024).

Lernen und Sozialisation zwischen Affirmation und Transgression

Folgt man dem hier vorgeschlagenen analytischen Ansatz, so ist die Frage nach einem veränderten Verständnis von Lernen und Sozialisation nicht zu trennen von der praktischen Bedeutung und Funktion, die den statistisch dominanten Bedeutungsstrukturen und Deutungsmustern in den jeweiligen Tätigkeitszusammenhängen beigemessen wird. Wird das statistisch Dominante als das praktisch Relevante begriffen oder wird auch dem Marginalen und Singulären eine Bedeutung beigemessen? Dient das statistisch Dominante als affirmativer Zielhorizont für individuelle Lern- und Sozialisationsprozesse oder sind es gerade diese in den Sprachmodellen sedimentierten Strukturen und Muster, die es im Rahmen von Lern- und Sozialisationsprozessen immer wieder herauszufordern, zu transformieren und zu unterlaufen gilt?

Entsprechend hängt etwa die praktische Bedeutung generativer Sprachmodelle bei der Erstellung von Texten entscheidend davon ab, welche Rolle und Funktion dem Schreiben in einem bestimmten soziokulturellen Kontext zugesprochen wird. So kann eine KI-basierte Schreibassistenz dabei unterstützen, einen Text zu verfassen, der gewissen sozialen Konventionen entspricht, beispielsweise indem orthographische und grammatikalische Fehler identifiziert werden oder der Text einem vordefinierten Sprachstil angepasst wird. Sie hilft aber nicht weiter, wenn das Ziel des Schreibens darin besteht, eine Erfahrung oder Erkenntnis zu artikulieren, die sich abseits der etablierten Konventionen und Sprachmuster bewegt und nicht in den gängigen Worten, Formulierungen und Bildern aufgeht. Die Entwicklung einer eigenen, widerständigen Sprache und damit einhergehender Formen der individuellen wie auch kollektiven Welterschließung ist hier letztlich nur in grundsätzlicher Abgrenzung zu den Vorschlägen der KI möglich.

Ähnlich verhält es sich mit der Verwendung generativer Sprachmodelle als Informationsquelle. Solange das Ziel darin besteht, geläufige (weil statistisch dominante) Muster und Interpretationen zu identifizieren, erscheinen generative Sprachmodelle hierfür als ein technisch probates Mittel. Dies gilt insbesondere dann, wenn davon ausgegangen werden kann, dass die entsprechenden Themen, Textgattungen und Fragestellungen im jeweils zugrunde liegenden Trainingsdatensatz hinreichend umfassend abgebildet sind. Je spezifischer und idiosynkratischer jedoch das epistemische Anliegen ist, je weiter sich die Lern- und Erkenntnisprozesse also von jenen statistisch dominanten Mustern und Interpretationen entfernen, desto weniger hilfreich werden die Produkte der generativen Sprachmodelle (auch wenn diese nicht müde werden, uns Antworten auf unsere Prompts zu generieren).

In den jeweiligen Regimen der Praxis ist damit auch angelegt, ob wir Lernen und Sozialisation als individuelle Aneignungsprozesse verstehen, die sich vor einem mehr oder weniger statischen soziokulturellen Hintergrund vollziehen

oder ob Lernen und Sozialisation vielmehr genau jene kollektiven und transgressiven Prozesse markieren, in denen unsere Kenntnisse, unsere Fertigkeiten und unsere Identitäten zuallererst ihre Form erlangen. In Abhängigkeit der Regime der Praxis treten generative Sprachmodelle dann entweder als nützliche Werkzeuge in Erscheinung, derer wir uns in zielführender Weise bedienen, oder aber als wirkmächtige Medien, die darüber mitbestimmen, was für wen artikuliert und wahrnehmbar wird. Die Konzeption generativer Sprachmodelle als Werkzeug negiert allerdings grundsätzlich, dass mit ihrer Verwendung bestimmte Strukturen, Verhältnisse und Ausdrücke reproduziert, ausgebaut und stabilisiert, somit also performativ werden. Es bleibt folglich eine entscheidende Frage, für wen, welche Identitäten und Tätigkeitszusammenhänge generative Sprachmodelle als Werkzeuge in Erscheinung treten und für welche nicht.

