

Tobias Albers-Heinemann & Lukas Spahlinger

Künstliche Intelligenz

verstehen und kompetent nutzen



WIR
mit Ki

2. Auflage



Zentrum Bildung und Gesellschaft
Evangelische Kirche in Hessen und Nassau

Herausgeber:

Tobias Albers-Heinemann & Lukas Spahlinger
Zentrum Bildung und Gesellschaft der EKH, Albert-Schweitzer 113-115, 55128 Mainz

Dieses Buch entstand im Rahmen des Projekts „AI Impact – Wir mit KI“, das vom Land Hessen im Rahmen von HessenCampus 2025 gefördert wurde.

2. Auflage
Erscheinungsjahr 2026

Illustrationen:

Benjamin Bertram (S.2)
notebookLM

Lizenz:

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International Lizenz (CC BY-SA 4.0)

Hinweis zur Lizenz:

Sie dürfen das Werk vervielfältigen, verbreiten und bearbeiten, sofern Sie die Urheber*innen nennen und das neue Werk unter denselben Lizenzbedingungen weitergeben.



Zentrum Bildung und Gesellschaft
Evangelische Kirche in Hessen und Nassau

Inhaltsverzeichnis

Vorwort	1
Einstieg in die Welt der Künstlichen Intelligenz.....	3
Prompten als angewandte Gesprächsführung: Was wir von Watzlawick und Schulz von Thun lernen können ...	7
KI-Sprachmodelle anpassen: Wie wir Ziele präziser erreichen und bessere Texte schreiben	12
KI, Vorurteile und Verantwortung – Wie wir KI diskriminierungssensibel nutzen können	16
Künstliche Intelligenz und der Datenschutz	21
KI, Urheberrecht und Kennzeichnungspflicht	25
Menschsein im KI-Zeitalter – Bildungsaufgaben angesichts eines veränderten Menschenbilds	29
Von piepsenden Tamagotchis und redseligen Robotern: eine kleine Geschichte der Mensch-Maschine-Interaktion	31
Wie Anthropomorphisierung von KI unsere Wahrnehmung von Maschinen verändert	34
Vertrauen, Beziehung und Reflexion: Medienpsychologische Perspektiven für die KI-Kompetenzbildung	39
Künstliche Intelligenz in der Familie.....	43
Künstliche Intelligenz und Einsamkeit.....	47
Künstliche Intelligenz und Nachhaltigkeit – Eine Perspektive der Bildung für nachhaltige Entwicklung	51
Digitaler Kolonialismus, globale Gerechtigkeit und Künstliche Intelligenz	55
Das Prinzip Barrierearmut – KI als Gestaltungsfaktor für Teilhabe	59
Einfache Sprache mit künstlicher Intelligenz	63
Lokale KI-Modelle und Open Source: Chancen, Grenzen und Relevanz für die Bildung	66
Prüfungsformate im Zeitalter von KI	69
KI-Methoden in der Bildungsarbeit	73
Ausblick: KI-Kompetenz in Zeiten gesellschaftlicher Transformation	79

VORWORT

Digitale Technologien verändern unseren Alltag, unsere Arbeit und unsere Kommunikation in rasantem Tempo – und wir alle spüren in unserer täglichen Arbeit, wie tiefgreifend diese Digitalisierungsprozesse auf unterschiedlichen Ebenen wirken. Arbeitsabläufe, Kommunikationswege und Lernformate verändern sich kontinuierlich. Wir bewegen uns dabei zwischen dem Staunen über täglich neu erscheinende KI-Tools und deren beeindruckende Möglichkeiten auf der einen Seite sowie Sorgen oder Ängsten vor Datenschutzproblemen und ungewollten Veränderungen von Arbeitsprozessen oder dem Wegfall ganzer Arbeitsfelder auf der anderen.

In der aktuellen Debatte um Künstliche Intelligenz (KI) entsteht zudem oft der Eindruck, es handele sich – ähnlich wie bei Apps und Social-Media-Diensten wie WhatsApp, TikTok, Instagram oder Snapchat – um eine Frage des persönlichen Nutzens oder Nicht-Nutzens. Doch dieser Vergleich greift zu kurz: Künstliche Intelligenz durchdringt als Form der Digitalisierung so viele Ebenen unseres Alltags, dass wir ihr häufig begegnen, ohne es überhaupt zu wissen. Es stellt sich also gar nicht die oft diskutierte Frage, ob wir KI nutzen wollen oder nicht, sondern vielmehr wie wir sie nutzen wollen und welche Rolle sie einnehmen soll.

Die Künstliche Intelligenz hat sich in den vergangenen Jahren von einer Nischentechnologie zu einem prägenden Thema entwickelt, das aus unserem Alltag kaum noch wegzudenken ist. Damit verbunden stellt sich nicht nur die Frage nach unserer persönlichen Haltung, sondern auch nach der Verantwortung der Bildungsakteure. Unabhängig von unserem Kompetenz- und Kenntnisstand müssen wir uns als mündige Bürger*innen einer digitalisierten Gesellschaft darin positionieren – auf der individuellen Ebene ebenso wie in professionellen Kontexten. Auf der persönlichen Ebene betrifft dies unser alltägliches Leben – ob wir es bewusst wahrnehmen oder nicht, sei es durch die Integration von KI in unsere Smartphones, in Online-Dienste wie die Google-Suche, in smarte Gegenstände zu Hause oder aber auch in Bewerbungsverfahren, Kreditvergabesystemen oder Telefonhotlines.

Zugleich stehen die Akteur*innen im Bildungsbereich in besonderer Verantwortung, Menschen auf diese umfassenden Entwicklungen vorzubereiten und Räume für Auseinandersetzung zu schaffen. Zudem verpflichtet seit Februar 2025 die Europäische KI-Verordnung in Artikel 4 KI-nutzende Einrichtungen dazu, KI-Kompetenz systematisch zu vermitteln, was zusätzliche institutionelle Anforderungen formuliert.

KI-Kompetenz meint dabei weit mehr als eine bloße Klickanleitung für populäre KI-Dienste. Sie umfasst sowohl ein grundlegendes technisches Verständnis und die Fähigkeit, Funktionsweisen kritisch einzuordnen, als auch die Kompetenz, Chancen und Risiken zu bewerten, gesellschaftliche und ethische Fragen zu reflektieren sowie KI verantwortungsvoll in der eigenen Arbeit einzusetzen.

AI IMPACT – WIR MIT KI

Mit dem vom Land Hessen geförderten Projekt „AI Impact – Wir mit KI“ konnten wir uns im Jahr 2025 intensiv mit den gesellschaftlichen, kulturellen und bildungsbezogenen Auswirkungen dieser Technologie auseinandersetzen. In Workshops, Fachtagen, Online-

Kursen und Diskussionsformaten sind vielfältige Perspektiven zusammengekommen – von Expert*innen, Multiplikator*innen und Menschen aus ganz unterschiedlichen Lebensbereichen.

Diese Veröffentlichung bündelt daher die wichtigsten Ergebnisse und Materialien, die unserer Auffassung nach der Auseinandersetzung mit dem Themenfeld KI dienlich sind. Sie enthält Grundlagenwissen zur Funktionsweise und Anwendung von KI sowie praxisorientierte Methoden, die den Einsatz in der Bildungsarbeit erleichtern. Zudem werden rechtliche Rahmenbedingungen wie Datenschutz, Urheberrecht und Kennzeichnungspflichten beleuchtet. Gesellschaftliche Herausforderungen werden kritisch reflektiert, etwa die neuen Anforderungen, die KI in Bereichen wie Familie, Beratung oder der Begleitung einsamer Menschen mit sich bringt, ebenso wie Fragen einer diskriminierungssensiblen Nutzung von KI. Gleichzeitig werden Chancen aufgezeigt, wie KI Barrieren abbauen, Inklusion fördern und neue Formen der Zusammenarbeit ermöglichen kann. Abgerundet wird dies durch Anregungen für eine kreative, verantwortungsbewusste und zielgruppenorientierte Gestaltung von Bildungsarbeit mit KI.

Unser Ziel ist es, Orientierung zu geben, Wissen zugänglich zu machen und zur kritischen Auseinandersetzung einzuladen. Künstliche Intelligenz kann wertvolle Unterstützung leisten – vorausgesetzt, ihr Einsatz wird bewusst geplant, transparent gestaltet und an den Bedürfnissen der Menschen ausgerichtet.



Auf unserer Website wir-mit-ki.de bieten wir dieses Buch als kostenlose PDF-Datei an – auch in Einfacher Sprache. Darüber hinaus besteht die Möglichkeit, sich über unsere Blogbeiträge, Wimmelbilder und anderen Materialien in die Thematik zu vertiefen. Gerne können Sie uns hier auch ein Feedback hinterlassen.

Lukas Spahlinger & Tobias Albers-Heinemann



EINSTIEG IN DIE WELT DER KÜNSTLICHEN INTELLIGENZ

von Tobias Albers-Heinemann

Künstliche Intelligenz (KI) ist mittlerweile so tief im täglichen Leben verankert, dass jede Person früher oder später mit ihr in Kontakt kommt – oft sogar ohne es bewusst zu merken. Für Bildungsakteure wird es daher zunehmend wichtiger, die Funktionsweise und Möglichkeiten von KI zu verstehen – und sie reflektiert in ihre Arbeit zu integrieren.

Damit verbunden ist der Aufbau sogenannter **KI-Kompetenz** – also der Fähigkeit, KI-Systeme bewusst, kritisch und konstruktiv zu nutzen. Ein zentrales Element von KI-Kompetenz ist die Fähigkeit, eigene Ziele und Erwartungen in der Arbeit mit KI zu formulieren und diese im Licht der gesetzlichen Rahmenbedingungen einzuordnen. Wer mit KI arbeitet, muss klar benennen können, was er oder sie erreichen möchte – etwa einen informativen Text, eine kreative Idee oder ein passendes Kommunikationsformat. Diese Kompetenz lässt sich nicht auslagern: Sie entsteht durch aktives Ausprobieren, kritisches Hinterfragen und kontinuierliches Lernen im Umgang mit den neuen Werkzeugen.

Darüber hinaus gehört zu KI-Kompetenz ein grundlegendes Verständnis davon, was Künstliche Intelligenz eigentlich ist und wie sie funktioniert. Gerade generative KI-Modelle wie ChatGPT, Mistral, Gemini oder Claude, die sehr schnell mit beeindruckenden Texten auf unsere Fragen und Eingaben reagieren, werfen viele Fragen auf: Wie entstehen die Antworten? Warum sind wir mit diesen nicht immer zufrieden? Was können wir tun, um zufrieden zu werden? Und wo liegen die Grenzen dieser Systeme? Diese Fragen sind kein technisches Detail, sondern Teil der Bildungsarbeit – und damit auch Teil von KI-Kompetenz.

WAS IST KI – UND WAS NICHT?

KI hat durch Medienberichte über allwissende Superintelligenzen oder gefährliche Maschinen sowie durch dystopische Darstellungen in Hollywood-Filmen und Büchern bereits ein eher negatives Image erhalten. Viele dieser Darstellungen sind fiktiv oder stark spekulativ – und weichen deutlich von den tatsächlichen technischen Möglichkeiten ab. Umso wichtiger ist es, sich nüchtern und sachlich damit auseinanderzusetzen, was KI tatsächlich ist – und was nicht.

