

Inez De Florio-Hansen

PÄDAGOGIK

KI-Tools für den Unterricht

2., vollständig überarbeitete Auflage

Mehr als
ChatGTP

BELTZ

Inez De Florio-Hansen ist ehemalige Lehrerin und Fachberaterin für Migration am Hessischen Kultusministerium. Seit 1996 bildet sie als Professorin für Fremdsprachenforschung/Interkulturelle Kommunikation an der Universität Kassel angehende Fremdsprachenlehrkräfte und Wirtschaftsromanisten aus. Ihr Forschungsinteresse gilt dem wissenschaftsorientierten Lehren und Lernen in Schule und Hochschule. In jüngster Zeit widmet sie sich verstärkt dem Lehren und Lernen mit digitalen Technologien, insbesondere KI, und der Ausbildung einer kritischen Medienkompetenz.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt.
Jede Verwertung ist ohne Zustimmung des Verlags unzulässig.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen
und die Einspeicherung und Verarbeitung in elektronische Systeme.
Die Verlagsgruppe Beltz behält sich die Nutzung ihrer Inhalte für Text und
Data Mining im Sinne von § 44b UrhG ausdrücklich vor.



Dieses Buch ist erhältlich als:
ISBN 978-3-407-25981-3 Print
ISBN 978-3-407-25982-0 (PDF)

2. Auflage 2025

© 2025 Verlagsgruppe Beltz
Werderstraße 10, 69469 Weinheim
service@beltz.de
Alle Rechte vorbehalten

Lektorat: Dr. Erik Zyber
Umschlaggestaltung: Michael Matl
Umschlagabbildung: © gettyimages/Salman Alfa

Satz und Herstellung: Michael Matl
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag (ID 15985-2104-1001).
Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

Teil I: Die Bedeutung von KI-Sprachmodellen für Schule und Unterricht	9
1. Funktionsweise und Nutzungsmöglichkeiten verschiedener KI-Sprachmodelle	11
1.1 Warum es wichtig ist, einen eigenen Stil auszubilden	11
1.2 Zur Funktionsweise von KI-Sprachmodellen	14
1.3 Auswirkungen auf Gesellschaft und Berufswelt	18
1.4 Lehren und Lernen mit KI-Sprachmodellen	26
2. Wichtige KI-Systeme im Überblick	32
2.1 Einleitende Bemerkungen	32
2.2 Für Schule und Unterricht immer noch aktuell: ChatGPT und Copilot	32
2.3 FelloFish und Fobizz	35
2.4 Weitere KI-Tools für Schule und Unterricht	37
2.5 KI-Systeme für spezielle Aufgaben	46
3. Zielführende Prompts verfassen	50
3.1 Definition	50
3.2 Beispiel für die Generierung von Kurzgeschichten	50
3.3 Hilfen beim Erstellen von Prompts	54
3.4 Weitere Beispiele für Prompts	57
3.5 Dem Ziel mit Geduld näher kommen	59
3.6 Wichtige Prompting-Ansätze	65
3.7 Prompt-Engineering für Lehrkräfte?	66

Teil II: Zum Umgang von Lehrpersonen und Lernenden mit KI-Sprachmodellen	69
4. Zum Leistungsspektrum von KI-Sprachmodellen	71
4.1 Ziele von KI-Sprachmodellen	71
4.2 Kurzgeschichten im Dialog verfassen	74
4.3 Unterstützung bei der Bildung von Kleingruppen	77
4.4 Die Kommunikation mit einzelnen Lernenden verbessern	81
5. Grundlegende Anwendungsmöglichkeiten für den Unterricht (Sek. I)	85
5.1 Voraussetzungen für die Nutzung von KI-Sprachmodellen und LGAIMs (<i>Large Generative AI Models</i>)	85
5.2 Was es bei KI im Unterricht generell zu bedenken gilt	86
5.3 Einleitung zu den folgenden Unterrichtsbeispielen	87
5.4 Unterrichtsbeispiele für die Klassen 5 bis 7	87
5.5 Unterrichtsbeispiele für die Klassen 8 bis 10	95
5.6 Prompts zum Ausprobieren	103
6. Grundlegende Anwendungsmöglichkeiten in der Sekundarstufe II	106
6.1 Weitere Hinweise zum Einsatz von KI-Systemen	106
6.2 Unterrichtsbeispiele für die Klassen 11 bis 13	107
7. Weitere Nutzungsmöglichkeiten	117
7.1 Vorschläge einschlägiger Autorinnen und Autoren	117
7.2 Multiple-Choice-Tests erstellen	118
7.3 Versuch, mit ChatGPT eine Mindmap zu generieren	120
7.4 Sich über Unterrichtsstrategien und -techniken informieren	123
8. Die Lernenden mit KI-Sprachmodellen vertraut machen	125
8.1 Grundlagen	125
8.2 Die etwas andere Einführung	126
8.3 Prompting für Lernende	130

