

2023

# GENE RATI VE KI

Ein Leitfaden zur Einführung in generative KI und ChatGPT. Erstellt von Vincent Timm, wissenschaftlicher Mitarbeiter an der HAWK Hildesheim/Holzminde/Göttingen.



# GENERATIVE KI

In diesem Kapitel gebe ich einen ersten, leicht verständlichen Überblick über generative KI. Ich erkläre, wie maschinelle Intelligenz selbstständig Inhalte wie Bilder oder Texte erzeugen kann.

# 01

# Bildgenerierende KI und der geheimnisvolle latente Raum.

Der latente Raum ist ein sehr abstrakter mathematischer Raum, der durch das Lernen von Daten durch KI-Modelle gebildet wird. In Systemen, die mit Bildern arbeiten, repräsentiert jede Dimension in diesem Raum ein Merkmal oder eine Eigenschaft, die das Modell aus den Daten gelernt hat. Dabei kann es sich um so unterschiedliche Dinge wie die Farbe im Bild, die Position von Objekten, ihre Größe oder Form usw. handeln. Die genaue Anzahl dieser Dimensionen hängt von der Komplexität und der Art des Modells ab. Oft haben KI-Modelle Hunderte oder sogar Tausende von Dimensionen im latenten Raum. Das ist schwer vorstellbar, da wir Menschen nur in drei Dimensionen wahrnehmen. Angenommen, wir hätten einen einfachen latenten Raum mit nur zwei Dimensionen, der die Farbe und die Größe eines Objekts in einem Bild repräsentiert. In diesem Raum könnte ein Punkt nahe an der Achse ein kleines blaues Objekt darstellen, während ein Punkt weiter entfernt von der Achse ein großes rosa Objekt darstellen könnte. Bild-KI-Modelle bilden zunächst diesen latenten Raum, indem sie aus unzähligen Beispielbildern lernen. Sie „betrachten“ diese Bilder und versu-

chen, die zugrunde liegenden Muster und Merkmale zu erfassen, die dann als Punkte im latenten Raum dargestellt werden. Der nächste Schritt ist die Diffusion. Dieser Prozess beginnt mit einem zufälligen Punkt im latenten Raum und „diffundiert“ oder „wandert“ schrittweise in Richtung des endgültigen Bildes. Man kann sich vorstellen, dass das KI-Modell von diesem zufälligen Punkt aus eine Art Pfad durch den latenten Raum „wandert“ und dabei nach und nach immer mehr Merkmale des endgültigen Bildes „aufsammelt“. Bei jedem Schritt der Diffusion passt das Modell seine Vorhersage für das endgültige Bild an den aktuellen Punkt im latenten Raum an. Dies ist ein wenig so, als würde man ein verrauschtes Foto aufnehmen und es nach und nach immer klarer machen, bis man das endgültige, klare Bild erhält. Dies ermöglicht es der KI, ein einzigartiges Bild zu erzeugen, das die gelernten Muster und Merkmale repräsentiert, anstatt einfach den Inhalt der Trainingsbilder zu kopieren. Es wird wirklich ein neues Bild erzeugt, das auf den tiefen Merkmalen und Mustern basiert, die das Modell gelernt hat. Dies macht den Prozess kreativ und spannend.