Generative Sprachmodelle als öffentliches Problem und kollektiver Lerngegenstand

Der in diesem Beitrag skizzierte Versuch einer Kritik generativer Sprachmodelle „von innen“, geht über die aktuelle Diskussion um Verzerrungen, Halluzinationen und toxische Sprachmuster hinaus und richtet den Fokus auf die Frage nach unserem kollektiven Umgang mit der Reproduktion sozial geteilter Klischees durch generative Sprachmodelle. Diese Klischees, so die Argumentation, sind eben keine Verzerrungen, die sich durch eine Anpassung des Trainingsmaterials oder ähnliche Vorkehrungen beheben ließen, sondern ein integraler Bestandteil der technischen Logik, die umso schwerer zu erkennen ist, je tiefer sich die entsprechenden Klischees in unser gemeinsames kulturelles Archiv eingeschrieben haben (vgl. Durt/Froese/Fuchs 2023).

Auch wenn eine solche Kritik nicht außerhalb der Regime der Praxis formuliert werden kann, so zeigt sie doch, dass die im aktuellen Diskurs vorherrschende Vorstellung generativer Sprachmodelle als einem zweckdienlichen Werkzeug nicht nur zu kurz greift, sondern selbst Teil einer hegemonialen Praxis ist, die darauf beruht, die in den Sprachmodellen sedimentierten Klischees als unproblematisch zu rahmen. Die auch in der erziehungswissenschaftlichen Diskussion immer wieder anzutreffende Formulierung eines Werkzeugcharakters sowie die Idee, es käme lediglich darauf an, die in den Sprachmodellen enthaltenen Verzerrungen zu identifizieren und gegebenenfalls zu beheben, übersieht, dass die jeweils statistisch dominanten Muster und Strukturen nie neutral sind, sondern immer schon Sedimentierungen bestehender sozialer und kultureller Ordnungen darstellen. Dies auszublenden bedeutet letztlich nichts anderes, als die mit diesen Ordnungen einhergehenden Formen epistemischer Gewalt ebenso auszuklamern wie die politische Dimension des Lernens und der Sozialisation. Generative Sprachmodelle als Werkzeuge zu betrachten, ist nur denjenigen möglich, die sich selbst in den jeweiligen Klischees wiederfinden.

Die in diesem Beitrag eingenommene Position löst sich damit letztlich auch von der Vorstellung, die Prozesse des Lernens und der Sozialisation seien gelöst von den jeweiligen Technologien und Regimen der Praxis zu begreifen. Lernen und Sozialisation sind hingegen ein integraler Bestandteil der sozialen Praxis selbst, sie sind Teil jenes kollektiven Ringens um den Sinn des gemeinsamen Miteinandertuns. In diesem Sinne sind generative Sprachmodelle weniger Werkzeuge als vielmehr ein „öffentliches Problem“ (Ananny 2024), ein kollektiver Lerngegenstand, der uns dazu auffordert, die komplexen Relationierungen von Wissen, Kultur, Sprache und Technik immer wieder neu auszuloten.

Christoph Richter, Dr., ist Wissenschaftlicher Mitarbeiter an der Abteilung für Medienpädagogik/Bildungsinformatik am Institut für Pädagogik der Christian-Albrechts-Universität zu Kiel.

Heidrun Allert, Prof. Dr., ist Professorin der Pädagogik, Schwerpunkt Medienpädagogik/Bildungsinformatik am Institut für Pädagogik der Christian-Albrechts-Universität zu Kiel.