8.4 Folgeprompts nutzen	131
8.5 Einen individuellen Schreibstil entwickeln	133
9. Der Hype geht weiter – wie damit umgehen?	135
Glossar	137
Literaturverzeichnis	141

Teil I:

Die Bedeutung von KI-Sprachmodellen für Schule und Unterricht

In den folgenden drei Kapiteln gehe ich in gebotener Kürze auf die Funktionsweise sowie Möglichkeiten und Grenzen von KI-Sprachmodellen und vorhersehbare Entwicklungen ein. **Kapitel 1** führt in wichtige Grundlagen ein, deren Kenntnis für eine sinnvolle Nutzung von KI-Sprachmodellen in Schule und Unterricht unerlässlich ist. Ausgehend von einem fiktiven Gespräch zwischen zwei Lehrpersonen (**1.1**) geht es im folgenden Abschnitt (**1.2**) um die technologischen Voraussetzungen, die es ermöglichen, mithilfe eines KI-Sprachmodells menschenähnliche Sprache zu generieren, besser gesagt: mit einem Chatbot in einen Dialog einzutreten. Bei vielen Generierungen sind inhaltliche Verzerrungen, aber auch sprachliche Unzulänglichkeiten nicht zu übersehen. In **1.3** geht es um erwartbare Auswirkungen der sich rasch weiterentwickelnden KI auf Gesellschaft und Arbeitsplätze. Was müssen Schülerinnen und Schüler lernen, um in ihrem späteren (Berufs-) Leben sinnvoll mit KI-Systemen umzugehen? Im letzten Abschnitt (**1.4**) werden einige Unzulänglichkeiten der Generierungen an konkreten Beispielen aufgezeigt, bei denen auch mehrere Prompts, sogenannte Folgeprompts, nicht zu einem zufriedenstellenden Ergebnis führten. Wenn Schülerinnen und Schüler sich auf KI-Anwendungen stützen, sind zudem Fragen des Copyrights und des Datenschutzes zu bedenken. Deshalb müssen die Lernenden angehalten werden, ihre Quellen in geeigneter Form zu benennen. Die wichtigste Möglichkeit, Täuschungsversuche zu unterbinden oder zumindest einzuschränken, besteht in einer grundlegenden Änderung der Aufgabenstellungen.

In **Kapitel 2** stelle ich KI-Systeme vor, die für die Nutzung in Schule und Unterricht besonders geeignet sind. Dazu gehört neben ChatGPT der Copilot von Microsoft (**2.1**). Im darauffolgenden Abschnitt (**2.2**) erläutere ich, warum nach meinem Dafürhalten ChatGPT und Microsoft Copilot für gängige Anwendungen beim schulischen Lehren und Lernen hinreichend sind. Nachdem bisher vor al-

lem Systeme aus den USA den Markt dominierten, gibt es inzwischen auch in Deutschland erstellte Tools, die speziell für den Unterricht ausgelegt sind: Fello-Fish (vormals Fiete.ai) und Fobizz (2.3). In Abschnitt 2.4 stelle ich weitere neue Tools vor: Google Gemini (vormals Bard) (2.4.1), Claude von Anthropic (2.4.2), Llama, Le Chat von Mistral (2.4.3) sowie das chinesische KI-Sprachmodell DeepSeek (2.4.4). Im letzten Abschnitt des Kapitels (2.5) geht es um KI-Systeme für spezielle Aufgaben: Perplexity (2.5.1), DeepL Write und DeepL Translate (2.5.2) sowie ChatPDF (2.5.3). Der Output dieser Tools wird mit Generierungen von ChatGPT verglichen. Es folgen abschließende Bemerkungen zu Bildgenerierungen (2.5.4).