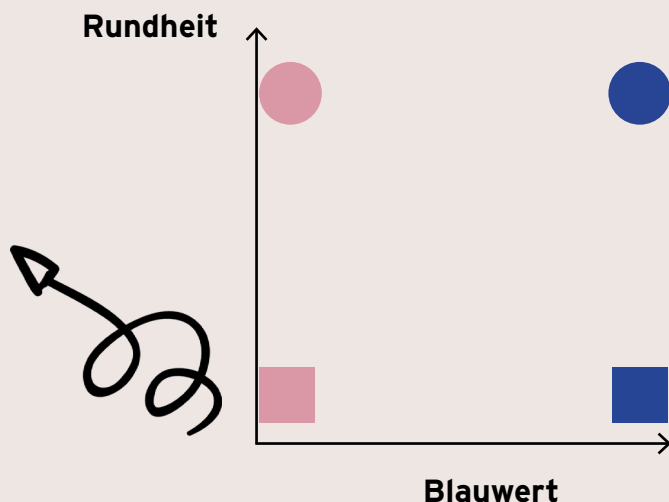


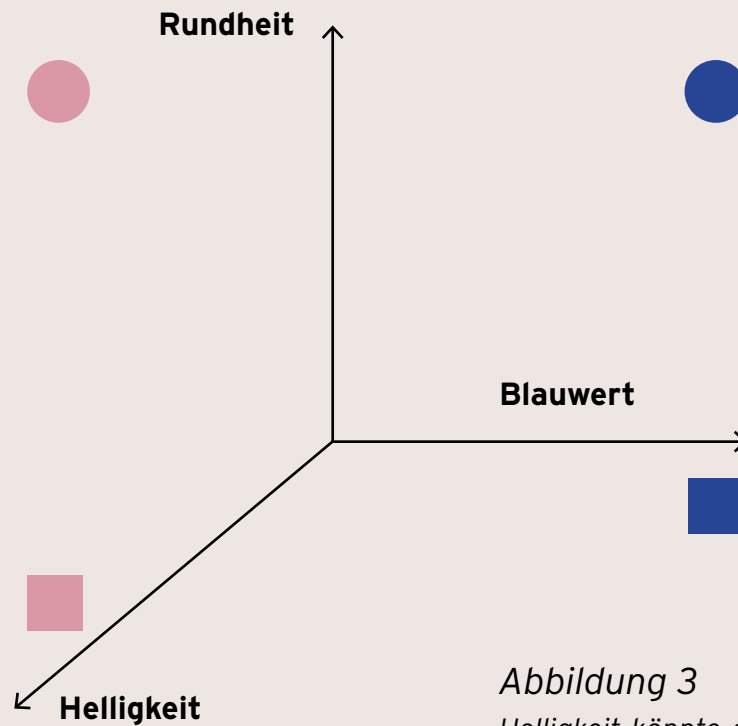
Abbildung 1

Aller Anfang ist leicht. Wir beginnen mit einer Dimension und bestimmen nur den Blauwert runder Objekte.

Abbildung 2

Sobald im Training erkannt wird, dass es nicht nur runde Objekte gibt, wird eine weitere Dimension zur Beschreibung des Objektes notwendig.





**Abbildung 3**

*Helligkeit könnte eine weitere Beschreibung sein, so dass die Objekte nun im dreidimensionalen Bereich positioniert werden müssen. In KI-Modellen gibt es jedoch Hunderte oder Tausende von Dimensionen, in denen die Daten angeordnet sind.*

## Gib mir Text, ich geb dir Sinn.

GPT-4 ist wie sein Vorgänger GPT-3 ein KI-Modell von OpenAI, das sich auf die Textgenerierung spezialisiert hat. Im Wesentlichen versucht es, auf der Basis dessen, was es bisher generiert hat, die nächste sinnvolle Eingabe vorherzusagen. Aber wie funktioniert das genau? Ähnlich wie das oben beschriebene Bild-KI-Modell lernt GPT-4 aus großen Mengen von Textdaten und erstellt einen latenten Raum. In diesem Raum repräsentiert jede Dimension jedoch nicht Farben oder Formen von Objekten, sondern Merkmale des Textes - wie Wortbedeutungen, grammatikalische Regeln, Kontext und so weiter. Man kann sich den latenten Raum von GPT-4 wie eine riesige Bibliothek vorstellen. Jedes Buch in der Bibliothek ist ein Punkt im Raum und repräsentiert einen bestimmten Aspekt des menschlichen Wissens. Wenn GPT-4 einen Text erzeugt, durchsucht es diese Bibliothek und kombiniert Informationen aus verschiedenen Büchern, um einen kohärenten und sinnvollen Text zu erzeugen. Der Diffusionsprozess in GPT-4 ist etwas anders. Anstatt an einem zufälligen Punkt im Raum zu beginnen, leitet GPT-4 seinen Weg durch den Raum auf

