

Working Paper: Sprachmodelle in der Lehre an der Technischen Fakultät der Universität Bielefeld

Beteiligte: Hanna Drimalla, Stefan Kopp, Markus Nebel, Benjamin Paaßen, Helge Ritter, Alexander Schönhuth

Einleitung

ChatGPT, Bard, Bing, Github Copilot und ähnliche Werkzeuge basieren auf *Sprachmodellen*. Sprachmodelle werden trainiert, für einen Eingabetext das nächste Token vorherzusagen. Wichtig ist: Die Modelle sind *nicht* dahingehend trainiert, Wissen zu verarbeiten oder faktisch korrekte Aussagen zu erzeugen. Alles, was wir an Bedeutung in den Ausgaben von Sprachmodellen erkennen, ist von uns hineininterpretiert (vgl. Bender *et al.*, 2021). Trotzdem ist das Potenzial von Sprachmodellen offensichtlich, Texte effizienter und effektiver zu erzeugen. Unsere Mission in der Lehre sollte es sein, Studierenden die nötigen Kompetenzen zu vermitteln, um Sprachmodelle als Werkzeug gebrauchen zu können, ohne selbst die Kontrolle über oder die Verantwortung für den erzeugten Text zu verlieren. Die nötigen Kompetenzen beziehen sich

1. auf das Fach, d.h. Studierende müssen immer noch inhaltlich bewerten können, ob der erzeugte Text fachlich korrekt und die Argumentation schlüssig ist,
2. auf die informatischen Grundlagen der Sprachmodelle, insbesondere *natural language processing* und Transformer-Architekturen und
3. auf die Anwendung der Sprachmodelle, d.h. Studierende müssen lernen, welche Teile von Arbeitsprozessen sich für die Unterstützung durch Sprachmodelle eignen und wie Aufgaben für Sprachmodelle zu stellen sind ("*prompt engineering*").

Insgesamt sollte es das Ziel sein, die Lehre an der technischen Fakultät "Sprachmodellgerecht" weiterzuentwickeln. Dies kann ein Qualitätsmerkmal der Bielefelder Lehre werden. Die Entwicklung der Sprachmodelle bleibt dabei hochgradig dynamisch. Dementsprechend muss sich die Lehre auch weiterhin an veränderte Möglichkeiten und Technologien anpassen.



Sprachmodelle als Schreibwerkzeuge und gute wissenschaftliche Praxis

Wir empfehlen Lehrenden den Einsatz von Sprachmodellen als Schreibwerkzeug grundsätzlich zu gestatten, wenn der Einsatz mit guter wissenschaftlicher Praxis vereinbar ist, das heißt insbesondere (vgl. Justizariat, Barber 2023; sowie Salden und Leschke, 2023):

- Studierende müssen anzeigen, ob Sprachmodelle als Hilfsmittel verwendet wurden und sollen Auskunft über Umfang und Art der Verwendung geben (z.B. über eine differenzierte Beschreibung mit Beispielprompts oder einen Link zum Gesprächsverlauf). Standardformulierungen zu diesem Zweck können den Studierenden etwa durch die Fakultät oder die Lehrenden bereitgestellt werden.
- Studierende müssen weiterhin die alleinige Verantwortung für die Arbeit übernehmen. Das schließt insbesondere ein: Die Korrektheit und Vollständigkeit von Zitaten (es dürfen sich also keine unbeabsichtigten Plagiate einschleichen) sowie die faktische Richtigkeit und argumentative Schlüssigkeit des Textes.
- Studierende dürfen aus Datenschutzgründen grundsätzlich keine persönlichen Daten Dritter (etwa von Versuchspersonen) in Sprachmodell-Interfaces eingeben.

Zusätzlich empfehlen wir den Lehrenden, die die Nutzung von Sprachmodellen gestatten wollen, Studierenden grundsätzliches Wissen und Kenntnisse für den Einsatz zu vermitteln, insbesondere:

- dass Sprachmodelle keine Garantie für faktische Richtigkeit bieten und insbesondere generierte Zitate frei erfunden sein können. Daher ist die inhaltliche, faktische Richtigkeit des Textes stets zu prüfen.
- dass es empfehlenswert ist, von Sprachmodellen generierte Texte nicht direkt zu übernehmen, um eine eigene Leistung sicher zu stellen und unbeabsichtigte Plagiate aus den Trainingsdaten zu vermeiden. Zwar verletzen von ChatGPT erzeugte Texte nicht per se das Urheberrecht – es wird aber dann kritisch, wenn von ChatGPT erzeugte Texte (unabsichtlich) Textabschnitten aus den Trainingsdaten erheblich ähneln.
- dass sich die Gefahr faktischer Fehler reduzieren lässt, wenn eigener Text als Grundlage bereitgestellt wird, z.B. Stichpunkte, aus denen das Sprachmodell dann einen Text formulieren soll.



- dass die Ausgabe von Sprachmodellen stark davon abhängt, wie die Aufgabe gestellt wird (z.B. wie viel Kontext bereitgestellt wird, welche Rolle dem Sprachmodell zugewiesen wird)
- dass sich die Qualität der Antwort eines Sprachmodells erhöhen kann, wenn eine umfassende Aufgabe in kleinere Teilaufgaben zerlegt wird ("*chaining*")
- dass insbesondere Literatursuche noch unzuverlässig funktioniert und daher bislang noch mit anderen Werkzeugen geschehen sollte und
- dass das Verhalten von Sprachmodellen insgesamt stark stochastisch und daher schwer vorhersagbar sind. Das Erstellen von *prompts* ist eher eine Kunst als eine Wissenschaft.