In **Kapitel 3** geht es um die schwierige, aber auch höchst interessante Aufgabe, ohne die von einem KI-Sprachmodell nichts zu erfahren ist, nämlich die sogenannte Eingabeaufforderung, den Prompt. Nach einer Definition (3.1) wird zunächst an einem Beispiel gezeigt, welches das Hessische Kultusministerium in einer Handreichung ausführlich kommentiert hat, was man beim Prompting bedenken muss. Auch wenn man noch so präzise formuliert, muss der Output nicht korrekt sein. Das kann an den Verzerrungen und Unzulänglichkeiten liegen, die von den Daten herrühren, mit denen das jeweilige Tool trainiert wurde. Deutlich wird dies am Beispiel von Wolfgang Borcherts Kurzgeschichte »Nachts schlafen die Ratten doch« (3.2). Verschiedene wissenschaftsorientierte Publikationen benennen wichtige Grundregeln für die Abfassung zielführender Prompts. Während sie anfangs nur in englischer Sprache vorlagen, finden sich inzwischen auch deutschsprachige Veröffentlichungen, unter anderem vom Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (3.3). Die englischsprachigen Quellen lassen sich jedoch mit Gewinn für den Englischunterricht nutzen (3.4). Generell ist es ratsam, beim Prompting (gleichgültig ob auf Deutsch oder Englisch) möglichst kleinschrittig vorzugehen. Das zeigt sich bei einem Vergleich zwischen der Suchmaschine Google Drive und ChatGPT zur Unterrichtsmethode Flipped Classroom, die häufig bei der Nutzung von KI-Sprachmodellen eingesetzt wird (3.5). Es folgen kurze Hinweise zu den wichtigsten Prompting-Ansätzen (3.6). Den Abschluss des Kapitels (3.7) bildet ein Ausblick auf komplexere Prompting-Ansätze für fortgeschrittene Lehrkräfte, vor allem Informatiklehrer.

1. Funktionsweise und Nutzungsmöglichkeiten verschiedener KI-Sprachmodelle

1.1 Warum es wichtig ist, einen eigenen Stil auszubilden

Als OpenAI Ende 2022 das KI-Sprachmodell ChatGPT kostenfrei im Web zur Verfügung gestellt hat, ging es in den Gesprächen im Lehrerzimmer vor allem darum, warum Verbote keine Lösung sind. Inzwischen sprechen Susanne B. und ihr Kollege Frederik M. immer häufiger über die möglichst sinnvolle Nutzung von KI im Unterricht. Ihre Hauptsorge ist die rasante Entwicklung der Tools. Nicht nur ChatGPT hat in den letzten zweieinhalb Jahren zahlreiche Updates erfahren; weitere KI-Sprachmodelle sind auf den Markt gekommen, die sich bisweilen sogar besser für das Lehren und Lernen in Schule und Unterricht eignen. Als Susanne den Kollegen im Lehrerzimmer entdeckt, ist Frederik gerade mit seinem Smartphone beschäftigt.

Susanne B.: Hallo, gibt es etwas Neues?

Frederik M.: Wie meinst du das? Bei diesen Tools gibt es fast täglich neue Nutzungsmöglichkeiten auch für den Unterricht.

Susanne B.: Und was hast du gerade entdeckt?

Frederik M.: Es gibt jetzt eine Plattform namens Fobizz, mit deren Hilfe auch Schülerinnen und Schüler arbeiten können, ohne dass ihre Daten irgendwie gespeichert werden. Den Zugriff ermöglichen wir.

Susanne B.: Interessant, wenn es tatsächlich stimmt, ist das ein großer Fortschritt.

Frederik M.: Ich werde es gleich heute Nachmittag ausprobieren.

Susanne B.: Es wäre ein echter Gewinn, wenn die Lernenden auch bei diesem Tool angehalten würden, einen eigenen Stil auszubilden.

Frederik M.: Ja, ich habe neulich einen Beitrag in einer Fachzeitschrift gelesen, wo der Lehrer von einem Schüler, dessen Hausarbeit er beurteilen soll, eine »eigene Stimme« fordert. Es ist also keineswegs damit getan, dass die Lernenden angeben, an welchen Stellen und in welcher Form sie welches Tool genutzt haben.