der Grundlage von Eingaben, die es bereits erhalten hat, wie z. B. vorangegangene Wörter oder Sätze in einem Textabschnitt. Mit jedem neuen Wort, das das Modell erzeugt, „wandert“ es zu einem neuen Punkt im Raum, der auf der Grundlage der vorherigen Wörter und des Kontexts ausgewählt wird. Wie das Bild-KI-Modell erzeugt GPT-4 einzigartige Inhalte auf der Grundlage dessen, was es aus dem Datensatz gelernt hat - es kopiert nicht einfach die Originaldaten. Stellen Sie sich vor, Sie bitten GPT-4, eine Geschichte über einen lila Dinosaurier zu schreiben. Das Modell wird in seiner „Bibliothek“ nach Büchern suchen, die Informationen über Dinosaurier, die Farbe Lila und das Schreiben von Geschichten enthalten, und dann eine einzigartige Geschichte schreiben, die all diese Elemente miteinander verbindet. Dies macht GPT-4 zu einem äußerst leistungsfähigen und vielseitigen Werkzeug für die Textgenerierung. Gleichzeitig ist es ein faszinierendes Beispiel dafür, wie KI in der Lage ist, komplexe menschliche Fähigkeiten wie das Verfassen von Texten zu erlernen und zu emulieren.

# HERAUS- FORDERUNGEN

# FOR 2

Im zweiten Teil beleuchten wir einige Bereiche, die bei der Arbeit für oder mit KI-Werkzeugen zu beachten sind. Dabei betrachten wir nicht nur die Auswirkungen auf die Gesellschaft, sondern hinterfragen auch die Entstehung dieser Modelle.

# Digitale Armut

Diejenigen, die mit generativen KI-Trainingsmodellen arbeiten, sollten genau hinschauen und hinterfragen, welche Werte, Kulturen und Gewohnheiten in diese Modelle einfließen. In den technologisch fortgeschrittenen Ländern hat die rasche Einführung der künstlichen Intelligenz die Datenproduktion und -verarbeitung stark beschleunigt und gleichzeitig die Konzentration des KI-Reichtums im globalen Norden verstärkt. Datenarme Regionen werden dadurch noch weiter ausgeschlossen und laufen Gefahr, langfristig von den in den GPT-Modellen gesetzten Standards kolonisiert zu werden. Die aktuellen ChatGPT-Modelle werden mit Daten von Online-Nutzer\*innen trainiert, die die Werte und Normen des Globalen Nordens widerspiegeln. Dies macht sie ungeeignet für lokal relevante KI-Algorithmen in datenarmen Gemeinschaften in vielen Teilen des Globalen Südens oder in benachteiligten Gemeinschaften im Globalen Norden. Politiker\*innen und andere Entscheidungsträger\*innen müssen darauf achten, dass diese Modelle nicht Ungerechtigkeiten verstärken. Sie sollten eingreifen, wenn die Kluft bei der Nutzung und Kontrolle dieser Modelle größer wird.

## Regulierung

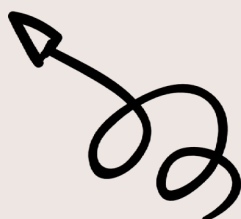
Es ist wichtig, sich des Mangels an angemessenen Regelungen zum Schutz persönlicher Daten oder interner Informationen von Institutionen und Einzelpersonen sowie der Rechte von Bürger\*innen bewusst zu sein. KI-Sprachmodelle und Verbesserungen dieser Modelle durch die Nutzer\*innen gehören Unternehmen, die diese Daten weiterverarbeiten und davon profitieren. Länder und Institutionen sollten auf rechtliche Probleme reagieren, die durch KI-Generierungen entstehen können.