Wir raten ausdrücklich davon ab, automatische Werkzeuge zur Detektion von Sprachmodellen einzusetzen. Nach aktuellem Stand der Wissenschaft sind diese Detektoren für fortgeschrittene Sprachmodelle (insb. ChatGPT) nicht zuverlässig genug und würden zu einer hohen Rate falscher Verdächtigungen führen. Dennoch kann es sein, dass Lehrende aufgrund Diskrepanz zwischen im Kurs und im Text präsentierten Wissens oder aufgrund "Sprachmodelltypischer" Schreibweise das Vorliegen eines ungekennzeichneten Einsatzes vermuten. Für solche Fälle bedarf es einer sensiblen Überprüfungsstrategie, die zwischen Lehrenden und mit Studierenden abzustimmen ist (etwa "klassische" Möglichkeiten, wie das direkte Gespräch mit Studierenden).

Der Technischen Fakultät als Institution empfehlen wir:

- Kurse zum wissenschaftlichen Schreiben/guter wissenschaftlicher Praxis anzubieten (idealerweise als Pflichtkurs und idealerweise bereits auf Bachelorniveau) und Sprachmodelle als einen Teil dieses Kurses zu verankern
- im fortgeschrittenen Bachelorniveau oder Masterniveau dedizierte und tiefgehende Lehrinhalte zu Sprachmodellen anzubieten und die Lehre diesbezüglich zu koordinieren
- eigene Instanzen von (quelloffenen) Sprachmodellen bereit zu stellen, um eine Abhängigkeit von proprietären Anbietern zu vermeiden und
- sich weiterhin der Dynamik im Bereich Sprachmodelle zu stellen.

Prüfungsformen

Für Prüfungen kann es erforderlich sein, die Nutzung von ChatGPT und anderen Hilfsmitteln zu verbieten, um die eigenständige Leistung der Studierenden sicher zu stellen. Zum jetzigen Zeitpunkt können wir dafür ausschließlich folgende Prüfungsformen empfehlen:

- Präsenz-Klausuren mit Papier & Stift oder mit standardisierten, von Lehrenden vorbereiteten Devices (z.B. *smart pens* mit vorbereiteten Tablets)
- mündliche Prüfungen (virtuell oder in Präsenz) und
- Präsentationen bzw. "Vorrechnen" (virtuell oder in Präsenz)

Alternativ sind *open book*-Klausuren möglich, bei denen der Einsatz von Sprachmodellen gestattet wird (unter Berücksichtigung guter wissenschaftlicher Praxis, siehe oben).

Programmierübungen

In der Technischen Fakultät spielen Programmierübungen eine zentrale Rolle. Diese Aufgaben sind durch Sprachmodelle besonders betroffen, weil viele moderne Sprachmodelle (insb. Github Copilot) spezifisch für die Generierung von Code trainiert wurden und die meisten Standard- Aufgaben in den Trainingsdaten enthalten sind. Daher ist es bei Programmierübungen besonders einfach, vollständige, richtige Antworten für typische Aufgabenstellungen zu generieren. Eine Eigenleistung der Studierenden besteht dann nur noch sehr eingeschränkt (im Gegensatz etwa zur Benutzung als Schreibwerkzeug, siehe oben).

Zum jetzigen Zeitpunkt haben sich noch keine abschließenden *best practices* für den Umgang von ChatGPT und Sprachmodellen für Programmieraufgaben etabliert. Wir können daher nur eingeschränkt Empfehlungen geben, etwa:

- Studierende dazu zu motivieren, selbst die Kompetenzen zu erwerben, um korrekte Programme zu produzieren – um zumindest die Ausgabe von Sprachmodellen prüfen zu können,
- Beispiel zu zeigen, bei denen ChatGPT/andere Sprachmodelle (subtil) falsche Antworten liefen, z.B. bei der Ausführung von Code (etwa: "Zeige Schritt für Schritt, wie sich das Programm für die folgende Eingabe verhält") oder bei ungewöhnlichen



Aufgabenstellungen (etwa: “Programmiere InsertionSort in $O(n * \log(n))$ mit Hilfe einer geeigneten Datenstruktur.”),

- Lösungen im Tutorium präsentieren zu lassen oder
- ChatGPT-Nutzung wenigstens bei einigen Aufgaben explizit zu gestatten und die Aufgaben entsprechend anzupassen.

Verweise

- Barber (2023). KI-Schreibwerkzeuge aus prüfungs- und urheberrechtlicher Sicht. <https://www.uni-bielefeld.de/einrichtungen/schreiblabor/chat-gpt/2023-07-05-Lehrbar-KI-schreibwerkzeuge.pdf>
- Justitiariat der Universität Bielefeld über rechtliche Fragen zur Nutzung von textgenerierenden digitalen Tools (inkl. Muster Eigenständigkeitserklärung). <https://www.uni-bielefeld.de/themen/pruefungsrecht/chatgpt/> Dezernat Studium und Lehre der Universität Bielefeld (2023). Digitale Technologien zur Textgenerierung und das Schreiben in Studium und Lehre. <https://www.unibielefeld.de/einrichtungen/schreiblabor/chat-gpt/>
- DFG (2023). Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG. https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2023/230921_stellungnahme_praesidium_ki_ai.pdf
- Leschke und Salden (Herausgeber, 2023). Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung. Zentrum für Wissenschaftsdidaktik der Ruhr-Universität Bochum. <https://doi.org/10.13154/294-9734>
- Schulministerium NRW (2023). Textgenerierende KI. <https://www.schulministerium.nrw/textgenerierende-ki>