Literatur

- Ananny, Mike (2024): Make AI a Public Problem. In: *Issues in Science and Technology* 40, 2, S. 88. <https://doi.org/10.58875/EHNY5426>.
- Bajohr, Hannes (2023): Dumb Meaning: Machine Learning and Artificial Semantics. In: *IMAGE* 37,1, S. 58-70. <https://doi.org/10.1453/1614-0885-1-2023-15452>.
- Boge, Florian J. (2022): Two Dimensions of Opacity and the Deep Learning Predicament. In: *Minds and Machines* 32,1, S. 43-75. <https://doi.org/10.1007/s11023-021-09569-4>.
- Burrell, Jenna (2016): How the Machine ‚Thinks‘: Understanding Opacity in Machine Learning Algorithms. In: *Big Data & Society* 3, 1. <https://doi.org/10.1177/2053951715622512>.
- Döbeli Honegger, Beat (2024): Was will uns ChatGPT sagen? In: *Pädagogik* 76, 3, S. 34-39. <https://doi.org/10.3262/PAED2403034>.
- Durt, Christoph/Froese, Tom/Fuchs, Thomas (2023): Large Language Models and the Patterns of Human Language Use: An Alternative View of the Relation of AI to Understanding and Sentience. [Preprint] <https://philsci-archive.pitt.edu/22744/> [Zugriff: 19. Februar 2025].
- Introna, Lucas D. (2016): Algorithms, Governance, and Governmentality: On Governing Academic Writing. Science. In: *Technology, & Human Values*, 41, 1, S. 17-49. <https://doi.org/10.1177/0162243915587360>.

- Kasneci, Enkelejda et al. (2023): ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. In: *Learning and Individual Differences* 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Krämer, Sybille (1998): Was haben die Medien, der Computer und die Realität miteinander zu tun? In: Krämer, Sybille (Hrsg.): *Medien Computer Realität – Wirklichkeitsvorstellungen und Neue Medien*. Frankfurt am Main: Suhrkamp, S. 9-26.
- Krämer, Sybille (2024): Die Nicht-Vernunft der Chatbots – Was macht auf Large Language Models beruhende Künstliche Intelligenz philosophisch interessant? In: Adolphi, Rainer/Alpsancar, Suzana/Hahn, Susanne/Kettner, Matthias (Hrsg.): *Philosophische Digitalisierungsforschung: Verantwortung, Verständigung, Vernunft, Macht*. Bielefeld: transcript, S. 297-314. <https://doi.org/10.1515/9783839474976-011>.
- Linde, Helmut (2023): So funktioniert ChatGPT. Golem.de. <https://www.golem.de/news/kuenstliche-intelligenz-so-funktioniert-chatgpt-2302-171644.html> [Zugriff: 19. Februar 2025].
- Loch, Werner (1968): Enkulturation als anthropologischer Grundbegriff der Pädagogik. In: *Bildung und Erziehung* 21, 3, S. 161-178. <https://doi.org/10.7788/bue-1968-jg19>.
- Matzner, Tobias (2024): Algorithmenkritik und die Suche nach dem „Außen“. In: Adolphi, Rainer/Alpsancar, Suzana/Hahn, Susanne/Kettner, Matthias (Hrsg.): *Philosophische Digitalisierungsforschung: Verantwortung, Verständigung, Vernunft, Macht*. Bielefeld: transcript, S. 137-151. <https://doi.org/10.1515/9783839474976-005>.
- Miragoli, Martin (2024): Conformism, Ignorance & Injustice: AI as a Tool of Epistemic Oppression. In: *Episteme, First View*, S. 1-19. <https://doi.org/10.1017/epi.2024.11>.
- Richter, Christoph/Allert, Heidrun (2024): Für eine „technologiebewusste Medienpädagogik“ jenseits der Digitalisierung. Ein Weg in die Archive der Technizität. In: *Jahrbuch Medienpädagogik*, 21, S. 323-350. <https://doi.org/10.21240/mpaed/jb21/2024.09.13.X>.
- Simon, Judith/Spiecker genannt Döhmman, Indra/von Luxburg, Ulrike (2024): Generative KI – jenseits von Euphorie und einfachen Lösungen (= Diskussion, 34). Halle an der Saale: National Akademie der Wissenschaften Leopoldina. https://doi.org/10.26164/LEOPOLDINA_03_01226.
- Ständige wissenschaftliche Kommission der Kultusministerkonferenz (2024): Large Language Models und ihre Potenziale im Bildungssystem. Impulspapier der Ständigen Wissenschaftlichen Kommission (SWK) der Kultusministerkonferenz. Bonn: SWK. <http://dx.doi.org/10.25656/01:28303>.