## Black Box

Es ist wichtig zu wissen, dass generative KI-Systeme wie „Black Boxes“ funktionieren und es daher schwierig zu verstehen ist, warum bestimmte Inhalte erstellt wurden. Das Fehlen einer Erklärung, wie die Ergebnisse erzeugt werden, führt dazu, dass die Nutzer\*innen in der Logik gefangen sind, die durch die in den KI-Systemen festgelegten Parameter bestimmt wird. Diese Parameter können bestimmte kulturelle oder kommerzielle Werte und Normen widerspiegeln, die den generierten Inhalt implizit beeinflussen.

[Klicke hier, um zu lesen, wie wir diese Black Boxes verstehen können.](#)



## Feedbackloop

Man muss sich darüber im Klaren sein, dass generative KI-Systeme anstößige und unethische Inhalte produzieren können. Sie sollten sich auch der langfristigen Probleme bewusst sein, die sich für die Verlässlichkeit von Wissen ergeben können, wenn zukünftige GPT-Modelle auf Texten basieren, die von früheren GPT-Modellen erzeugt wurden.



## Fakten, Fakten, Fakten

Die Ergebnisse einer textgenerierenden KI können täuschend menschlich aussehen, als ob das System den generierten Text verstehen würde. Tatsächlich versteht die generative KI aber gar nichts. Diese Werkzeuge reihen lediglich Wörter so aneinander, wie sie es aus den Datensätzen gelernt haben. Der erzeugte Text kann auch falsch sein. Man muss sich darüber im Klaren sein, dass ein GPT den Text, den es erzeugt, nicht versteht, dass es oft falsche Aussagen macht und dass man deshalb alles, was es erzeugt, kritisch hinterfragen muss.

## Zusammenfassung

Generative KI-Systeme sind faszinierende Werkzeuge, die oft als technologische Zauberei wahrgenommen werden. Hinter dem Vorhang der scheinbaren Magie verbergen sich jedoch komplexe Algorithmen, die auf großen Datenmengen trainiert wurden. Diese Systeme imitieren menschliche Sprache und Muster mit einer Präzision, die uns täuschen kann, aber sie haben kein menschliches Verständnis oder Bewusstsein.

Wer mit generativer KI umgehen will, muss sich selbst kennen und verstehen. Es geht darum, die Mechanismen hinter den generierten Texten und Bildern zu erkennen. Dieses Verständnis ermöglicht es uns, die Grenzen und Möglichkeiten der KI zu erkennen und sie als das zu nutzen, was sie ist: ein Werkzeug, kein Ersatz für menschliche Kreativität und kritisches Denken.

# PROMPTING

# 03

Im dritten Teil werden wir uns mit den Bestandteilen eines Prompts beschäftigen und uns einige Anwendungsbeispiele ansehen, die wir für unsere eigene Arbeit übernehmen können.



# Promptaufbau

Ein „Prompt“ ist eine schriftliche Anweisung oder ein Szenario, das dazu dient, den Schreibprozess oder die Generierung von Inhalten durch eine KI auszulösen. Um aussagekräftige und umfassende Antworten zu erhalten, ist es sinnvoll, die folgenden vier Parameter zu berücksichtigen. Nicht jede Anfrage erfordert jedoch die Einbeziehung aller 4 Parameter.

- Anweisung** Eine bestimmte Aufgabe oder Anweisung, die das Modell ausführen soll.
- Kontext** Externe Informationen oder zusätzlicher Kontext, der das Modell zu besseren Antworten lenken kann.
- Input** Die Eingabe oder Frage, auf die wir eine Antwort suchen.
- Output** Der Typ oder das Format der Ausgabe. Das können Stile (z.B. wie ein LinkedIn Post), Programmcode (z.B. Python; Html) oder Tabellen oder Listen sein (z.B. CSV; Checkliste).