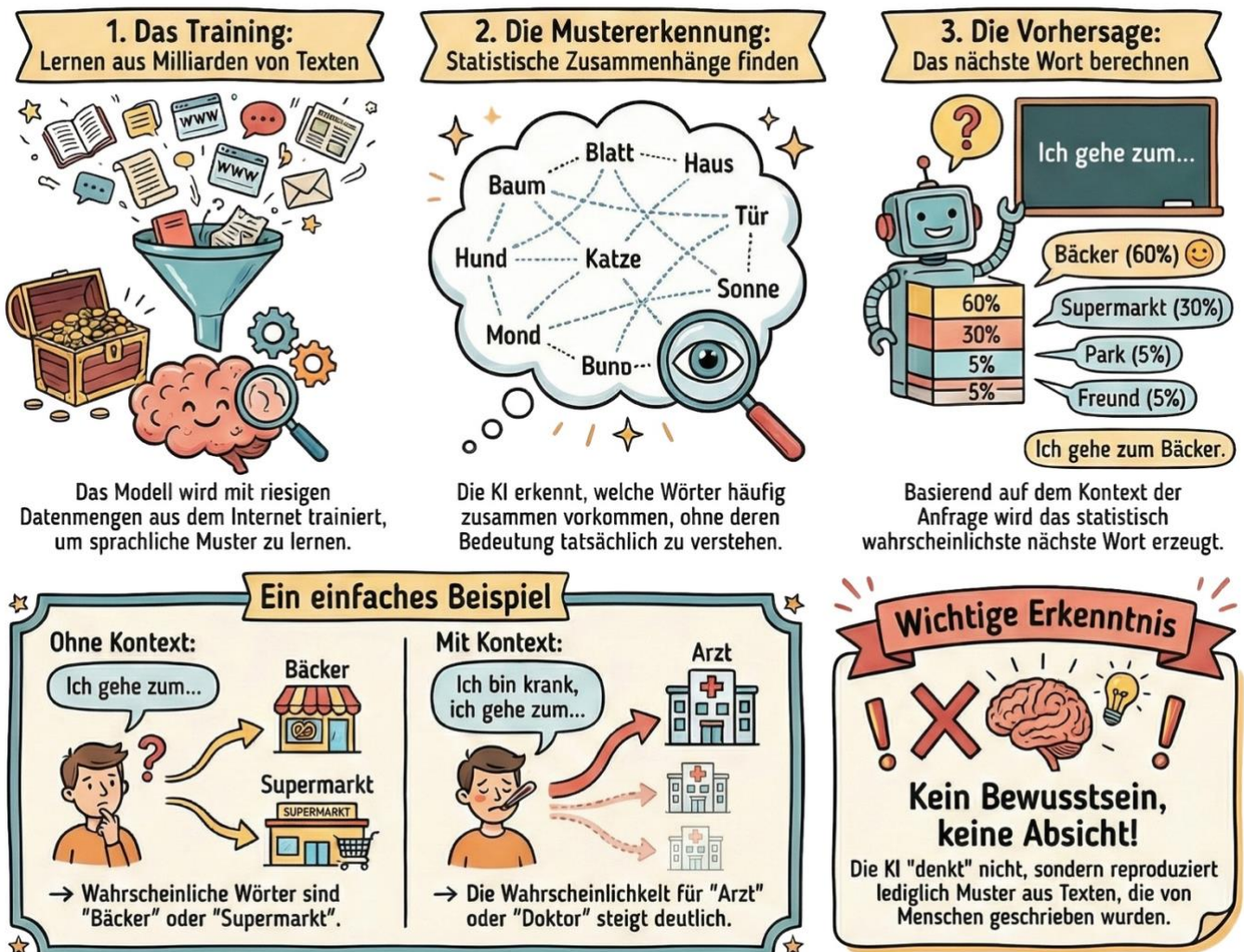
Der Begriff „Künstliche Intelligenz“ bezeichnet in erster Linie Systeme, die Aufgaben übernehmen, die zuvor menschliche Intelligenz erforderten, z. B. Spracherkennung, Textproduktion oder Datenanalyse. In der öffentlichen Debatte wird dabei oft vor allem an ChatGPT oder Bildgeneratoren gedacht. KI umfasst jedoch weit mehr – sie ist eine technische Entwicklung innerhalb der Digitalisierung mit vielfältigen Formen und Anwendungen. KI ist allerdings keineswegs „intelligent“ im menschlichen Sinn: Sie hat kein Bewusstsein, keine Empathie und keine eigenen Absichten. Vielmehr handelt es sich um rechnergestützte Verfahren, die Wahrscheinlichkeiten berechnen und so möglichst passende Antworten oder Ergebnisse erzeugen. Ein realistisches Verständnis dieser Funktionsweise und ihrer Grenzen ist entscheidend, um KI verantwortungsvoll einschätzen und nutzen zu können.

Gleichzeitig ist das Spektrum von Künstlicher Intelligenz enorm: Es reicht von Navigationssystemen, automatisierter Übersetzung, Bilderkennung und medizinischer Diagnostik bis hin zu Robotik, Sprachassistenten, Überwachungs- und Empfehlungssystemen. Aufgrund dieser Vielfalt müssen Anwendungen je nach Bereich

unterschiedlich bewertet werden – mal als Chance, mal mit Blick auf erhebliche Risiken. In diesem Buch konzentrieren wir uns jedoch bewusst auf generative KI-Modelle, insbesondere Large Language Models (LLMs), weil sie durch ihre Fähigkeit der Text-, Bild- und Videoproduktion sowie der Analyse und Bearbeitung komplexer Daten für die Bildungsarbeit und den Aufbau von KI-Kompetenz derzeit die größte Relevanz haben.

WIE FUNKTIONIEREN GENERATIVE KI-MODELLE?

Generative Modelle wie ChatGPT, Claude oder Gemini erzeugen neue Inhalte – etwa Texte, Bilder oder Audiodateien – auf der Grundlage gelernter Muster. Sie werden mit Milliarden von Texten und anderen Daten trainiert und berechnen bei jeder Eingabe Wort für Wort, Pixel für Pixel oder Ton für Ton die wahrscheinlichste Fortsetzung. Auf diese Weise entstehen scheinbar eigenständige Inhalte, die jedoch ausschließlich statistisch erzeugt sind.



Um diesen Mechanismus besser zu verstehen, lohnt ein Blick auf den Begriff GPT, den viele aus „ChatGPT“ kennen. GPT steht für „**Generative Pretrained Transformer**“ – ein vortrainiertes Sprachmodell, das auf der sogenannten Transformer-Architektur basiert.

Grundlage solcher Modelle sind enorme Datenmengen, aus denen Muster und Strukturen abgeleitet werden. Weil die Datenbasis so groß ist, spricht man von „**Large Language Models**“ (LLMs). Diese gehören zu den generativen KI-Modellen und sind darauf ausgelegt, auf Basis dieser riesigen Datenmengen Muster zu erkennen und daraus neue Inhalte zu erzeugen. Dabei ist wichtig zu wissen: Ein Anbieter wie zum Beispiel ChatGPT kann unterschiedliche Large Language Models einsetzen (GPT-4o-mini, GPT-4o, GPT-5 ...), die in verschiedener Qualität trainiert wurden – weshalb auch die Ergebnisse inhaltlich und sprachlich deutlich variieren. Aus diesem Funktionsprinzip ergeben sich für die generativen Text-KIs typische Stärken und Schwächen:

- Sie können schnell plausible Inhalte wie zum Beispiel Texte, Bilder, Videos etc. erzeugen.
- Sie verfügen jedoch über kein eigenes Wissen, sondern erzeugen sprachliche Wahrscheinlichkeiten.
- Sie liefern mitunter überzeugende, aber falsche Informationen – sogenannte Halluzinationen.

Ein Beispiel: Auf „*Ich gehe zum ...*“ könnte die generative Text-KI Wörter wie „*Bäcker*“ oder „*Supermarkt*“ ergänzen, weil diese im Training häufig in ähnlichen Zusammenhängen vorkamen. Wird der Kontext verändert, etwa zu „*Ich bin krank, ich gehe zum ...*“, steigt die Wahrscheinlichkeit deutlich für Begriffe wie „*Arzt*“ oder „*Doktor*“. Dieses einfache Beispiel zeigt, wie stark der Kontext die Textgenerierung beeinflusst. Die KI berechnet dabei keine Bedeutung im menschlichen Sinn, sondern lediglich statistische Wahrscheinlichkeiten dafür, welche Wörter aufeinander folgen. Gerade hier wird deutlich: Die Ergebnisse hängen stark davon ab, mit welchen Daten die Modelle trainiert wurden. Da diese Daten überwiegend aus dem globalen Norden stammen, reproduzieren viele Modelle westliche Werte und Stereotype. Dadurch können Diskriminierungen oder Einseitigkeiten entstehen, etwa bei Geschlechterrollen oder kulturellen Darstellungen. Deshalb ist eine kritische Perspektive auf die Datenbasis unverzichtbar. Genauer werden wir uns mit diesem Punkt aber im nächsten Kapitel beschäftigen.

PROMPTING: DER SCHLÜSSEL ZUR QUALITÄT

Nachdem wir gesehen haben, wie stark die Ergebnisse von den Trainingsdaten geprägt sind und welche gesellschaftlichen Verzerrungen dadurch entstehen können, lohnt nun der Blick auf einen zweiten entscheidenden Faktor: die Eingabe selbst. Denn neben den Daten, auf denen Modelle basieren, bestimmt auch die Art und Weise, wie wir mit ihnen kommunizieren, maßgeblich die Qualität der Resultate. Das Ergebnis eines KI-Systems hängt daher entscheidend vom sogenannten **Prompt** ab – also von der Art und Weise, wie die Eingabe formuliert ist. Ein Prompt ist im Grunde eine Aufgabenstellung oder Frage, die der KI vorgibt, in welcher Rolle sie agieren, welchen Kontext sie berücksichtigen und welches Ergebnisformat sie liefern soll. Je präziser und kontextreicher ein Prompt, desto passgenauer und brauchbarer sind in der Regel die Antworten. Zwei einfache Beispiele verdeutlichen den Unterschied:

- „Erkläre KI für Grundschulkinder.“ → ergibt eine einfache, kindgerechte Erklärung mit Beispielen.

- „Definiere KI im Stil eines wissenschaftlichen Abstracts.“ → führt zu einem stark verdichteten, akademischen Text mit Fachbegriffen.

Gutes Prompting ist keine technische, sondern eine kommunikative Fähigkeit. Es erfordert die bewusste Auswahl von Perspektive, Zielgruppe und Darstellungsform. Hilfreich können zum Beispiel Angaben sein wie:

- Rolle („Du bist Sprachwissenschaftler*in“),
- Zielgruppe („für Jugendliche ab 14“),
- gewünschte Form („in einfacher Sprache, als Liste mit 5 Punkten, als Liedtext“ etc.)

Darüber hinaus gehört zu gelungenem Prompting die Klarheit über die eigenen Ziele: Möchte ich einen informativen Überblick, ein praktisches Beispiel, einen kreativen Impuls oder eine wissenschaftliche Definition? Erst wenn diese Fragen klar sind, lässt sich ein Prompt so formulieren, dass die KI sinnvolle und nützliche Ergebnisse liefert. Prompts lassen sich nicht einfach von außen vorgeben, sondern müssen individuell entwickelt werden. So entstehen Ergebnisse, die passgenau zum eigenen Anliegen und zum jeweiligen Kontext passen.

PRAKTISCHER EINSTIEG



Zum Kennenlernen eignen sich niederschwellige Tools wie Duck.AI, die ohne Anmeldung direkt datensparsam im Browser laufen. Ein einfacher erster Prompt könnte sein: *„Schreibe ein unterhaltsames Gedicht über Computer in der Schule im Stil von Goethe“*. Variiere danach die Formulierung oder die Sprache und beobachte, wie sich die Ergebnisse verändern. Der erste Schritt ist, einfach auszuprobieren.

PROMPTEN ALS ANGEWANDTE GESPRÄCHSFÜHRUNG: WAS WIR VON WATZLAWICK UND SCHULZ VON THUN LERNEN KÖNNEN

von Tobias Albers-Heinemann

Die Fähigkeit, effektiv mit generativen KI-Modellen zu kommunizieren – das sogenannte Prompten – wird oft als rein technisches Skill wahrgenommen. Influencer und Tech-Enthusiasten preisen die „richtigen Prompts“ als Geheimwaffe für Produktivität und Effizienz an. Das zeigen auch Schulungen und Kurse, in denen darum gebeten wird, die richtigen Prompts für ein bestimmtes Ergebnis mitzuteilen. Doch diese Sichtweise greift zu kurz. Prompten ist als kommunikative Basis der Mensch-Maschine-Interaktion vielmehr eine Form der angewandten Gesprächsführung, die gelernt und geübt werden muss – ähnlich wie zwischenmenschliche Kommunikation. Aktuelle Forschung bestätigt, dass Prompt Engineering nicht nur technisches Know-how erfordert, sondern auch kommunikative Kompetenzen.

In diesem Text geht es nicht um einen 1:1-Vergleich zwischen Prompten und kommunikationswissenschaftlichen Theorien. Vielmehr zeigen Prinzipien von Kommunikationswissenschaftlern wie Paul Watzlawick und Friedemann Schulz von Thun, dass Ideen aus dem zwischenmenschlichen Bereich auch für die Interaktion mit KI-Systemen hilfreich sein können – trotz der offensichtlichen Unterschiede. KI kommuniziert nicht im eigentlichen Sinne, sondern simuliert kommunikative Muster, die Menschen als dialogische Antworten interpretieren. Gerade weil Nutzer*innen diese Antworten wie Äußerungen eines Gegenübers deuten, entstehen Kommunikationsdynamiken, die sich mit kommunikationswissenschaftlichen Modellen erklären lassen – obwohl die KI kein bewusstes Subjekt ist und weder Wissen noch Absichten besitzt.

WATZLAWICKS AXIOME UND DIE KUNST DES PROMPTENS

Um die Dynamik des Promptens besser zu verstehen, lohnt sich ein Blick auf die fünf Axiome von Paul Watzlawick. Diese grundlegenden Prinzipien der menschlichen Kommunikation – oft als Leitsätze oder selbstverständliche Annahmen formuliert – bieten wertvolle Einblicke, wie wir mit KI-Systemen interagieren. Sie zeigen, warum manche Prompts besser funktionieren als andere und wie wir unsere Kommunikation mit KI optimieren können.

1. „Man kann nicht nicht kommunizieren“

Watzlawicks erstes Axiom besagt, dass jede Handlung – oder auch das Unterlassen einer Handlung – eine Botschaft sendet. Im zwischenmenschlichen Kontext umfasst dies sowohl verbale als auch nonverbale und paralinguistische Signale wie Gestik, Mimik oder Tonfall. Bei der Interaktion mit aktuell genutzten generativen KI-Modellen beschränkt sich die Kommunikation auf die verbale und schriftliche Ebene, da nonverbale Phänomene in dieser Form der Interaktion aktuell (noch) keine Rolle spielen.