- Susanne B.: Zumal es bis heute keine Detektoren gibt, die einwandfrei angeben können, ob eine Schülerin oder ein Schüler geschummelt hat.
- Frederik M.: Diese Sorge habe ich inzwischen weniger. Meine Oberstufenschüler haben erkannt, dass man wichtige Informationen anhand anderer Quellen überprüfen muss. Sie haben verstanden, dass solche statistischen Systeme an ihre Grenzen stoßen, vor allem, weil die Tools nur in Ausnahmefällen die Generierung auf eine Anfrage ablehnen.
- Susanne B.: Genau, die Systeme produzieren geschmeidige Texte, um den Usern gefällig zu sein. Das gilt nicht nur für ChatGPT, sondern bis zu einem gewissen Maß auch für viele andere Systeme.
- Frederik M.: Und die falschen Aussagen und Verzerrungen haben auch nicht abgenommen. Du hattest ganz recht, als du gesagt hast, es werde immer wieder von Zeitersparnis geredet, das sei aber falsch. Aus meiner Sicht hat der Zeitaufwand in letzter Zeit sogar zugenommen.
- Susanne B.: Da hast du vollkommen recht. Ich nutze inzwischen nur noch zwei bis drei KI-Tools, jedes für andere Aufgaben. Manchmal ist die Versuchung natürlich groß, den Output verschiedener KI-Systeme auf der Grundlage desselben Prompts zu vergleichen.
- Frederik M.: Und manchmal ist der Vergleich durchaus hilfreich, besonders wenn man daran denkt, was auf die Lernenden in ihrem späteren beruflichen Leben zukommt.
- Susanne B.: Weißt du was? Wir sollten uns mal wieder mit anderen Kolleginnen und Kollegen darüber austauschen, wie sie mit den Neuerungen umgehen.
- Frederik M.: Das ist eine gute Idee. Das machen wir!

In ihrer kurzen Unterredung sprechen die beiden Lehrkräfte – direkt oder indirekt – drei wichtige Gesichtspunkte an, die mit der Nutzung von KI-Sprachmodellen (*Large Language Models*, LLMs) beim schulischen Lehren und Lernen verbunden sind. In den folgenden Abschnitten gehe ich kurz auf diese Aspekte ein, denn sie bilden die Grundlage für eine sinnvolle Verwendung von KI-Systemen im Unterricht.

Zunächst behandle ich in gebotener Kürze die technischen Grundlagen von KI-Sprachmodellen (vgl. 1.2): Sich einzuarbeiten, um die Tools sinnvoll nutzen zu können, bedarf einer gewissen Anstrengung und ist mit Zeitaufwand verbunden. Dabei spielt es keine große Rolle, ob sie gratis zur Verfügung stehen oder ob sie kostenpflichtig sind. Experten wie Ethan Mollick weisen immer wieder darauf hin, dass Lehrpersonen fünf bis zehn Stunden investieren müssen, bevor sie einigermaßen angemessen mit den Tools im Unterricht arbeiten können (Mollick; Blogpost: One useful thing 17.03.2023; Mollick 2024). Damit ist nicht die vergleichsweise

einfache Einrichtung eines KI-Systems auf dem eigenen Laptop oder Smartphone gemeint (vgl. Kap. 2). Vielmehr geht es darum, das Leistungsspektrum der Tools kennenzulernen: Inwieweit können KI-Systeme zur Verbesserung der fachlichen und überfachlichen Lernergebnisse beitragen? Darüber hinaus muss jede Lehrkraft sich Gedanken darüber machen, welche Auswirkungen die Tools auf die Persönlichkeitsentwicklung der Schülerinnen und Schüler haben können.

Im folgenden Kapitel (Kap. 2) zeige ich kurz auf, wie man die neuen Tools am einfachsten nutzen kann. Vor allem geht es darum, in welchen Bereichen ihre besonderen Stärken und Schwächen mit Blick auf schulisches Lehren und Lernen liegen. Kapitel 3 ist dem Prompting vorbehalten: Noch immer bestimmen passende Anweisungen den Output von KI-Systemen, gleichgültig ob ihr Anwendungsspektrum allgemein ausgerichtet ist oder speziell für Schule und Unterricht konzipiert wurde. Zwar zeichnet sich ab, dass die Systeme früher oder später das Autoprompting, also die selbstständige Gestaltung der Prompts nach kurzen Eingabeaufforderungen, übernehmen werden. Generell ist es aber beim schulischen Lehren und Lernen überaus sinnvoll, dass vor allem Schülerinnen und Schüler klar formulieren können, welche Aufgaben die Tools konkret erledigen sollen.

Im obigen Meinungsaustausch zwischen den beiden Lehrpersonen wird – über die technischen Voraussetzungen hinaus – der erwartbare Einfluss der Tools auf die Gesellschaft und die Berufswelt kurz angesprochen. Ist es sinnvoll, KI-Sprachmodelle wie ChatGPT mit der Erfindung des Buchdrucks oder des Internets zu vergleichen? Werden die Berufe, auf die sich Schülerinnen und Schüler heute vorbereiten, in Zukunft noch existieren? Werden sie überhaupt einen Arbeitsplatz finden? Vor allem aber: Was sollten sie lernen, um in einer von künstlicher Intelligenz dominierten Welt zu bestehen? Mögliche Antworten auf diese Fragen gebe ich in Abschnitt 1.3.