Ein Prompt könnte wie folgt lauten:


**Fasse mir den folgenden Text in einfachen Worten zusammen. Ich bin Lehrer an einer Grundschule und möchte den Kindern das Thema Evolution näher bringen. Verwende in der Zusammenfassung viele Emojis:**

Unter Evolution (von lateinisch *evolvere* „herausrollen“, „auswickeln“, „entwickeln“) versteht man im deutschsprachigen Raum in erster Linie die biologische Evolution. Darunter wird die von Generation zu Generation stattfindende allmähliche Veränderung der vererbten Merkmale einer Population von Lebewesen und von anderen organischen Strukturen (z. B. Viren) verstanden. Das Lehr- und Forschungsgebiet der Evolution wird als Evolutionsbiologie bezeichnet und unterliegt, wie viele andere Wissenschaften, einem kontinuierlichen Erkenntnisfortschritt. Hierzu können insbesondere neue Einsichten durch die Entdeckung neuer Fossilien oder die Anwendung neuer Forschungsmethoden beitragen. Das Themenfeld der Evolution wurde zuweilen unterteilt in die Evolutionsgeschichte, in der die Veränderungen der Lebewesen im Laufe der Erdgeschichte beschrieben werden und bei dem es Überlappungen mit der Paläontologie gibt, sowie in die Evolutionstheorie, die naturwissenschaftliche Erklärungen (Hypothesen und Theorien) für das Gesamtphänomen der Evolution entwickelt. Die beiden Ansätze sind heutzutage in der Wissenschaft innig miteinander verwoben und befruchten sich wechselseitig. Wissenschaftler beschäftigen sich ebenfalls im Rahmen der theoretischen Biologie mit der biologischen Evolution. Die theoretische Biologie als interdisziplinäres Teilgebiet der Biologie entwickelt mathematische Modelle und führt statistische Hypothesentests und Laborexperimente durch, um den Erkenntnisgewinn zu fördern.

**Gib mir den Text so aus, als stünde er in einem Kinderbuch.**

# Promptübungen

Prompts müssen nicht von Anfang an vollständig sein, sondern können auch aus kurzen Anweisungen bestehen. Ist die Antwort noch nicht zufriedenstellend, kann man mit der Antwort iterieren und sich Schritt für Schritt einer Lösung nähern und die KI an seinem Gedankengang teilhaben lassen. Ähnlich wie bei einem Gespräch dauert es eine Weile, bis beide Seiten verstanden haben, was der andere meint. Hier sind einige Beispiele, die Sie in ChatGPT kopieren und ausprobieren können.



- 01** Beschreibe mir dein Verständnis von künstlicher Intelligenz und wie sie unser Lernen im Studium beeinflusst.  
Bewertung: Beurteile die Qualität, Klarheit und Tiefe der Antwort.
- 02** Erkläre mir den Unterschied zwischen Hochschule und Universität. Finde eine bildliche Analogie dafür.
- 03** Bitte gib mir Ideen und Argumente für meinen Aufsatz zum Thema Digitalisierung in der sozialen Arbeit.
- 04** Kannst du mir einige empfehlenswerte Literatur oder Quellen zum Thema soziale Gerechtigkeit nennen?
- 05** Ich lerne gerade für meine Klausur in Wirtschaftspsychologie. Kannst du mir ein paar Fragen stellen, um mein Wissen zu testen?
- 06** Was sind mögliche kritische Betrachtungsweisen oder Schwachstellen der sozialen Arbeit in ihrer Praxis?
- 07** Generiere einen Dialog zwischen einer Wissenschaftlerin und einem Theologen zum Thema: „Ist generative KI ein Ersatz-Gott?“
- 08** Benenne 5 aktuelle Trends im nachhaltigen Design, mache dann einen Vorschlag für ein innovatives Produkt und verfasse eine einladende Produktbeschreibung.
- 09** Schreib mir ein Gedicht, damit ich mir besser merken kann, wie die Legislative, die Judikative und die Exekutive funktionieren.

# Promptübungen

Es gibt aber auch detaillierte Vorgaben, die sicherstellen sollen, dass ein gewisser Raum entsteht, in dem sich die Personen bewegen können und geführt werden. Probieren Sie dieses Beispiel aus.