Übertragen auf das verbale und schriftliche Prompten bedeutet das: Jeder Input, den wir einer KI geben, ist eine Form der Kommunikation. Gibt man einer generativen KI keine spezifischen Anweisungen, reagiert diese auf der Grundlage vordefinierter Systemprompts und des gegebenen Interaktionskontexts – etwa mit einer freundlichen Begrüßung oder einem Angebot allgemeiner Hilfsbereitschaft. Diese Voreinstellungen sind bereits eine Form der Kommunikation und steuern das Verhalten der KI. Selbst eine scheinbar einfache

Eingabe wie das Wort „Hallo“ wird von der KI als Aufforderung interpretiert, darauf zu reagieren – sei es mit einer Begrüßung oder einer Frage nach weiteren Anweisungen.

Studien¹ zum Prompt Engineering zeigen, dass selbst implizite Prompts – also Systemvorgaben oder Kontextinformationen – das Verhalten des Modells steuern. Nutzer*innen sollten sich bewusst sein, dass jeder Input – oder sogar das Fehlen eines expliziten Prompts – eine Botschaft an die KI sendet. Klare und präzise Prompts führen zu besseren und zielgerichteteren Ergebnissen.

2. Inhalts- und Beziehungsaspekt

Jede Kommunikation hat zwei Ebenen: den Inhaltsaspekt, also das, was gesagt wird, und den Beziehungsaspekt, also die Art und Weise, wie etwas gesagt wird. Beim Prompten enthält jeder Prompt sowohl einen Sachinhalt, wie etwa „Erkläre mir das Thema X in 200 Wörtern“, als auch eine Beziehungsaussage im Sinne einer Rollenzuweisung, wie zum Beispiel „Du bist meine Expertin“.

Rollenformulierungen wie „Agier als Coach“ oder „Sprich mit mir wie mit einem Einsteiger“ verändern das Antwortverhalten der KI deutlich. Die Art und Weise, wie wir einen Prompt formulieren, beeinflusst somit nicht nur den Inhalt der Antwort, sondern auch deren Ton und Tiefe.

3. Zirkulärer Ursache-Wirkungs-Charakter

Kommunikation ist ein kreisförmiger Prozess, bei dem jede Reaktion eine Antwort auf eine vorherige Botschaft ist. In Dialog-Interfaces mit KI wird dies besonders deutlich: Jede Antwort der KI reagiert auf vorherige Prompts und setzt den Rahmen für folgende Nachfragen. Ein unklarer Prompt führt zu einer unklaren Antwort, die wiederum zu weiteren unklaren Nachfragen führen kann. Der gesamte Chatverlauf ist eine zirkuläre Interaktion, keine lineare Frage-Antwort-Reihe². Nutzer*innen sollten Prompts als Teil eines fortlaufenden Dialogs betrachten und bei Bedarf nachschärfen oder kontextualisieren.

4. Analog und digital

Neben der „digitalen“ Wortwahl sind auch „analoge“ Signale wie Format-Marker, Listen, Emojis oder typografische Hervorhebungen wichtig. Diese Signale wirken als starke Kontextsignale und lenken den Output der KI in Stil und Struktur. Anders als beim Menschen, der solche Elemente emotional oder inhaltlich deuten kann, interpretiert die KI sie rein formal – als strukturelle Muster, die den Aufbau, die Hierarchie oder die Betonung der Antwort steuern. Die Nutzung von Aufzählungen, Fettdruck oder Emojis wird von der KI nicht als Ausdruck von Gefühlen oder Stimmungen verstanden, sondern als formale Anweisung zur Gestaltungslogik der Antwort. Diese funktionale Interpretation macht analoge Signale zu einem mächtigen Werkzeug im Prompten: Sie helfen, die Antwort der KI strukturierter, übersichtlicher und zielgerichteter zu³. Visuelle und strukturelle Elemente optimieren so die maschinelle Verarbeitung und verbessern die Qualität der Interaktion – auch wenn sie keine emotionale oder subjektive Bedeutung tragen.

¹ Song, Y. et al. (2024): A Communication Theory Perspective on Prompting Engineering Methods for Large Language Models. Verfügbar unter: <https://arxiv.org/abs/2310.18358> [Zugriff: 25.11.2025].

² Paul Watzlawick Institut. (o. J.): Die fünf Axiome der Kommunikation. Verfügbar unter: <https://www.paulwatzlawick.de/axiome.html> [Zugriff: 25.11.2025].

³ Schmidmaier, M. et. al. (2024): *Increasing Large Language Model Performance through Structural Prompting*. Verfügbar unter: <https://www.mmi.ifi.lmu.de/pubdb/publications/pub/schmidmaier2024increasinglarge/schmidmaier2024increasinglarge.pdf> [Zugriff: 25.11.2025].

5. Symmetrische und komplementäre Beziehungen

Watzlawick unterscheidet zwischen symmetrischen (gleichberechtigten) und komplementären (ergänzenden) Beziehungen. Beim Prompten können Nutzer*innen die KI als gleichberechtigten Sparringspartner ansprechen, etwa durch Formulierungen wie „Wir erarbeiten gemeinsam eine Lösung“, oder als komplementäres Gegenüber, zum Beispiel mit „Du bist die Expertin, ich bin der Lernende“. Untersuchungen belegen, dass die explizite Definition der Rolle der KI die Interaktion gezielt steuert: Systemprompts fungieren als „operationaler Bauplan“ für KI-Agenten und definieren Rolle, Verhalten und Antwortstil vor der eigentlichen Interaktion. Gut gestaltete Systemprompts sorgen für konsistente, zuverlässige und kontextgerechte Ausgaben, was die Performance in komplexen Abläufen signifikant verbessert⁴. Die Wahl der Beziehungsebene beeinflusst dabei maßgeblich die Dynamik der Interaktion und die Qualität der Ergebnisse.

SCHULZ VON THUNS KOMMUNIKATIONSQUADRAT UND PROMPTS.

Neben den Axiomen von Paul Watzlawick bieten auch andere kommunikationswissenschaftliche Modelle wertvolle Ansätze, um die Interaktion mit KI-Systemen zu verstehen und zu optimieren. Besonders ergiebig ist hier das Vier-Seiten-Modell der Kommunikation von Friedemann Schulz von Thun, das nicht nur zwischenmenschliche Gespräche analysiert, sondern sich auch auf die Mensch-Maschine-Kommunikation übertragen lässt. Dieses Modell unterscheidet vier Ebenen jeder Botschaft: die Sachseite, die Selbstoffenbarung, den Beziehungshinweis und den Appell. Jede dieser Ebenen spielt auch beim Prompten eine zentrale Rolle und hilft, die Qualität der Interaktion mit KI-Systemen zu verbessern.

Die **Sachseite** eines Prompts enthält den eigentlichen Auftrag an die KI – also das Thema, den gewünschten Umfang und das Format der Antwort. Ein präzise formulierter Prompt wie „Erstelle eine Gliederung für einen Vortrag über Nachhaltigkeit mit fünf Hauptpunkten und jeweils drei Unterpunkten“ gibt der KI klare inhaltliche Vorgaben. Je konkreter die Sachseite beschrieben wird, desto zielgerichteter und nützlicher fallen die Antworten aus. Unklare oder vage Formulierungen führen dagegen oft zu unscharfem oder irrelevantem Output. Eine präzise Beschreibung der Aufgabe ist somit der erste Schritt zu klaren und verwertbaren Ergebnissen.

Die **Selbstoffenbarung** bezieht sich auf alle Informationen, die Nutzer*innen über sich selbst, ihre Rolle oder ihren Kontext preisgeben. Wenn eine Lehrkraft in der Erwachsenenbildung etwa angibt: „Ich habe keine Vorkenntnisse zu diesem Thema und benötige eine einfache Erklärung“, ermöglicht dies der KI, ihre Antworten an das Vorwissen und die Bedürfnisse der Nutzer*innen anzupassen. Selbstoffenbarung schafft Kontext und hilft der KI, den Ton, die Tiefe und die Komplexität der Antworten passend zu gestalten. Ohne solche Hinweise muss die KI Annahmen treffen, die möglicherweise nicht zur Zielgruppe passen. Selbstoffenbarung ist daher ein zentrales Element, um die Antworten der KI zielgruppengerecht und verständlich zu gestalten.

Der **Beziehungshinweis** definiert das Verhältnis zwischen Nutzer*in und KI und beeinflusst maßgeblich den Ton und die Art der Antwort. Formulierungen wie „Bitte erkläre

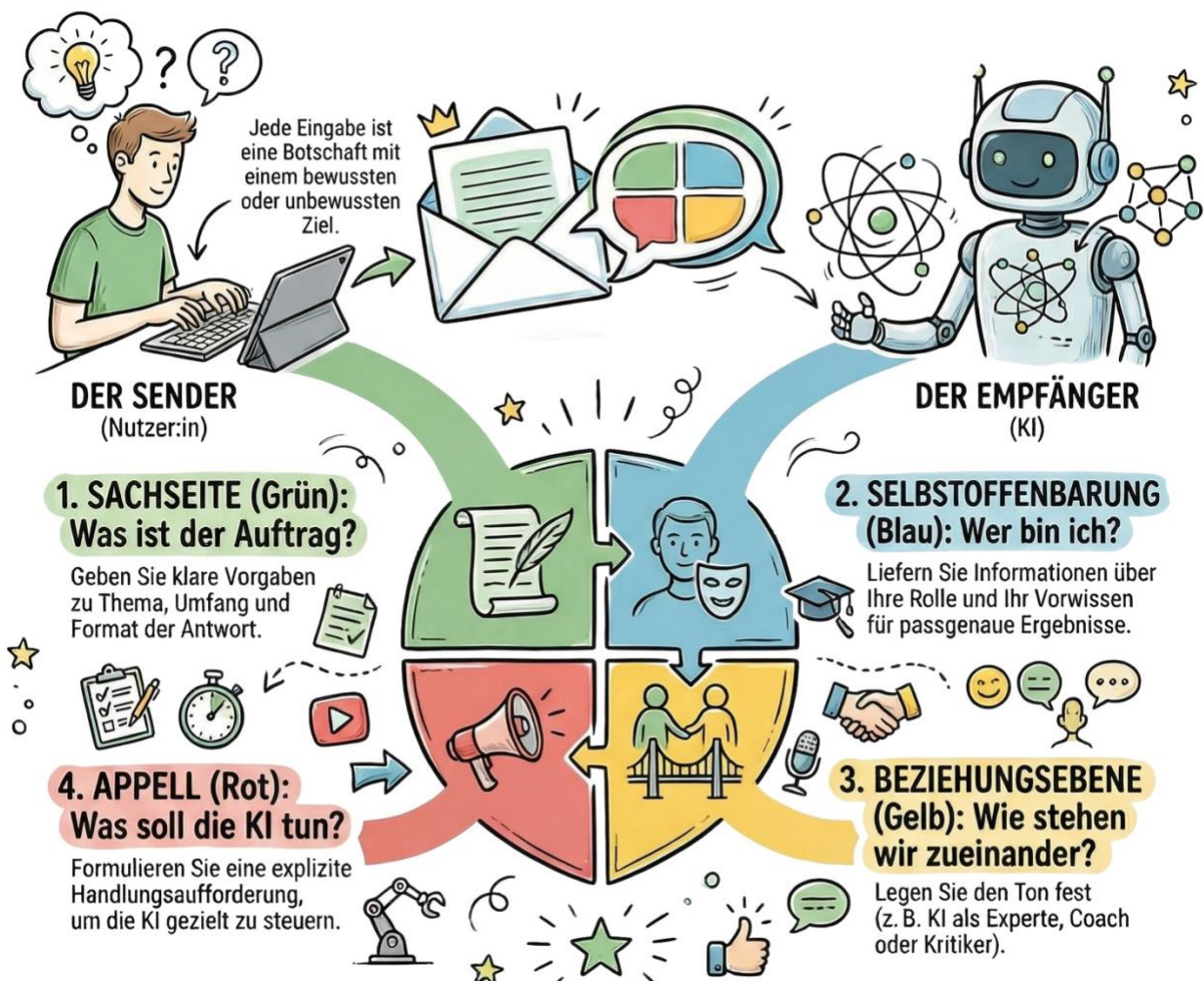
⁴ Shah, Kanya (2025): *The Importance of System Prompts in Shaping AI Agent Responses*. Verfügbar unter: <https://www.getmaxim.ai/articles/the-importance-of-system-prompts-in-shaping-ai-agent-responses/> [Zugriff: 1.12.2025].

es mir, als wäre ich ein Laie“ oder „Sei kritisch und hinterfrage meine Argumente“ setzen den Rahmen für die Interaktion. Sie signalisieren der KI, ob eine unterstützende, herausfordernde oder neutrale Haltung erwünscht ist. Klare Beziehungshinweise vermeiden Missverständnisse und sorgen dafür, dass die KI-Antworten im gewünschten Stil und mit der passenden Distanz oder Nähe formuliert werden. Diese Ebene ist besonders wichtig, um eine vertrauensvolle und produktive Interaktion zu gestalten.