Daran schließen sich erste Überlegungen in Bezug auf die Nutzung der Systeme für schulisches Lehren und Lernen an (vgl. 1.4). Ein konkretes Generierungsbeispiel vermittelt nicht nur erste Einblicke in die Funktionsweise von ChatGPT im Vergleich zum Copilot von Microsoft. Die Gegenüberstellung belegt auch die mehr oder weniger große Unzulänglichkeit der Bots, nämlich über dem Wortreichtum die inhaltlichen Aspekte zu vernachlässigen. Ein weiteres Beispiel gibt erste Hinweise auf Nutzungsmöglichkeiten nicht nur für Lehrkräfte, sondern auch für Schülerinnen und Schüler. Daraus ergeben sich Fragen des Urheberrechts, des Datenschutzes sowie des Umgangs mit Täuschungsversuchen (vgl. Rieck 2023). Hauptsächlich aber geht es um die längst fällige Neuorientierung des Unterrichts: weg von reproduktiven Aufgaben, die KI-Sprachmodelle erledigen können, hin zu Aktivitäten, bei denen kritisches Denken und persönliche Stellungnahmen im Vordergrund stehen.

Generell kann man sagen, dass es bei KI in Schule und Unterricht nicht vorrangig um Arbeitserleichterung beim fachlichen Lehren und Lernen durch den sinn-

vollen Einsatz der Tools geht. Zwar ist die Frage berechtigt: »Was bringt uns das?«, aber wichtiger ist die Einsicht, dass KI die Neuausrichtung von Lehren und Lernen – gewollt oder ungewollt – vorantreibt. Eine wichtige Frage lautet: Wie können wir KI so in Schule und Unterricht integrieren, dass die Persönlichkeitsentwicklung der Schülerinnen und Schüler mit Blick auf ihr künftiges (Berufs-)Leben gefördert wird?

Die mögliche Hilfe beim schulischen Lehren und Lernen darf nicht dazu führen, dass Lehrpersonen und Lernende dies für den wichtigsten Aspekt im Zusammenhang mit dem Einsatz von KI-Systemen halten. Die Arbeitserleichterung stellt einen Anreiz dar, sich mit darüberhinausgehenden Aspekten zu beschäftigen: KI gibt Anstöße, die längst fällige Reform des schulischen Lernens endlich ernsthaft in Angriff zu nehmen. Mehr noch: Schülerinnen und Schüler müssen damit vertraut gemacht werden, dass es in einer von KI dominierten Welt ganz wesentlich darauf ankommt, den Wandel mitzugestalten, nicht zuletzt um für die Zukunft gerüstet zu sein. Der (möglichst) kompetente Umgang mit KI – Mollick spricht von Zusammenarbeit mit smarten Maschinen (2024) – stellt erhebliche Anforderungen an Lehrkräfte und Lernende, denen man nicht ausweichen kann und darf. Also nicht: Wenn wir die Systeme kennengelernt haben, ist dies hinreichend, denn wir profitieren für das Lehren und Lernen davon, sondern vielmehr: Jetzt sind wir motiviert, den sich durch Digitalisierung und KI abzeichnenden gesellschaftlichen Wandel aktiv und konstruktiv mitzugestalten.

1.2 Zur Funktionsweise von KI-Sprachmodellen

Der Begriff »Künstliche Intelligenz« taucht erstmals im Zusammenhang mit dem *Dartmouth Summer Research Project on Artificial Intelligence* auf, einer Tagung, die 1956 am Dartmouth College in New Hampshire stattfand. Eine der Grundannahmen der Initiatoren hat auch heute noch Gültigkeit: KI stützt sich auf die Annahme, dass grundsätzlich alle Aspekte des Lernens und andere Merkmale der Intelligenz so genau beschrieben werden können, dass eine Maschine zur Simulation dieser Vorgänge gebaut werden kann (vgl. auch zum Folgenden De Florio-Hansen 2020, S. 42 ff.). Die Auseinandersetzung mit der Frage, ob es möglich ist, menschliche Intelligenz in Maschinen nachzubilden, hat zuvor schon Alan Turing (1912–1952), den Urheber des Turing-Tests, beschäftigt (vgl. De Florio-Hansen 2020, S. 42–47).