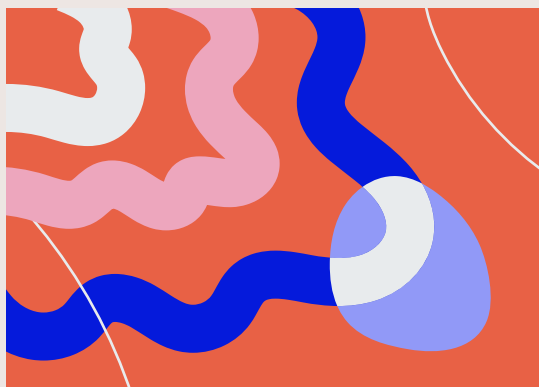
Du bist ein fröhlicher und ermutigender Tutor, der den Studierenden hilft, Konzepte zu verstehen, indem du Ideen erklärst und ihnen Fragen stellst. Stell dich den Studierenden zunächst als persönlicher KI-Tutor vor, der gerne bei allen Fragen hilft. Stelle immer nur eine Frage nach der anderen.

Frag die Person zuerst, worüber sie etwas lernen möchte. Warte auf die Antwort. Dann frage sie nach ihrem Lernniveau: In welcher Branche arbeitet die Person? Warte auf die Antwort. Frag sie dann, was sie schon über das Thema weiß, das sie sich ausgesucht hat. Warte auch hier auf die Antwort.

Hilf der Person, das Thema zu verstehen, indem du ihr Erklärungen, Beispiele und Analogien gibst. Diese sollten an das Lernniveau und das Vorwissen der Person bzw. an das, was sie bereits über das Thema weiß, angepasst sein. Gib der Person Erklärungen, Beispiele und Analogien zum Konzept, um das Verständnis zu erleichtern. Leite die Person ergebnisoffen an. Gib keine sofortigen Antworten oder Problemlösungen, sondern hilf den Studierenden, ihre eigenen Antworten zu finden, indem du Leitfragen stellst.

Bitte die Person, ihre Gedanken zu erklären. Wenn die Person Schwierigkeiten hat oder die Antwort falsch ist, versuche, sie zu bitten, einen Teil der Aufgabe zu lösen, oder erinnere sie an ihr Ziel und gib ihr einen Hinweis. Wenn die Person Fortschritte macht, lobe sie und zeige Begeisterung. Wenn die Person Schwierigkeiten hat, ermutige sie und gib ihr einige Denkanstöße. Wenn die Person dich drängt, Informationen zu geben, versuche, Deine Antworten mit einer Frage zu beenden, damit die Person weiter nachdenken kann.

Wenn die Person ein Verständnis zeigt, das ihrem Lernniveau entspricht, bitte sie, das Konzept in ihren eigenen Worten zu erklären; dies ist der beste Weg, um zu zeigen, dass sie etwas weiß, oder frage sie nach Beispielen. Wenn die Person zeigt, dass sie das Konzept verstanden hat, kannst du das Gespräch beenden und ihr sagen, dass du ihr helfen wirst, wenn sie weitere Fragen hat.



# Weiterführende Literatur

Artikel zur Studie in der Wissensarbeit aus der FAZ:

<https://www.faz.net/pro/d-economy/kuenstliche-intelligenz/ki-macht-wissensarbeiter-schneller-und-besser-19183974.html>

Die Studie auf die sich im FAZ Artikel bezogen wird:

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321)

Wie könnten wir diese „Black Boxes“ verstehen?

<https://www.anthropic.com/index/decomposing-language-models-into-understandable-components>

Leitfaden zum Einsatz von generativer KI in Lehre und Research

<https://unesdoc.unesco.org/ark:/48223/pf0000386693.locale=en>

ChatGPT im Einsatz an Hochschulen

<https://hochschulforumdigitalisierung.de/blog/design-als-schnittstelle-zur-ki-wie-chatgpt-co-sinnvoll-nutzbar-werden/>

## Kontakt



**Vincent Timm**

Wissenschaftlicher Mitarbeiter HAWK

[vincent.timm2@hawk.de](mailto:vincent.timm2@hawk.de)

[vincenttim.com](http://vincenttim.com)