Jeder Prompt enthält zudem einen **Appell** – eine Aufforderung an die KI, eine bestimmte Handlung auszuführen. Dies kann explizit sein, wie in „Stelle zuerst Rückfragen, bevor du antwortest“, oder implizit, wie in „Erstelle am Ende eine Checkliste“. Explizite Handlungsaufforderungen reduzieren „Dekodierungsfehler“, also Situationen, in denen die KI die Erwartungen der Nutzer:innen nicht trifft. Durch klare Appelle wird die KI gezielt gesteuert, was zu präziseren und nützlicheren Ergebnissen führt. Diese Ebene entspricht der Klärung der Beziehungsebene bei Schulz von Thun, bei der es darum geht, Erwartungen und Ziele transparent zu kommunizieren.

PROMPTEN ALS DIALOGISCHER PROZESS

Die Parallelen zwischen Schulz von Thuns Modell und dem Prompten zeigen, dass gute KI-Kommunikation dieselben Prinzipien erfordert wie erfolgreiche zwischenmenschliche Gespräche: Klarheit, Kontextbewusstsein und die Bereitschaft, die eigene Perspektive und



die des Gegenübers – oder in diesem Fall des KI-Systems – zu reflektieren. Wer diese Ebenen bewusst gestaltet, kann die Interaktion mit KI-Systemen nicht nur effizienter, sondern auch natürlicher und zielgerichteter gestalten.

ITERATIVER DIALOG

Prompten ist kein einmaliger Akt, sondern ein iterativer Prozess, der sich durch kontinuierliches Nachschärfen, Nachfragen und Kontextualisieren auszeichnet. Diese Dynamik ähnelt den metakommunikativen Prozessen, wie sie in Beratungsgesprächen oder Unterrichtssituationen vorkommen: Genauso wie in zwischenmenschlichen Dialogen nicht jede erste Antwort perfekt ist, erfordert auch die Interaktion mit KI oft mehrere Schleifen, um zu optimalen Ergebnissen zu gelangen. Nutzer*innen haben dabei die Möglichkeit, durch gezielte Fragen, Präzisierungen oder Anpassungen des Prompts die Antworten der KI schrittweise zu verbessern. Jede Rückmeldung an die KI – sei es eine Korrektur, eine Ergänzung oder eine neue Perspektive – wirkt wie ein Feedbackloop, der die Qualität der Interaktion steigert.

Diese Vorgehensweise spiegelt direkt die Prinzipien der zirkulären Kommunikation nach Watzlawick wider, bei der jede Äußerung eine Reaktion hervorruft, die wiederum den Rahmen für die nächste Interaktion setzt. Gleichzeitig greift hier auch Schulz von Thuns Konzept der Metakommunikation: Indem Nutzer*innen reflektieren, wie ihre Prompts verstanden wurden, und bewusst steuern, wie sie ihre Anliegen formulieren, schaffen sie Klarheit über Erwartungen, Ziele und mögliche Missverständnisse. Multi-Turn-Prompting – also das mehrmalige, aufeinander aufbauende Prompten – nutzt genau diese Feedback-Schleifen, um die Interaktion mit der KI zu verfeinern. Studien zeigen, dass dieser dialogische Ansatz nicht nur die Präzision der KI-Antworten erhöht, sondern auch die Fähigkeit der Nutzer*innen stärkt, ihre eigenen Anliegen klarer zu artikulieren und die KI als aktiven Gesprächspartner zu begreifen.

PROMPTEN ALS ZUKUNFTSSKILL DER MENSCH-MASCHINE-INTERAKTION

Der Schlüssel zu erfolgreichem Prompten liegt somit nicht nur in der initialen Formulierung, sondern auch in der Bereitschaft, die Interaktion als lernenden und kommunikativen Prozess zu verstehen. Prompten ist keine statische Fähigkeit, bei der jemand von außen sagen kann: „Mit diesem Prompt kommst du zu diesem Ergebnis.“ Vielmehr handelt es sich um eine dynamische Kompetenz der Mensch-Maschine-Interaktion, die starke Parallelen zu kommunikationswissenschaftlichen Prinzipien aufweist. Ein offener und reflexiver Umgang mit den Antworten der KI führt zu tiefergehenden, zielgerichteteren und letztlich nützlicheren Ergebnissen. Dieser Ansatz unterstreicht, dass gutes Prompten nicht nur technisches Know-how erfordert, sondern vor allem kommunikative Kompetenz: die Fähigkeit, Feedback zu geben, Perspektiven zu wechseln und gemeinsam mit der KI Lösungen zu erarbeiten. Die Ähnlichkeiten zu kommunikationswissenschaftlichen Modellen – wie den Axiomen von Watzlawick oder dem Vier-Seiten-Modell von Schulz von Thun – zeigen, dass Prompten als angewandte Gesprächsführung verstanden werden muss. Es wird daher eine zentrale Aufgabe im Bildungsbereich sein, diese Parallelen zu erkennen und in didaktische Konzepte zu integrieren. Nur so kann die Fähigkeit, effektiv mit KI-Systemen zu kommunizieren, als Zukunftsskill vermittelt und gefördert werden.

KI-SPRACHMODELLE ANPASSEN: WIE WIR ZIELE PRÄZISER ERREICHEN UND BESSERE TEXTE SCHREIBEN

von Tobias Albers-Heinemann

Wie bereits deutlich wurde, spielt das Prompting – also die Fähigkeit, die richtigen Eingaben zu formulieren – eine wichtige Rolle, um von einer KI die gewünschten Ergebnisse zu erhalten. Dabei geht es nicht nur um die reine Formulierung der Frage oder Aufgabe, sondern vor allem um den Kontext, den wir dem Modell mitgeben. Je detaillierter wir die Rolle des Modells, die Zielgruppe, das Sprachniveau und den gesamten Anwendungskontext beschreiben, desto präziser und passgenauer werden die Ergebnisse.

Doch die Möglichkeiten, den Kontext zu individualisieren und zu professionalisieren, gehen weit über das reine Prompting hinaus. Viele KI-Dienste – wie auch Mistral oder OpenAI – bieten die Funktion der individuellen Anpassung von Sprachmodellen. Der entscheidende Unterschied liegt im Grundkonzept der Datenbasis: Während die Standardversionen dieser Sprachmodelle mit riesigen und intransparenten Datensätzen aus dem gesamten Internet trainiert sind, bieten dieselben Anbieter mit ihren anpassbaren Optionen die Möglichkeit, eine eigene, kuratierte Wissensbasis aufzubauen. Nutzer*innen können dabei Dokumente hochladen, spezifische Rollen definieren und klare Anweisungen hinterlegen – ohne auf externe Anbieter oder zusätzliche Tools angewiesen zu sein. OpenAI nennt diese Funktion **CustomGPT**, bei Mistral sind es **Agenten**, die mit **Bibliotheken** arbeiten.

Auf diese Weise entsteht ein System, das exakt auf die eigenen Inhalte, Qualitätsstandards und Anwendungsfälle zugeschnitten ist – sei es für Förderrichtlinien, theologische Texte, pädagogische Leitfäden oder interne Richtlinien. Die KI greift dann nicht mehr auf das unüberschaubare Internet zurück, sondern auf die genau definierten Quellen, die für den jeweiligen Einsatzzweck relevant sind.

BEISPIEL: SACHBERICHTERSTELLUNG IN EINEM FÖRDERFÄHIGEN PROJEKT

Fragt man ein Standard-Sprachmodell, wie ein Sachbericht gegliedert sein muss, durchsucht es seine allgemeine Datenbasis nach allen verfügbaren Sachberichten und schlägt auf Basis statistischer Wahrscheinlichkeiten ein Ergebnis vor. Dieses sieht zwar gut aus und klingt plausibel, entspricht aber sehr wahrscheinlich nicht den spezifischen Kriterien der jeweiligen Förderlinie oder des Projekts.

Ein angepasstes Sprachmodell hingegen kann mit den konkreten Anforderungen der Förderlinie trainiert werden – etwa durch das Hochladen der Förderrichtlinien, früherer Berichte oder interner Vorlagen. Die KI erstellt dann nicht nur eine Gliederung, die den formalen Vorgaben entspricht, sondern kann auch inhaltliche Schwerpunkte setzen, die für das Projekt relevant sind, und sicherstellen, dass alle notwendigen Informationen enthalten sind. So wird aus einer allgemeinen Vorlage ein maßgeschneiderter, förderkonformer Sachbericht, der direkt weiterverwendet werden kann.

DER TON MACHT DIE MUSIK: SYSTEM-PROMPTS UND DATENBANKEN ALS GRUNDLAGE

Bevor ein solch angepasstes Sprachmodell jedoch arbeiten kann, muss es eingerichtet und konfiguriert werden. Der erste Schritt ist dabei die Erstellung eines **Systemprompts** – quasi die Arbeitsanweisung für die KI. Dieser Prompt definiert, welche Rolle die KI einnimmt, welche Aufgaben sie erfüllt und in welchem Rahmen sie agiert. Parallel wird eine Wissensdatenbank mit relevanten Dokumenten aufgebaut, auf die das Modell zugreifen kann. Die Erstellung eines Systemprompts klingt jetzt erst einmal herausfordernd, aber: Wer unsicher ist, kann einfach mit einer generativen KI in den Dialog treten und um Unterstützung bitten.

Ein Beispiel für eine solche Anfrage: *„Ich möchte einen Systemprompt für einen Agenten schreiben, der die Rolle eines Förderberaters für soziale Projekte übernimmt. Der Agent soll Sachberichte nach den Vorgaben einer bestimmten Förderlinie erstellen. Hilf mir dabei, den Prompt zu formulieren.“*

Die KI unterstützt dann bei der Formulierung und gibt Beispiele, wie der Prompt aussehen könnte. So wird der Einstieg in die Anpassung von Sprachmodellen auch ohne technische Vorkenntnisse möglich. Allerdings ist auch zu beachten, dass neben dem Inhalt des Systemprompts auch das Format eine entscheidende Rolle für die Qualität der Ergebnisse spielt.

BEISPIELE FÜR VERSCHIEDENE FORMATE (EXEMPLARISCHE AUSWAHL):

Die folgenden Beispiele zeigen, wie unterschiedlich Systemprompts aussehen können – von einfachen Textformaten bis hin zu strukturierten Datenformaten - nur, dass Sie es einmal gesehen haben. Welches Format sich letztendlich eignet, hängt vom jeweiligen Anwendungsfall und der gewünschten Präzision ab.

Fließtext (einfache Textform):

- 1 Du bist ein Förderberater für soziale Projekte. Deine Aufgabe ist es, ein
- 2 Der Bericht soll folgende Abschnitte enthalten:
- 3 1. Projektbeschreibung mit Zielsetzung, Zielgruppe und Maßnahmen
- 4 2. Finanzielle Abrechnung mit Kostenaufstellung und Nachweisen
- 5 Der Stil soll präzise, sachlich und förderkonform sein.

Markup / Markdown

```
1 ## Rolle
2 Förderberater für soziale Projekte
3
4 ## Aufgabe
5 Erstelle einen Sachbericht nach den Vorgaben der Förderlinie XY.
6
7 ## Struktur
8 - Projektbeschreibung
9   - Zielsetzung
10  - Zielgruppe
11  - Maßnahmen
12 - Finanzielle Abrechnung
13   - Kostenaufstellung
14   - Nachweise
15
16 ## Stil
17 präzise, sachlich, förderkonform
```

XML

```
1 <SystemPrompt>
2   <Rolle>Förderberater für soziale Projekte</Rolle>
3   <Aufgabe>Erstelle einen Sachbericht nach den Vorgaben der Förderlinie X
4   <Struktur>
5     <Abschnitt name="Projektbeschreibung">
6       <Inhalt>Zielsetzung</Inhalt>
7       <Inhalt>Zielgruppe</Inhalt>
8       <Inhalt>Maßnahmen</Inhalt>
9     </Abschnitt>
10    <Abschnitt name="Finanzielle Abrechnung">
11      <Inhalt>Kostenaufstellung</Inhalt>
12      <Inhalt>Nachweise</Inhalt>
13    </Abschnitt>
14  </Struktur>
15  <Stil>
16    <Eigenschaft>präzise</Eigenschaft>
17    <Eigenschaft>sachlich</Eigenschaft>
18    <Eigenschaft>förderkonform</Eigenschaft>
19  </Stil>
20 </SystemPrompt>
```

JSON (JavaScript Object Notation)

```
1 {  
2   "Rolle": "Förderberater für soziale Projekte",  
3   "Aufgabe": "Erstelle einen Sachbericht nach den Vorgaben der Förderlini  
4   "Struktur": [  
5     {  
6       "Abschnitt": "Projektbeschreibung",  
7       "Inhalt": ["Zielsetzung", "Zielgruppe", "Maßnahmen"]  
8     },  
9     {  
10      "Abschnitt": "Finanzielle Abrechnung",  
11      "Inhalt": ["Kostenaufstellung", "Nachweise"]  
12    },  
13    "Stil": ["präzise", "sachlich", "förderkonform"]  
14  }
```