Trotz beeindruckender Fortschritte gibt es bisher keine Ausprägung von KI, welche die kognitiven, emotionalen und sozialen Fähigkeiten des Menschen aufweist, und es ist fraglich, ob es sie jemals geben wird. Das Problem dabei ist nicht die in Science-Fiction häufig thematisierte Machtübernahme durch KI-gestützte Systeme (*AI takeover*), sondern unser Umgang mit ihnen. Bereits 2019 (deutsch

2020) hat der bekannte KI-Forscher Stuart Russell, Professor für Informatik an der University of California in Berkeley, ein hilfreiches Buch mit dem Titel *Human Compatible – Künstliche Intelligenz und wie der Mensch die Kontrolle über superintelligente Maschinen behält* veröffentlicht.

Schaut man sich den Output von KI-Sprachmodellen wie ChatGPT an, ist es nicht verwunderlich, dass viele User Bewusstsein, Empathiefähigkeit und Denkvermögen mit den Tools verbinden. Das gilt insbesondere für die neueren KI-Systeme, die das *Reasoning* praktizieren. Die Tools überprüfen ihren Output vor der Generierung mehrmals. So hochentwickelt sie auch sein mögen: KI-Sprachmodelle beruhen auf mathematischen Berechnungen in Form statistischer Modelle. Sie haben keinerlei Weltverständnis und kennen die Bedeutung der Wörter und Texte nicht, die sie generieren. Darauf weist unter anderem Stefan Buijsman immer wieder hin (Buijsman 2021, S. 15).

Seit den 1950er Jahren hat sich die KI-basierte Forschung trotz einiger Rückschläge kontinuierlich weiterentwickelt. Ab 2010 ist ein exponentielles Wachstum zu verzeichnen. Das ist unter anderem auf die immer ausgeklügelteren Algorithmen zurückzuführen, welche die Grundlage heutiger KI-Systeme bilden (zu den Fachtermini vgl. Glossar).

Ein Algorithmus ist eine eindeutige Handlungsanweisung zur Lösung eines Problems. Folglich beruht ein Kochrezept auf einem Algorithmus, obwohl der Terminus in der Regel für mathematische Instruktionen bei der Programmierung von Computern verwendet wird. Eine wissenschaftsorientierte Definition von Algorithmus lautet:

1. Eine präzise, das heißt in einer festgelegten Sprache abgefasste, endliche Beschreibung eines allgemeinen Verfahrens unter Verwendung elementarer Verarbeitungsschritte zur Lösung einer gegebenen Aufgabe.
2. Lösungsverfahren in Form einer Verfahrensanweisung, die in einer wohldefinierten Abfolge von Schritten zur Problemlösung führt (vgl. Gabler Wirtschaftslexikon 2023).

So präzise die Definitionen auch sein mögen: Algorithmen sind immer nur so gut wie die Daten und Anweisungen, auf denen sie basieren.

Darüber hinaus beruhen die bedeutenden Fortschritte des letzten Jahrzehnts auf dem sogenannten maschinellen Lernen (*Machine Learning*). Diese Art von KI erkennt in Daten, mit denen das System trainiert wurde, Muster und kann diese Erkenntnisse anschließend auf neue Daten anwenden, ohne dass ein Programmierer eingreifen muss. Die Systeme arbeiten mit Tokens, technischen Einheiten, die sich wechselseitig bedingen. Vereinfacht ausgedrückt, berechnet ein KI-System wie ChatGPT auf der Grundlage eines eingegebenen Textsegments, welches Wort bzw. welche Wörter mit hoher Wahrscheinlichkeit auf den Input folgen, und verarbeitet sie zum Output. Was solchen Systemen vor allem fehlt, ist Intentionalität.

Im Unterschied zum Menschen verfolgen sie mit ihren Generierungen keinerlei Absicht.

Ein Grund für die Erfolge des maschinellen Lernens besteht darin, dass es sich mehr und mehr auf die Nachbildung von neuronalen Netzen (*Neural Networks*) des menschlichen Gehirns stützt. Neuronen können Informationen von außen oder von anderen Neuronen aufnehmen und modifiziert an andere Neuronen weiterleiten oder als Endergebnis ausgeben. Zwischen der Eingabe und der Ausgabe können verschiedene Schichten liegen, die – je nach Anwendung – das Endergebnis bestimmen.

Im Zusammenhang mit künstlichen neuronalen Netzwerken ist häufig von *Deep Learning* (etwa »tief gehendes Lernen«) die Rede. Bisher wurde nicht festgelegt, wie viele Schichten zwischen der Eingabe und der Ausgabe liegen sollten, damit maschinelles Lernen als *Deep Learning* gelten kann. Je zahlreicher die zwischen Ein- und Ausgabe aufeinander folgenden Schichten sind, umso schwerer ist es für Menschen, auch für KI-Experten und -Expertinnen, die Lösung nachzuvollziehen.

