

# Was sind textgenerierende Technologien?

## Zum Begriff „KI“ und zu großen Sprachmodellen

Das, was hier mit textgenerierenden Technologien gemeint ist, wird heute meist als „generative KI (Künstliche Intelligenz)“ oder auch einfach nur als „KI“ bezeichnet. Der Begriff birgt allerdings einige Herausforderungen. Eine davon ist die Verwechslungsgefahr: Nachdem Ende des Jahres 2022 das große Sprachmodell ChatGPT veröffentlicht wurde, ist es im öffentlichen Diskurs fast synonym zum Begriff „KI“ geworden. Das zeigt sich in Fragen wie: „Wie wird KI den wissenschaftlichen Arbeitsprozess verändern?“ Die Antwort ist: „KI“ beeinflusst unsere Arbeitsprozesse schon seit vielen Jahren – neu hinzugekommen ist jetzt die Möglichkeit der Produktion von Texten und anderen Inhalten (die es ebenfalls schon länger gibt, die durch ChatGPT nur sehr viel zugänglicher geworden ist).

Der Begriff „KI“ ist, weil er so diffus ist, auch aus anderen Perspektiven eher untauglich, um sich produktiv darüber auszutauschen. In diesem Beitrag vertrete ich den Ansatz, dass es hilfreicher ist, die Anwendungen so konkret wie möglich zu benennen. In diesem Beitrag geht es in erster Linie um sogenannte große Sprachmodelle (Large Language Models, kurz LLM), die Text generieren können – im Gegensatz zum Beispiel zu Modellen, die mithilfe statistischer Verfahren wie Clusteranalyse Muster in Datensätzen erkennen können, was für die Forschung durchaus ein wichtiges Hilfsmittel ist.

Mithilfe von Sprachmodellen wird versucht, menschliche Sprache maschinell zu verarbeiten, zu modellieren und auszugeben. Das Feld, das sich damit beschäftigt, ist das Natural Language Processing (NLP). Da geht es um Textproduktion und um andere Anwendungsgebiete, bei denen maschinelle Sprachverarbeitung schon seit längerem genutzt wird: Zur Spracherkennung für automatische Transkription, sprach-gesteuerte Assistenzsysteme im Haushalt, automatische Übersetzung und so weiter.

Große Sprachmodelle funktionieren aufgrund von Wahrscheinlichkeiten: Anhand großer Datenmengen berechnet das Modell, wie wahrscheinlich ein Wort in einem gegebenen Kontext in Bezug auf andere Wörter auftritt. Das hat verschiedene Konsequenzen. Zum Beispiel kommt es dazu, dass Text generiert wird, der Informationen enthält, die wir als falsch beurteilen. Im wissenschaftlichen Bereich sind das zum Beispiel Quellen: Es kann vorkommen, dass ein Artikel

referenziert wird, der sich zwar so anhört, als ob es ihn geben könnte – bei dem aber nach kurzer Recherche auffällt, dass es ihn gar nicht gibt. Da das Modell anhand eines bestimmten Corpus von Daten trainiert wurde, sind die Ergebnisse besser, je mehr dazu in den Daten vorhanden ist, und schlechter, je spezifischer das Thema ist. Wichtig ist auch, dass aufgrund der genutzten Daten bestimmte Biases, also Verzerrungen, auftreten können – z. B. Stereotype über Geschlechter.

## Kann man maschinell generierten Text erkennen?

Eine Frage, die Lehrende häufig stellen und die Sie sich vielleicht ebenfalls bereits gestellt haben, ist die nach der Nachweisbarkeit von Texten, die mithilfe textgenerierender Technologien erstellt wurden. Das ist verständlich: Es ist eine unangenehme Vorstellung, dass man einer gut gemachten Täuschung unterliegen könnte, die dazu führt, dass sich ein\*e Student\*in eine gute Note erschleicht, die – im Vergleich zu anderen Studierenden, die selbst an ihrem Text gearbeitet haben – nicht auf eigener Leistung beruht.

Tatsächlich wäre es wünschenswert, maschinell generierte Inhalte, die der Täuschung dienen sollen, nachweisen zu können: Denken Sie nur an Deepfakes, die im Internet kursieren. Das scheint zum derzeitigen Standpunkt technologisch jedoch nicht umsetzbar, wie eine Studie (Weber-Wulff et al., 2023) und Erfahrungsberichte zeigen. Erkennungssoftware, die angibt, ob ein Inhalt maschinell generiert ist, funktioniert nicht zuverlässig. Was Sie keinesfalls tun sollten, ist, eine Anwendung wie ChatGPT selbst als Mittel zur Erkennung zu nutzen: Dazu sind diese Modelle nicht in der Lage.

Gleichzeitig muss der Einsatz textgenerierender Technologien nicht automatisch mit einem **Täuschungsversuch** einhergehen. Hier spielt die **persönliche Haltung** eine Rolle: Misstrauen – bis hin zu einem **Generalverdacht** gegenüber den Studierenden – ist alles andere als zuträglich, um eine Umgebung zu schaffen, in der Studierende motiviert lernen. Hilfreicher ist es, davon auszugehen, dass Studierende **Interesse an ihrem eigenen Lernerfolg** haben. Ausgehend davon können Sie mit ihnen zusammen Möglichkeiten und Grenzen des Einsatzes textgenerierender Technologien diskutieren. Denn wenn Studierende solche Technologien sinnvoll und mit der notwendigen Kennzeichnung einzusetzen, entstehen Texte, die ihre eigenen sind und in der textgenerierende Technologien nur ein weiteres Hilfsmittel sind, um zu wissenschaftlicher Erkenntnis zu gelangen.

---

## Autor\*in

**Nadine Lordick**, Mitarbeiterin im Projekt KI:edu.nrw und im Schreibzentrum des Zentrums für Wissenschaftsdidaktik der Ruhr-Universität Bochum. Arbeitsschwerpunkt: Künstliche Intelligenz (KI) und wissenschaftliches Schreiben.