VERGLEICH DER FORMATE UND IHRE AUSWIRKUNGEN AUF DIE KI

Auf den ersten Blick mögen diese Formate kompliziert wirken – besonders XML oder JSON. Doch keine Sorge: Mit der Unterstützung einer KI lassen sich solche System-Prompts ganz einfach erstellen oder umwandeln. Ein kurzer Befehl wie "Wandle diesen Fließtext in ein JSON-Format um" genügt, und die KI übernimmt die Arbeit. Die Wahl des Formats hat jedoch deutliche Auswirkungen auf das Verhalten des Sprachmodells:

- **Fließtext** ist zwar einfach zu schreiben, aber die KI muss die Struktur selbst interpretieren. Das kann zu Unschärfen führen, besonders bei komplexen Anforderungen. Fließtext eignet sich gut für einfache, klare Anweisungen, bei denen keine detaillierte Struktur nötig ist.
- **Markup / Markdown** macht Hierarchien und Schwerpunkte sichtbar – etwa durch Überschriften, Listen oder Hervorhebungen. Die KI erkennt so besser, welche Teile besonders wichtig sind. Markup ist ideal, wenn Sie lesbare Struktur mit visueller Klarheit verbinden möchten.
- **XML** bietet eine klare, hierarchische Struktur und ist besonders in technischen Umgebungen verbreitet. Es eignet sich gut, wenn Sie mit anderen Systemen oder Datenbanken arbeiten, die XML verarbeiten.
- **JSON** ist das präziseste Format für maschinelle Verarbeitung. Es gibt eine klare, maschinell auswertbare Struktur vor und eignet sich ideal für komplexe Anforderungen oder wenn die Ergebnisse weiterverarbeitet werden sollen. Mit JSON lassen sich logische Abfolgen, Datenfelder und Bedingungen besonders gut abbilden – das führt zu deutlich präziseren und zuverlässigeren Ergebnissen.

Einfach ausprobieren! Die beste Methode, um das passende Format zu finden, ist, selbst zu experimentieren. Probieren Sie aus, wie sich Ihre System-Prompts in den verschiedenen Formaten verfassen lassen – und scheuen Sie sich nicht, die KI um Hilfe zu bitten. Sie werden überrascht sein, wie stark sich das Verhalten des jeweiligen Modells ändert, wenn Sie nur das Format anpassen.

KI, VORURTEILE UND VERANTWORTUNG – WIE WIR KI DISKRIMINIERUNGSSSENSIBEL NUTZEN KÖNNEN

von Lukas Spahlinger, Mike Breitbart und Tobias Albers-Heinemann

Bilder in Lehrmaterial, Fotodatenbanken, Stellenanzeigen oder Nachrichtenbeiträgen prägen unsere Vorstellung davon, was als „normal“ oder als „ungewöhnlich“ gilt. Solche Muster wirken – sie lenken Aufmerksamkeit, prägen Selbstbilder und Entscheidungen. Sie sind nicht neu. Neu ist, wie stark digitale Systeme sie vervielfältigen können. Die einfache Bedienung und schnellen Ergebnisse von Chatbots oder Bildgeneratoren werden im Alltag oft als praktische Unterstützung oder clevere Hilfe wahrgenommen. Dabei wird Künstliche Intelligenz schnell für neutral gehalten – dabei ist genau das ein Trugschluss.

KI-Systeme sind nicht neutral – sie spiegeln die Wertesysteme, Vorannahmen und Vorurteile ihrer Entwickler*innen und der zugrunde liegenden Trainingsdaten wider. Sie übernehmen stereotype Muster, die aus gesellschaftlicher Realität stammen, und verfestigen diese – oft unbemerkt, aber mit spürbaren Konsequenzen für die Lebenswirklichkeit vieler Menschen.

Was aber bedeutet das konkret? Wie genau gelangen diese Verzerrungen in KI-Systeme? Welche Verantwortung tragen Entwickler*innen – und auch Nutzer*innen? Und wie können wir KI so einsetzen, dass sie gesellschaftliche Vielfalt abbildet statt sie zu verzerren? Dieser Text geht diesen Fragen nach und zeigt Strategien für eine diskriminierungssensible KI-Nutzung auf.

WAS IST EIN BIAS?

Bias ist der englische Begriff für Verzerrung, Vorurteil oder Voreingenommenheit. Im Zusammenhang mit statistischen Modellen wird der Begriff verwendet, um deutlich zu machen, dass die Datengrundlage nicht den realen Zusammenhängen entspricht. Dadurch werden Zusammenhänge, Aussagen oder Darstellungen verzerrt wiedergegeben. Dies ist kein seltenes Phänomen. Daten in ihrer vielfältigen Form (als Texte, Bilder oder Umfragedaten) wurden von Menschen produziert. Sie schreiben Texte, wählen Motive für Bilder aus oder erstellen Fragekataloge für Umfragen. Dabei sind Menschen nicht neutral. Das hat mit der selektiven Wahrnehmung von uns Menschen zu tun. Wir lernen aufgrund unserer Beobachtungen und unserer Erfahrungen. Diese werden sortiert und kategorisiert, um die Vielfalt an Informationen verarbeiten zu können – das kann hilfreich sein, aber auch Vorurteile erzeugen. Wenn Menschen Texte schreiben oder auf andere Art und Weise die Realität dokumentieren, entstehen Inhalte mit unterschiedlichen Sichtweisen und Perspektiven, aber auch mit Vorurteilen und Stereotypen.

WIE BIAS IN KÜNSTLICHER INTELLIGENZ ENTSTEHT

KI-Modelle basieren auf riesigen Datenmengen. Die Daten bilden die Welt, die wir sehen und erleben können, in sehr großen Teilen ab, einschließlich der Verzerrungen, Vorurteile, Vorannahmen oder Ungerechtigkeiten. Die Daten sind nicht neutral, genauso wenig die KI-Modelle, die auf Basis der Daten trainiert werden. Das heißt: Wenn KI-Systeme aus solchen Daten lernen, übernehmen sie nicht nur Fakten, sondern auch soziale Verzerrungen. Stereotype Muster entstehen nicht erst im System, sondern sind in Form von unausgewogenen Darstellungen, diskriminierenden Sprachmustern oder einseitigen

Kategorisierungen direkt in die Trainingsdaten eingeschrieben – oft unbemerkt, aber mit weitreichender Wirkung. Generative Künstliche Intelligenz basiert auf Wahrscheinlichkeiten und Kontextualisierung. Sie gibt das wieder, was auf Basis von statistischen Häufigkeiten und Zusammenhängen und den durch die Eingabe gegebenen Kontext das wahrscheinlichste Ergebnis ist. Man kann also sagen: Maschinen lernen aus Texten, Bildern oder anderen Datenquellen, in denen beispielsweise Führungskräfte fast ausschließlich männlich und weiß dargestellt werden. Sie übernehmen Begriffe und Formulierungen, die bestimmte Gruppen benachteiligen oder unsichtbar machen. Bekannte Beispiele verdeutlichen die Auswirkungen solcher Verzerrungen in der Praxis:

- Amazons Recruiting-KI bewertete Bewerbungen von Frauen systematisch schlechter, weil das Trainingsmaterial fast ausschließlich männlich geprägte Karriereverläufe enthielt. Ganz unabhängig davon, dass Frauen gleiche Qualifikationen und Fähigkeiten mitbrachten wie ihre männlichen Mitbewerber.
- Gesichtserkennungssysteme zeigten bei dunkelhäutigen Personen deutlich höhere Fehlerraten als bei weißen Personen – mit teils gravierenden Folgen im Sicherheitsbereich. Denn die Systeme wurden hauptsächlich mit Bildern trainiert, die weiße Personen zeigten.
- Die Apple Card gewährte Frauen geringere Kreditlimits als Männern, obwohl sie über vergleichbare finanzielle Voraussetzungen verfügten.

Diese Beispiele zeigen: KI erkennt zwar Muster, versteht aber den gesellschaftlichen Kontext nicht. Die Systeme analysieren Korrelationen, ohne die Bedeutungen oder Diskriminierungsrisiken zu erfassen. Hinzu kommt, dass Systemprompts – das sind systeminterne Anweisungen und Voreinstellungen, die festlegen, wie ein KI-Modell auf bestimmte Anfragen reagieren soll – erheblichen Einfluss auf die Ergebnisse haben. Sie bestimmen unter anderem, welche Themen priorisiert oder welche Formulierungen vermieden werden. Diese Prompts sind meist unsichtbar für die Nutzer*innen, können aber entscheidend dafür sein, ob eine Antwort stereotypen Mustern folgt oder Vielfalt abbildet. Auch die Zusammensetzung der Entwicklerteams spielt eine zentrale Rolle: Sind diese homogen – etwa überwiegend männlich, weiß und akademisch geprägt – fehlen diverse Perspektiven. Das Risiko, Vorurteile zu reproduzieren, steigt dadurch erheblich.

WIE NUTZER*INNEN DEN BIAS SELBST VERSTÄRKEN

Es gibt sehr viele unterschiedliche Formen dieser Bias. Ein zentrales Phänomen ist der sogenannte Confirmation Bias: die Tendenz, Informationen so auszuwählen und zu interpretieren, dass sie bestehende Überzeugungen und Wertemuster bestätigen. Wer in seinem Kopf bestimmte Bilder abgespeichert hat – etwa vom „typischen Chef“ als weißem Mann im Anzug – wird genau solche Darstellungen auch unbewusst als normal empfinden. KI-Modelle wie ChatGPT lernen in den Trainingsphasen aus den Datensätzen. Sobald wir die abgeschlossenen Large Language Models (LLMs) verwenden, passen sie sich daran an, was wir ihnen zeigen – und wie wir mit ihnen umgehen. Wer immer wieder ähnliche Fragen stellt, bekommt ähnliche Antworten. Die Gefahr: Die KI bestätigt uns in unseren Vorstellungen. Sie reflektiert nicht nur unsere Erwartungen, sondern bestärkt sie – wie ein Spiegel, der nicht nur zeigt, was ist, sondern das Bild zurückwirft. Wer Vielfalt sehen will, muss sie bewusst einfordern. Hier helfen sogenannte sensible Prompts. Anstatt von einer

„normalen Familie“ zu sprechen, lässt sich gezielt formulieren: „Erstelle ein Bild einer Familie mit verschiedenen Konstellationen, Hautfarben, Altersstufen und Geschlechtern.“

Auch Reflexionsfragen an die KI können helfen, sichtbar zu machen: „Welche Perspektiven fehlen in deiner Darstellung?“ oder „Welche möglichen Probleme siehst du selbst?“ – Solche Nachfragen öffnen den Raum für mehr Vielfalt und zeigen, dass es auch im Umgang mit KI auf unsere Haltung ankommt.

STEREOTYPE REPRODUKTION IM ALLTAG

Dass diese Verzerrungen in den KI-Daten keine Theorie sind, zeigen zahlreiche Untersuchungen. Eine Analyse der UNESCO aus dem Jahr 2024 stellte fest, dass große Sprachmodelle wie ChatGPT und andere systematisch geschlechtsbezogene, rassistische und homophobe Stereotype reproduzieren.

KI-generierte Inhalte können dadurch Diskriminierung verstärken:

- Frauen werden als weniger kompetent oder vor allem in häuslichen Rollen dargestellt.
- BIPOC-Personen (engl. Begriff für Black, Indigenous and People of Color) werden entweder ignoriert oder stereotyp dargestellt.
- Menschen mit abweichender sexueller Orientierung werden unsichtbar gemacht oder diskriminierend beschrieben.

Diese Beispiele zeigen, dass KI die Vielfalt der Gesellschaft nicht von selbst abbildet. Hier braucht es bewusste Steuerung und ein Bewusstsein für die stereotypen Muster von KI-Systemen – sowohl auf technischer als auch auf gesellschaftlicher Ebene.