Die rasante Verbesserung der genannten Bereiche hat in den letzten Jahren die Entwicklung von *Large Generative AI Models (LGAIMs)* ermöglicht, das heißt von KI-Systemen, die Ergebnisse in verschiedenen Bereichen (z. B. Text, Bild, Video, Audio; vgl. Kap. 2) generieren. Eine spezielle Form stellen *Large Language Models (LLMs)* dar, die auf dem *Natural Language Processing (NLP)* beruhen. NLP – das Verfahren ist bereits um 1950 entstanden – ist eine KI-Methode, die sich mit der Analyse, dem Verständnis und der Generierung von natürlicher Sprache beschäftigt. Inzwischen geht es darum, Rechner so zu programmieren, dass sie große Mengen natürlicher Sprachdaten mündlich und/oder schriftlich verarbeiten und analysieren können. Der beeindruckende Output von KI-Sprachmodellen beruht auf Techniken mathematischer Optimierung, die während des Trainings immer wieder angewendet werden, bis das System den Input in wünschenswerten Output transformiert. Die Konvertierung wurde und wird vor allem dadurch unterstützt, dass immer größere Datenmengen und leistungsstärkere Computer zur Verfügung stehen. Trotz der Fortschritte stellt die Vielfalt menschlicher Sprache, einschließlich Ironie und Mehrdeutigkeiten – man denke nur an Wörter wie »Bank« oder »Gericht« –, auch weiterhin eine große Herausforderung dar.

Häufig ist auch von *Foundation Models* (etwa: »Grundlagenmodelle«) die Rede. Sie verfügen über vielfältige Verwendungsmöglichkeiten und liegen den meisten spezialisierten Systemen zugrunde. Zum Vorgehen von OpenAI, dem Unternehmen, das ChatGPT geschaffen und in Umlauf gesetzt hat, schreibt Helen Toner in einem Blogpost vom Mai 2023, dass ChatGPT das KI-Sprachmodell GPT-3.5 zugrunde liegt, dieses Foundation Model jedoch von OpenAI an die Nutzung in einem Chatbot-Setting angepasst wurde (<https://time.com/6280533/ai-chatbots-improv-machines/>; letzter Zugriff Jan. 2025). Inzwischen basiert die Gratisversion des Tools, ChatGPT 4o mini, ebenso wie der Copilot auf GPT-4 (vgl. Kap. 2).

Die bisher kurz erläuterten Grundlagen gelten für (fast) alle LLMs. Die meisten werden über das *Unsupervised Learning* (Unüberwachtes Lernen) hinaus auch durch *Supervised Learning* (Überwachtes Lernen) optimiert. Beim Überwachten Lernen werden dem Tool Unterhaltungen vorgelegt, in denen menschliche Trainer beispielsweise beide Rollen übernehmen: die des Users und die des Chatbots. Man kann sich fragen, welche Vorstellungen die menschlichen Trainer zugrunde legen bzw. wie unvoreingenommen sie tatsächlich sind. Was also macht ChatGPT so besonders? Obwohl OpenAI nicht alle Verfahren zur Optimierung seiner Sprachmodelle offengelegt hat (es gibt kein Handbuch oder auf *Peer-Review* gegründete wissenschaftliche Veröffentlichungen), haben der CEO Sam Altman und weitere Experten auf folgende Besonderheit von ChatGPT hingewiesen: ChatGPT wurde zusätzlich mithilfe von Reinforcement Learning optimiert (<https://dida.do/de/blog/wie-chatgpt-mit-hilfe-von-reinforcement-learning-optimiert-wird>; letzter Zugriff August 2023). Durch *Reinforcement Learning from Human Feedback* kann ein zugrunde liegendes KI-Sprachmodell durch die Verbindung mit einem weiteren KI-System deutlich verbessert werden. Das zusätzliche KI-System, mit dem ChatGPT optimiert wurde, ist ein sogenanntes Reward Model («Belohnungsmodell»). Durch ein Reward Model wird festgelegt, welche Antworten menschliche Nutzerinnen und Nutzer in bestimmten Fällen bevorzugen. In der Regel besteht die Auswahl, an der sich das Reward Model orientiert, aus einer vorgegebenen Liste von Möglichkeiten. Da sich OpenAI und andere Herstellerfirmen hinsichtlich der Nutzung und Zusammenführung verschiedener KI-Systeme aus Gründen der Konkurrenz bedeckt halten, sind konkretere Angaben nicht möglich.