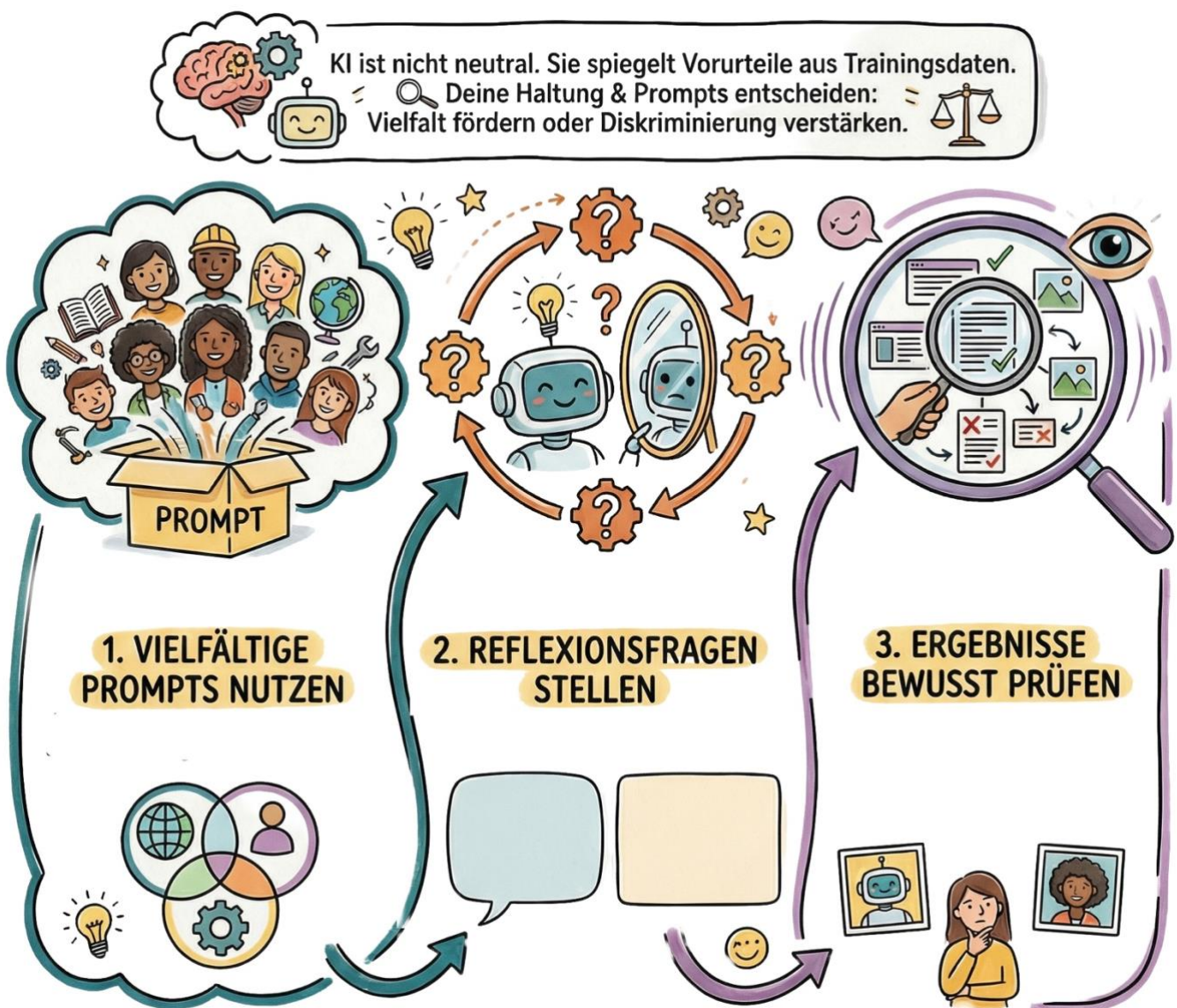
VERANTWORTUNG IN ENTWICKLUNG UND ANWENDUNG

Verzerrungen entstehen nicht nur durch Daten, sondern auch durch Designentscheidungen. So können bestimmte Optimierungsziele – etwa maximale Reichweite oder Nutzerbindung – dazu führen, dass polarisierende Inhalte bevorzugt werden. Ein weiterer Einflussfaktor sind die Systemprompts – Anweisungen, die das Verhalten der KI steuern. Hier entscheidet sich, welche Informationen die KI berücksichtigt und welche ausgeblendet werden. Fehlende Transparenz und verdeckte Verzerrungen in der Entwicklung verstärken diese Effekte.

Verantwortung bedeutet also nicht nur, gute Daten zu nutzen, sondern auch, Vielfalt im Entwicklerteam zu fördern und die Systeme kontinuierlich zu überprüfen. Denn nur ein diverses Team bringt auch eine Vielfalt an Erfahrungen, Blickwinkeln und Lebensrealitäten mit. Wo unterschiedliche Perspektiven zusammentreffen, entstehen andere Fragen, sensiblere Entscheidungsprozesse – und eine kritischere Auseinandersetzung mit potenziellen Verzerrungen. Ein diverses Team wirkt wie ein Schutzfilter gegen blinde Flecken: Es ist wahrscheinlicher, dass diskriminierende Strukturen auffallen, wenn Menschen mit unterschiedlichen Hintergründen gemeinsam an Lösungen arbeiten. Wer KI für alle bauen will, muss auch alle im Blick haben – und das beginnt bei den Menschen, die die Systeme entwickeln.

STRATEGIEN FÜR EINE DISKRIMINIERUNGSSENSIBLE KI-NUTZUNG

Der Umgang mit Künstlicher Intelligenz ist kein rein technischer Vorgang – er ist eine Frage der Haltung. Wer glaubt, KI sei automatisch fair, unterschätzt, wie stark unsere eigenen Vorannahmen in die Ergebnisse einfließen. Deshalb beginnt diskriminierungssensible KI-Nutzung bei uns selbst – bei dem Bewusstsein dafür, dass wir mit jeder Anfrage, jedem Prompt und jeder Reaktion mitgestalten, was die KI für normal, relevant oder sichtbar hält. Wer Vielfalt sehen will, muss sie bewusst einfordern. Denn ohne klare Impulse übernimmt die KI nur, was in den Daten am lautesten vertreten ist – und das sind oft stereotype Vorstellungen.



Wer dagegen aktiv andere Perspektiven einbringt, kann dazu beitragen, dass KI nicht nur Bestehendes reproduziert, sondern neue Sichtweisen ermöglicht. Das gelingt durch:

- **Vielfältige Prompts:** gezielte Aufforderungen an die KI, verschiedene Perspektiven einzubeziehen.
- **Reflexionsfragen:** die KI um Einschätzungen zu möglichen Verzerrungen bitten.
- **Bewusste Nutzung:** sich nicht auf die erste Antwort verlassen, sondern gezielt nach Alternativen suchen. KI ist kein Ersatz für kritisches Denken. Sie kann ein Werkzeug sein, um neue Perspektiven zu entdecken – wenn wir sie dazu auffordern.

Künstliche Intelligenz ist weder objektiv noch neutral. Sie spiegelt die Werte und Vorurteile unserer Gesellschaft. Wenn wir lernen, verantwortungsvoll mit KI umzugehen, können wir Vielfalt sichtbar machen und Diskriminierung abbauen. Dafür ist eine Sensibilität für Ungleichheiten, Vorurteilsstrukturen und Diskriminierung notwendig, um diese in den Ergebnissen wie Texten oder Bildern zu erkennen.

Dazu braucht es technische Verantwortung – etwa durch vielfältige Entwicklerteams, die ihre unterschiedlichen Erfahrungen und Perspektiven in die Systemgestaltung einbringen und auf gute, ausgewogene Daten achten. Ebenso entscheidend ist die gesellschaftliche Verantwortung: Nutzer*innen müssen lernen, KI bewusst, kritisch und reflektiert einzusetzen. Schließlich braucht es politische Verantwortung – also klare gesetzliche Rahmenbedingungen und verbindliche Standards, die faire und inklusive Systeme fördern.

Der Weg zu einer gerechteren KI beginnt bei uns selbst – mit jedem Prompt, jeder Frage und jeder Entscheidung, Vielfalt einzufordern und zu fördern.

KÜNSTLICHE INTELLIGENZ UND DER DATENSCHUTZ

von Tobias Albers-Heinemann

Das Thema des Datenschutzes stellt eine der größten Herausforderungen der digitalen Transformation dar – insbesondere im Zusammenhang mit dem Einsatz von Künstlicher Intelligenz. Denn während KI-Systeme große Datenmengen benötigen, um effektiv zu funktionieren, ist der Schutz personenbezogener Informationen ein zentrales Anliegen demokratischer Gesellschaften. Wie kann der Einsatz von KI also gelingen, ohne gegen Datenschutzgrundsätze zu verstoßen? Und wie lassen sich innovative Entwicklungen mit den Rechten und Bedürfnissen der Nutzer*innen in Einklang bringen? Es stellt sich zudem die grundsätzliche Frage: Worum geht es beim Datenschutz eigentlich genau – und ist es tatsächlich möglich, KI-Tools so zu gestalten und einzusetzen, dass sie sowohl nützlich als auch datenschutzkonform sind?

ZWISCHEN TECHNISCHER INNOVATION UND SCHUTZBEDÜRFNIS

Im Zentrum des Spannungsfeldes zwischen technischer Innovation und dem Schutzbedürfnis der Nutzer*innen stehen zwei zentrale Interessen: Auf der einen Seite verfolgen Anbieter von KI-Diensten das Ziel, möglichst viele Daten zu erfassen, zu analysieren und weiterzuverarbeiten – sei es zur Optimierung bestehender Systeme, zur Entwicklung neuer Produkte oder zur wirtschaftlichen Verwertung, etwa durch Weitergabe an Dritte. Ohne diese umfassende Datenverarbeitung wäre ein Großteil der rasanten technologischen Fortschritte in den Bereichen maschinelles Lernen und Sprachverarbeitung kaum denkbar.

Auf der anderen Seite steht das berechtigte Bedürfnis von Nutzer*innen nach Schutz ihrer Privatsphäre, nach Transparenz und Kontrolle über die eigenen Daten. Besonders deutlich wird diese Diskrepanz, wenn vermeintlich kostenlose Apps und Dienste im Hintergrund personenbezogene Informationen sammeln – häufig ohne explizite Zustimmung oder verständliche Hinweise. Das erzeugt nicht nur ein Gefühl der Ohnmacht, sondern gefährdet auch das Vertrauen in digitale Technologien.

Gerade in sensiblen Bereichen wie Bildung, Gesundheit oder sozialer Arbeit ist daher besondere Sorgfalt geboten. Denn Datenschutz schützt nicht bloß Daten, sondern vor allem die Menschen hinter den Daten – ihre Würde, ihre Rechte, ihre Freiheiten. Ohne ausreichende Kontrollmechanismen kann es zu gravierenden Folgen kommen: Diskriminierung durch algorithmische Verzerrungen, politische Einflussnahme oder die Einschränkung persönlicher Entfaltungsmöglichkeiten. Ein aktuelles Beispiel unterstreicht diese Gefahr: Im Jahr 2025 wurde mehreren europäischen Bürger*innen die Einreise in die USA verweigert, weil sie sich in sozialen Medien kritisch gegenüber dem amerikanischen Präsidenten geäußert hatten. Solche Vorfälle zeigen eindrucksvoll, wie zentral der Schutz persönlicher Daten für demokratische Teilhabe und individuelle Freiheit ist.

WAS ZÄHLT ZU PERSONENBEZOGENEN DATEN?

Neben offensichtlichen Angaben wie Namen oder Adressen gelten auch IP-Adressen, Standortdaten, Chatverläufe oder Stimmprofile als personenbezogen – insbesondere, wenn sie durch KI-gestützte Analyse miteinander verknüpft werden können. Auch scheinbar anonyme Daten können bei langfristiger Nutzung Rückschlüsse auf Einzelpersonen

zulassen. Datenschutzrechtlich relevant ist jede Verarbeitung personenbezogener Daten – unabhängig davon, ob sie gespeichert, analysiert oder nur kurzfristig verarbeitet werden. Dabei ist stets eine sorgfältige Abwägung erforderlich: Eine sogenannte Risikofolgenabschätzung hilft dabei, potenzielle Gefährdungen für die betroffenen Personen einzuschätzen und geeignete Schutzmaßnahmen zu treffen. So ist beispielsweise die Weitergabe einer IP-Adresse im regulären Arbeitsalltag oft unproblematisch – für politisch verfolgte Personen, etwa in totalitären Staaten oder unter repressiven Regimen, kann sie hingegen existenzbedrohende Konsequenzen haben. Datenschutz muss daher immer im Kontext möglicher Risiken gedacht werden.

DATENSCHUTZRECHTLICHE GRUNDLAGEN

Die Datenschutz-Grundverordnung (DSGVO) sowie ergänzende Datenschutzgesetze auf Landesebene und spezifische kirchliche Datenschutzgesetze – etwa das DSG-EKD – bilden den rechtlichen Rahmen. Die DSGVO richtet sich dabei nicht an einzelne Privatpersonen, sondern verpflichtet öffentliche Stellen, Unternehmen und sonstige Einrichtungen, die personenbezogene Daten verarbeiten, zur Einhaltung datenschutzrechtlicher Grundsätze. Diese parallelen Regelwerke führen jedoch in der Praxis häufig zu Unsicherheit und Verwirrung: Je nach Trägerschaft einer Einrichtung (staatlich, kommunal oder kirchlich) und je nach Bundesland können unterschiedliche Datenschutzvorgaben gelten. Hinzu kommt, dass die zuständigen Datenschutzbeauftragten diese Regelungen teils unterschiedlich auslegen. So kann es etwa passieren, dass ein bestimmtes KI-Tool in einem Bundesland oder bei einem bestimmten Träger eingesetzt werden darf, während es in einer anderen Einrichtung untersagt ist. In dieser komplexen Gemengelage fühlen sich viele Einrichtungen alleingelassen oder durch pauschale Verbote gehemmt. Dabei bietet die Rechtslage durchaus Spielräume – vorausgesetzt, es gibt ausreichende Kenntnisse, klare Strukturen und verantwortungsbewusste Entscheidungen.

DATENSCHUTZ IST MÖGLICH – MIT KLAREN PRINZIPIEN

Für jede Verarbeitung personenbezogener Daten muss eine gültige Rechtsgrundlage vorliegen. Diese kann sich unter anderem auf die Einwilligung der betroffenen Person (Art. 6 Abs. 1 lit. a DSGVO), die Erfüllung eines Vertrags (lit. b), rechtliche Verpflichtungen (lit. c), lebenswichtige Interessen (lit. d), die Wahrnehmung einer Aufgabe im öffentlichen Interesse (lit. e) oder berechtigte Interessen des Verantwortlichen (lit. f) stützen.

Umgekehrt gilt: Wenn keine personenbezogenen Daten verarbeitet werden, besteht in der Regel auch keine datenschutzrechtliche Relevanz. Dennoch ist eine sorgfältige Prüfung stets notwendig, da scheinbar neutrale Informationen durch KI-gestützte Analyse personenbezogen werden können. Ein Beispiel: Die Berufsbezeichnung „Bildungsreferent“ ist zunächst recht allgemein und anonym. Wird sie jedoch ergänzt zu „Referent für digitale Bildung der EKHN“, kann diese Information sehr schnell eine konkrete Person identifizierbar machen. Solche Kontexte müssen stets mitgedacht werden, um den Schutz personenbezogener Daten zu gewährleisten.

Datenschutzkonformer KI-Einsatz ist nicht nur möglich, sondern auch notwendig. Entscheidend sind transparente Prozesse, fundierte Entscheidungen und eine kritische Reflexion über Zweck, Umfang und Notwendigkeit der Datenverarbeitung. Bei der Nutzung

eines KI-gestützten Textgenerators sollte daher vorab geprüft werden, ob personenbezogene Daten eingegeben werden müssen – etwa biografische Angaben von Teilnehmenden in einem Bildungsprojekt. Wird dies vermieden, entfällt oftmals die datenschutzrechtliche Relevanz.

Auch beim Hochladen von Dokumenten in cloudbasierte KI-Tools ist besondere Vorsicht geboten – etwa bei Anträgen, in denen Namen, Adressen oder Diagnosen enthalten sind. Hier sollten entweder anonymisierte Fassungen erstellt oder lokale Tools verwendet werden, bei denen keine Übertragung an Dritte erfolgt.

STRATEGIEN FÜR DATENSPARSAME KI-NUTZUNG

Im praktischen Alltag gibt es eine Reihe von Strategien, mit denen der Einsatz von KI datenschutzfreundlich gestaltet werden kann. Sie helfen dabei, Risiken zu minimieren und gleichzeitig handlungsfähig zu bleiben:

- Anonymisierung und Pseudonymisierung: z. B. durch allgemeine Nutzerkennungen oder abstrahierte Inhalte
- Privacy by Design: Tools wählen, bei denen die Grundeinstellungen auf Datenschutz ausgerichtet sind wie zum Beispiel duck.ai
- Lokal statt Cloud: lokale Tools wie LM Studio bevorzugen
- Transparenz schaffen: Datenverarbeitungen intern dokumentieren
- Richtlinien entwickeln: z. B. Checklisten oder Teamvereinbarungen für den KI-Einsatz
- Geeignete Anbieter wählen: europäische oder DSGVO-konforme Alternativen nutzen wie LeChat von Mistral
- Risiken regelmäßig neu bewerten und Prozesse bei Bedarf anpassen

DATENSCHUTZ BRAUCHT KOMPETENZ

Datenschutz ist weit mehr als eine technische oder juristische Herausforderung – er ist ein zentrales Element digitaler Bildungskompetenz. Pädagogische Fachkräfte, Entscheidungsträger*innen und Bildungseinrichtungen stehen in der Verantwortung, digitale Technologien nicht nur funktional, sondern auch rechtlich und ethisch fundiert einzusetzen. Dazu gehört die Fähigkeit, Risiken realistisch einzuschätzen sowie die rechtlichen Rahmenbedingungen – wie DSGVO, landes- oder kirchenspezifische Datenschutzgesetze – in die tägliche Praxis zu integrieren.

Gerade in der Bildungsarbeit, die auf Vertrauen, Teilhabe und Schutz sensibler Informationen angewiesen ist, darf Datenschutz nicht als hinderlich oder hemmend verstanden werden. Vielmehr sollte er als gestaltbares Element begriffen werden, das die digitale Souveränität stärkt und Innovation verantwortungsvoll begleitet.

Um das zu erreichen, braucht es zielgerichtete Fortbildungen, verständliche Handreichungen und Räume für kollegialen Austausch. Interdisziplinäre Kooperation – etwa zwischen Pädagogik, IT, Recht und Datenschutzbeauftragten – schafft dabei die notwendige Basis, um konkrete und tragfähige Lösungen zu entwickeln. So entsteht eine Kultur des Vertrauens und der digitalen Souveränität.

KI sicher & datensparsam nutzen

Der Einsatz von Künstlicher Intelligenz wirft wichtige Datenschutzfragen auf. Die folgenden Strategien helfen dabei, Risiken zu minimieren, personenbezogene Daten zu schützen und KI-Tools verantwortungsvoll und im Einklang mit der DSGVO zu nutzen.

1. Anonymisieren & Pseudonymisieren

Persönliche Informationen vor der Eingabe entfernen oder verfremden.

2. "Privacy by Design" nutzen

Tools wählen, deren Grundeinstellungen auf Datenschutz ausgerichtet sind.

3. Lokal statt Cloud

Lokale Tools bevorzugen, damit Daten auf dem eigenen Gerät bleiben.

5. Richtlinien entwickeln

Gemeinsame Checklisten oder Team-Vereinbarungen für den KI-Einsatz festlegen.

4. Transparenz schaffen

Datenverarbeitungen im Team oder in der Organisation dokumentieren.

7. Risiken regelmäßig bewerten

Genutzte Prozesse und Tools immer wieder prüfen und bei Bedarf anpassen.

6. Sichere Anbieter wählen

Europäische oder DSGVO-konforme Alternativen bevorzugen.

KI, URHEBERRECHT UND KENNZEICHNUNGSPFLICHT

von Tobias Albers-Heinemann

Fragen wie „Darf ich das einfach nutzen?“ oder „Muss ich das kennzeichnen?“ begegnen vielen Menschen, die sich mit Künstlicher Intelligenz beschäftigen. Gerade bei der Nutzung von KI-Tools zur Erstellung von Arbeitsmaterialien, Präsentationen oder Texten ist das Urheberrecht ein zentrales Thema. Es wirft nicht nur Fragen zur Nutzung bestehender Inhalte auf, sondern auch zur rechtlichen Bewertung von KI-generierten Produkten. Ebenso unklar ist oft, ob und wie eine Kennzeichnung solcher Inhalte erfolgen muss.

DAS URHEBERRECHT – KURZ ERKLÄRT

Das Urheberrecht regelt, wie Werke genutzt werden dürfen – also wer etwas wie lange, in welcher Form und unter welchen Bedingungen verwenden darf. Dabei schützt es nicht bloße Ideen oder Informationen, sondern die Art und Weise, wie sie konkret gestaltet und umgesetzt wurden. Ein Werk im Sinne des Urheberrechts ist zum Beispiel ein gemaltes Bild, ein geschriebener Text oder ein komponiertes Lied – eben etwas, das erkennbar eine eigene, schöpferische Leistung zeigt. Entscheidend ist, dass es nicht nur „irgendetwas“ ist, sondern eine erkennbare persönliche Note trägt, die über das rein Alltägliche hinausgeht.

Ob ein kreatives Ergebnis wirklich als „Werk“ gilt, hängt maßgeblich davon ab, ob es eine sogenannte persönliche geistige Schöpfung darstellt – so steht es auch in § 2 Absatz 2 des Urheberrechtsgesetzes (UrhG). Damit ist gemeint: Es muss sich um etwas handeln, das individuell gestaltet ist und die Handschrift einer bestimmten Person trägt. Fachlich spricht man dabei von der sogenannten Schöpfungshöhe. Diese zeigt sich zum Beispiel in der besonderen Art, wie ein Thema umgesetzt wird, in einer originellen sprachlichen Gestaltung oder im individuellen Stil. Man kann sich das vorstellen wie bei einem Gedicht: Ein einfacher Satz wie „Die Sonne scheint“ ist zu gewöhnlich. Wird daraus jedoch ein eigenständiger, bildhafter Text mit kreativem Ausdruck, kann dieser durchaus unter den Schutz des Urheberrechts fallen.

Entscheidend ist immer: Die kreative Leistung einer bestimmten Person muss erkennbar sein. Letztendlich ist jedoch zu beachten, dass die Bewertung, ob eine Schöpfungshöhe tatsächlich erreicht ist, stark vom Einzelfall abhängt. Der Begriff selbst ist recht unbestimmt und bietet viel Interpretationsspielraum – was zu Diskussionen und Konflikten führen kann, die in letzter Instanz gerichtlich gelöst werden müssen.

KI-generierte Inhalte stellen hier jedoch eine Besonderheit dar: Da sie nicht von einem Menschen, sondern von einer Maschine erzeugt wurden, gelten sie im klassischen Sinne nicht als urheberrechtlich geschützt. Nach derzeitiger Rechtslage kann nur ein Mensch Urheber*in sein. Das bedeutet: Werke, die vollständig durch KI erzeugt wurden, genießen keinen originären Schutz durch das Urheberrecht. Allerdings kann es zu rechtlichen Auseinandersetzungen kommen, wenn KI-Inhalte bestehenden geschützten Werken zu ähnlich sind oder auf solchen basieren.

KI UND DAS URHEBERRECHT – ZWEI PERSPEKTIVEN

Bevor wir uns aber mit diesen rechtlichen Grauzonen beschäftigen, müssen wir zuerst einen Blick auf die Perspektiven des Urheberrechts im Themenfeld der Künstlichen Intelligenz werfen.

Nutzung von geschützten Inhalten zum Training von KI

Beim Training von Künstlicher Intelligenz kommen häufig große Datenmengen zum Einsatz – darunter auch Texte, Bilder oder Musikstücke, die urheberrechtlich geschützt sein können. Um diese Inhalte systematisch auswerten zu können, wird auf ein Verfahren namens „Text und Data Mining“ (TDM) zurückgegriffen. Dabei handelt es sich um das automatisierte Durchsuchen und Analysieren von Daten durch Maschinen.

Nach deutschem Urheberrecht (§ 44b Urheberrechtsgesetz) ist TDM grundsätzlich zulässig, sofern die genutzten Inhalte rechtmäßig zugänglich sind und die Rechteinhaber*innen der Nutzung nicht ausdrücklich widersprochen haben. Für wissenschaftliche Einrichtungen und sogenannte Gedächtnisinstitutionen (wie Bibliotheken oder Museen) gelten hier großzügigere Regelungen: Sie dürfen TDM auch dann anwenden, wenn ein solcher Widerspruch vorliegt (§ 60d UrhG).

Für kommerzielle Anbieter*innen – darunter fallen viele Entwickler*innen von KI-Systemen – ist die Lage strenger: Sie dürfen Inhalte nur dann per TDM analysieren, wenn kein erklärtes Nutzungsverbot (sogenanntes Opt-out) vorliegt. Dieses Opt-out muss von den Rechteinhaber*innen aktiv in maschinenlesbarer Form beispielsweise in den Metadaten oder auch in den Nutzungsbedingungen gesetzt werden. Dabei handelt es sich um eine digitale Kennzeichnung, die automatisiert von Suchmaschinen oder KI-Systemen erkannt und ausgelesen werden kann.

Nutzung KI-generierter Inhalte für eigen Zwecke

Auch bei der Verwendung von mittels KI erzeugten Inhalten bestehen rechtliche Risiken. Grundsätzlich fallen KI-Outputs – also von Maschinen generierte Ergebnisse – nicht unter den klassischen Urheberrechtsschutz (§ 2 Urheberrechtsgesetz), da eine persönliche geistige Schöpfung natürlicher Personen fehlt. Doch auch wenn solche Inhalte formal nicht geschützt sind, können vertragliche Einschränkungen durch die Lizenzbedingungen des jeweiligen KI-Dienstes eine entscheidende Rolle spielen.

Diese Lizenzbedingungen regeln oft, wie die generierten Inhalte genutzt werden dürfen – etwa ob eine kommerzielle Nutzung erlaubt ist oder ob bestimmte Quellenangaben erforderlich sind. Wer gegen solche Vorgaben verstößt, riskiert Abmahnungen oder andere rechtliche Konsequenzen, auch wenn kein Urheberrecht im engeren Sinne verletzt wurde. Es ist daher wichtig, die Nutzungsbedingungen der eingesetzten KI-Plattform sorgfältig zu prüfen und bei Unsicherheiten auf alternative Quellen oder rechtlich geprüfte Inhalte zurückzugreifen.