Das ausgefeilte Training von KI-Sprachmodellen bringt einen entscheidenden Nachteil mit sich: Nicht die inhaltliche Wahrheit einer Generierung, sondern die wohlklingende Formulierung hat für die Systeme Priorität. Alle bisher auf dem Markt befindlichen KI-Systeme weisen mehr oder weniger große Unzulänglichkeiten auf: Wenn die Systeme überfordert sind, kommt es zu sogenannten *Hallucinations*, zuverlässig klingenden Aussagen, die durch die Trainingsdaten nicht gerechtfertigt sind. Die umfangreichen Trainingsdaten der meisten LLMs sind außerdem von einer westlichen Sicht auf die Welt geprägt. Diese Verzerrungen werden als *Bias* bezeichnet. Die Tools generieren stets das, was dem statistischen Mittelwert entspricht: Lässt man eine Liebesgeschichte generieren, geht es immer um ein heterosexuelles Paar. Die Hersteller sind zwar bemüht, sexistische sowie rassistische Diskriminierungen zu unterbinden. Fragt man aber geschickt nach, sind diese Mängel immer noch vorhanden.

Außerdem können KI-Sprachmodelle oft keine Quellen für ihre Generierungen benennen. Letzteres führt beispielsweise bei ChatGPT dazu, dass das System auf Nachfrage sogar Quellen erfindet: Es benennt Publikationen einschließlich Autoren und ISBN-Nummern, die nicht existieren. Bei den konkreten Unterrichtsbei-

spielen in Teil II haben wir Gelegenheit, uns mit diesen Einschränkungen auseinanderzusetzen.

Festzuhalten gilt der von den Anbietern und in wissenschaftlichen Publikationen mehr oder weniger ausführlich behandelte Hinweis: Jede Generierung eines KI-Sprachmodells bedarf der Überprüfung anhand von eigenem Wissen sowie von zuverlässigen Quellen. Auf der Startseite zu ChatGPT heißt es: »ChatGPT kann Fehler machen. Überprüfe wichtige Informationen.«

Mit derartigen Fehlinformationen und Vorurteilen beschäftigt sich auch das Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen in seinem im März 2023 veröffentlichten Handlungsleitfaden *Umgang mit textgenerierenden KI-Systemen*: »Die produzierten Inhalte dürfen also nicht unreflektiert übernommen, sondern müssen von den Lernenden kontrolliert werden. Dies ist nur möglich, wenn die Lernenden ausreichend eigenes Wissen über den Gegenstand erworben haben. Damit gewinnt Medienkompetenz noch einmal an Bedeutung. Gerade die Fähigkeit, Fake News von Fakten auf der Grundlage eines gesicherten Wissens zu unterscheiden, wird immer wichtiger« (MSB NRW 2023, S. 4).

Fitzpatrick und seine Mitautoren, die den genannten NRW-Handlungsleitfaden in englischer Übersetzung als Zusammenfassung ihres Buches verwenden, empfehlen für die Überprüfung und Verbesserung von KI generierten Inhalten das Verfahren EDIT (Fitzpatrick u. a. 2023, S. 100 ff.):

Evaluate your AI output content for language, facts and structure.

Determine accuracy and corroborate with source.

Identify biases and misinformation in output.

Transform content to reflect adjustments and new findings.¹

1.3 Auswirkungen auf Gesellschaft und Berufswelt

Selbst wenn man die Nutzung von KI-Sprachmodellen beim fachlichen Lehren und Lernen mit Skepsis betrachtet, darf man die Behandlung generativer KIs im Unterricht keinesfalls aussparen. Jede Lehrperson hat die Aufgabe, die Schülerinnen und Schüler über fachliche Ziele und Inhalte hinaus auf ihr künftiges Leben, vor allem die Berufswelt, vorzubereiten. Dazu bedarf es verschiedener Einblicke in KI-Systeme, vor allem in LLMs wie ChatGPT.

1 »Evaluieren (Überprüfen) Sie den Inhalt des Outputs in Bezug auf Sprache, Fakten und Struktur. Determinieren (Bestimmen) Sie die Genauigkeit und untermauern Sie sie mit Quellen. Identifizieren (Erkennen) Sie Verzerrungen und Fehlinformationen im Output. Transformieren (Verändern) Sie den Inhalt durch Anpassungen und neue Einsichten.«
(eigene Übersetzung)