Gerade im Bildungsbereich und in anderen nicht-kommerziellen Zusammenhängen stellt sich die Frage, wie man rechtlich mit KI-Inhalten umgehen soll – ohne tiefgreifende juristische Prüfungen durchführen zu müssen. Da solche Prüfungen in der Praxis kaum leistbar sind, bietet sich ein pragmatischer Umgang an. Verantwortliche sollten deshalb darauf achten, ob generierte Inhalte auffällige Ähnlichkeiten mit bekannten Werken aufweisen – etwa durch typische Formulierungen, Motive oder eine bestimmte stilistische

Handschrift. Wo Unsicherheiten bestehen, empfiehlt es sich, auf eigene Formulierungen zurückzugreifen oder mit offen lizenzierten Inhalten zu arbeiten.

Darüber hinaus kann Transparenz über die Herkunft der Inhalte zusätzliche Klarheit schaffen. Dazu gehört etwa die Nennung des verwendeten KI-Dienstes sowie eine allgemeine Beschreibung des Prompts. Solche Hinweise ersetzen zwar keine rechtliche Prüfung, stärken aber die Glaubwürdigkeit – insbesondere im pädagogischen Raum – und helfen, Missverständnisse im Vorfeld zu vermeiden.

RECHTLICHE GRAUZONEN UND AKTUELLE ENTWICKLUNGEN

Die aktuelle Rechtslage führt zunehmend zu Diskussionen und Streitfällen, insbesondere wenn KI-Tools Werke erzeugen, die bestehenden urheberrechtlich geschützten Inhalten ähneln oder aus urheberrechtlich geschützten Quellen stammen, die entgegen der gesetzlichen Regelung des TDM trainiert wurden. Ein sehr aktuelles Beispiel aus dem Jahr 2025 ist der Fall GEMA vs. Suno, bei dem es um Musik geht, die mit KI generiert wurde. Die GEMA argumentiert, dass die zugrunde liegenden Trainingsdaten urheberrechtlich geschützte Musikstücke enthalten und somit Rechte verletzt sein könnten. Zum Zeitpunkt der Veröffentlichung dieses Artikels gibt es in diesem Zusammenhang jedoch noch keine wegweisenden oder allgemein übertragbaren Urteile, sodass die Rechtslage weiterhin unsicher bleibt.

Ein Schritt in die richtige Richtung ist der AI Act der Europäischen Union. Diese Verordnung soll unter anderem regeln, wie KI-Systeme kategorisiert, zugelassen und überwacht werden. Zwar liegt der Fokus des AI Acts auf Transparenz, Sicherheit und ethischen Standards, doch enthält er auch Ansatzpunkte, die urheberrechtlich relevant sein können. Besonders die Pflicht zur Dokumentation der verwendeten Trainingsdaten und zur Offenlegung der Funktionsweise von KI-Systemen kann in Zukunft dazu beitragen, Urheberrechtsverletzungen besser nachvollziehen und bewerten zu können.

KENNZEICHNUNGSPFLICHT – TRANSPARENZ BEI KI-NUTZUNG

Eine der häufigsten Fragen im Umgang mit KI-Tools lautet: „Woran erkenne ich eigentlich, ob ein Inhalt von einer Maschine stammt – und muss ich das offenlegen?“ Mit der zunehmenden Qualität und Täuschungsfähigkeit von KI-generierten Texten, Bildern, Audios oder Videos wird diese Frage immer drängender. Oft sind maschinell erstellte Inhalte kaum noch von menschlich geschaffenen zu unterscheiden – gerade in der Bildungsarbeit, wo Vertrauen und Transparenz zentrale Werte sind, wird dieses Spannungsfeld sichtbar.

Daraus ergibt sich die Forderung nach einer **Kennzeichnungspflicht**: Nutzer*innen sollen erkennen können, ob ein Inhalt durch KI erzeugt wurde. Die europäische KI-Verordnung – der sogenannte AI Act – greift dieses Anliegen auf und schafft ab dem 02. August 2026 mit Artikel 50 eine wichtige Grundlage für mehr Transparenz. Dort ist festgelegt, dass Anbieter*innen von KI-Systemen verpflichtet sind, darauf hinzuweisen, wenn Inhalte durch Künstliche Intelligenz erzeugt, verändert oder manipuliert wurden – insbesondere, wenn eine Täuschung von Personen nicht ausgeschlossen werden kann. Ziel dieser Regelung ist es, maschinell erzeugte Inhalte erkennbar zu machen und so Vertrauen sowie einen verantwortungsvollen Umgang mit KI zu fördern. Diese Transparenzpflicht gilt insbesondere

für KI-Systeme, die mit einem erhöhten Risiko verbunden sind – etwa bei täuschend echten Texten, Bildern oder synthetischen Stimmen. Für die Praxis bedeutet das:

- KI-generierte Inhalte müssen stets als künstlich erzeugt oder manipuliert gekennzeichnet werden, wenn sie ohne menschliche Redaktion direkt veröffentlicht werden.
- Für redaktionell verantwortete Texte entfällt die Kennzeichnungspflicht, sofern eine natürliche oder juristische Person die redaktionelle Verantwortung übernimmt; für synthetische Audio-, Bild- oder Videoinhalte bleibt die Kennzeichnung jedoch in jedem Fall erforderlich.
- Erzeugte Audio-, Bild- oder Videoinhalte mit täuschend echt wirkendem Charakter („Deep Fakes“) müssen stets als künstlich generiert gekennzeichnet werden, auch wenn sie vor Veröffentlichung von einer menschlichen Redaktion geprüft wurden.

Wir müssen uns eins immer vor Augen führen: Das deutsche Urheberrechtsgesetz (UrhG) wurde 1965 aus zwei bedeutend älteren Gesetzen geschaffen: aus dem Gesetz betreffend das Urheberrecht an Werken der Literatur und der Tonkunst (LUG) von 1901 und dem Gesetz betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie (KUG) von 1907. Damit basiert unser heutiges Urheberrecht auf Grundlagen, die aus einer analogen Zeit stammen, in der digitale Technologien noch nicht absehbar waren. Die technische Entwicklung – insbesondere im Bereich der Künstlichen Intelligenz – verläuft jedoch heute in einem Tempo, das die Anpassungsfähigkeit des bestehenden Rechtsrahmens bei Weitem übersteigt. Viele Fragen, die sich aus der Nutzung von KI-Systemen ergeben, sind bislang ungeklärt und werden erst im Laufe der kommenden Jahre durch rechtliche Auseinandersetzungen und gerichtliche Entscheidungen konkretisiert werden.

Diese rechtlichen Anforderungen lassen sich gut in die pädagogische Praxis übertragen. Fachkräfte sollten KI als unterstützendes Werkzeug nutzen – nicht als alleinige Quelle von Inhalten – und sich ihrer Verantwortung bewusst sein. Inhalte, die mithilfe von KI erstellt wurden, sollten klar als solche gekennzeichnet werden, insbesondere wenn sie täuschend echt wirken oder ohne redaktionelle Prüfung veröffentlicht werden. Gleichzeitig ist es sinnvoll, Angaben zu verwendeten Prompts oder Trainingsdaten zu dokumentieren. Auch Lernende profitieren von dieser Transparenz, wenn offengelegt wird, wo KI zum Einsatz kam. So wird Vertrauen geschaffen, Urheberrechtsbewusstsein gefördert und die Nutzung von KI in einem ethisch verantwortungsvollen Rahmen vermittelt. Multiplikator*innen im Bildungsbereich können dabei eine Schlüsselrolle übernehmen, indem sie diese Prinzipien aktiv in ihre Bildungsangebote integrieren und als Vorbilder für einen reflektierten Umgang mit KI-Inhalten wirken.

Dieser Text gibt eine praxisnahe Einführung in zentrale Aspekte des Urheberrechts, erhebt jedoch keinen Anspruch auf Vollständigkeit. Für eine verbindliche rechtliche Einschätzung oder Beratung im Einzelfall empfiehlt sich stets die Konsultation einer Fachanwältin oder eines Fachanwalts für Urheber- und Medienrecht.

MENSCHSEIN IM KI-ZEITALTER – BILDUNGSAUFGABEN ANGESICHTS EINES VERÄNDERTEN MENSCHENBILDS

von Lukas Spahlinger

Wer oder was ist der Mensch in einer Welt, in der Maschinen zunehmend Aufgaben übernehmen, die lange als Ausdruck menschlicher Vernunft oder Kreativität galten? Maschinen verfassen Gedichte oder komponieren Musik. Algorithmen erstellen in kürzester Zeit medizinische Diagnosen oder generieren Reden, die emotional berühren und überzeugend wirken. Solche Szenarien sind längst Realität und führen dazu, grundlegende Fragen nach dem Wesen des Menschen neu zu stellen. Schon in der Vergangenheit haben technische Innovationen dazu geführt, dass Menschenbilder infrage gestellt wurden.

In Bildungsprozessen gilt es daher, diese Veränderungen nicht nur technisch zu begleiten, sondern auch emotional, ethisch und philosophisch zu reflektieren. Was macht das Menschsein aus? Welche Werte, Fähigkeiten und Haltungen sollten im Zeitalter der Künstlichen Intelligenz bewahrt, welche weiterentwickelt werden? Welche Kompetenzen sind erforderlich, um diesen Wandel nicht nur zu überstehen, sondern aktiv und mündig mitzugestalten.

WAS DEN MENSCHEN AUSMACHT UND WIE KI IHN HERAUSFORDERT

Die Unterscheidung zwischen Mensch und Maschine war lange relativ klar: Bewusstsein, Moral, Kreativität und Sprache galten als exklusiv menschliche Fähigkeiten. Doch mit dem Fortschritt generativer KI-Systeme wird diese Abgrenzung zunehmend unscharf. Lange war der Mensch als dasjenige Wesen definiert, das Vernunft, Sprache und moralisches Handeln in sich vereint. KI-Systeme, die Texte verfassen, medizinische Diagnosen stellen oder strategische Entscheidungen treffen stellen diese Zuschreibungen infrage.

Immanuel Kant betonte zum Beispiel die Mündigkeit des Menschen – also seine Fähigkeit, sich seines Verstandes ohne Anleitung eines anderen zu bedienen. Heute stellt sich die Frage, ob diese Fähigkeit durch KI eher gestärkt oder untergraben wird. Das ist zum jetzigen Zeitpunkt unklar. Zugleich hat Michel Foucault gezeigt, dass Macht nicht nur durch Zwang wirkt, sondern auch durch Wissen und Normierung. KI-Systeme, die Verhalten analysieren, Entscheidungen vorstrukturieren oder Inhalte filtern, formen mitunter das Selbstverständnis der Menschen – oft ohne dass dies bewusst geschieht.

KRITISCH BEGLEITEN: POST- UND TRANSHUMANISMUS IM DISKURS

Bewegungen wie der Transhumanismus propagieren das Ideal, den Menschen durch Technik zu optimieren oder gar zu überwinden. Der Posthumanismus fordert, die Sonderstellung des Menschen als Maß aller Dinge zu hinterfragen und ihn als Teil vernetzter, nichtmenschlicher Systeme zu denken. Beide Strömungen bieten wichtige Impulse – etwa zur Infragestellung überkommener Grenzziehungen –, bergen aber auch Herausforderungen: etwa die Gefahr, körperliche Endlichkeit, Verletzlichkeit oder moralische Verantwortung zu entwerten. Bildung braucht hier einen kritischen Rahmen, um spekulative Zukunftsbilder ethisch einordnen und hinterfragen zu können.

VERÄNDERT: DAS VERHÄLTNIS ZU UMWELT, TIER UND TRANSZENDENZ

Der Mensch als soziales Wesen ist nie unabhängig von der Umwelt und passt sich an ihre Bedingungen an. Gleichzeitig fordern diese Technologien dazu heraus, das Menschenbild weniger anthropozentrisch zu denken. Die feministische Wissenschaftstheoretikerin Donna Haraway hat mit ihrem „Cyborg“-Konzept auf die Auflösung starrer Grenzen zwischen Natur und Technik, Mensch und Maschine hingewiesen. KI-gestützte Systeme ermöglichen neue Formen des Zusammenspiels zwischen Natur, Mensch, Technik und Tieren – etwa im Artenschutz oder in der Umweltüberwachung. Gleichzeitig verschwimmt die klare Abgrenzung des Menschen als „vernunftbegabtes Gegenüber“ der Natur. Das anthropozentrische Selbstbild wird dadurch relativiert.

Ähnlich zeigen sich Verschiebungen im religiösen Bereich: Der Mensch hat mit der Künstlichen Intelligenz eine Technologie geschaffen, die Charakteristika aufweist, die bisher nur einer göttlichen Kraft zugeschrieben wurden: allmächtig, allwissend oder allgegenwärtig zu sein. Doch der genaue Blick auf Künstliche Intelligenz zeigt auch ihre Grenzen auf. Large Learning Models (LLMs) repräsentieren menschliche Wertesysteme. Sie funktionieren nur, solange sie Zugriff auf das Internet haben. Außerdem sind sie fehleranfällig. Die Mängel werden problematisch, wenn KI-Systemen blind vertraut wird. Bildungsprozesse sollten diesen Wandel nicht nur beschreiben, sondern ethisch einordnen – etwa im Blick auf Authentizität, Empathie oder das Verständnis von Schöpfungsverantwortung.

MÜNDIGKEIT IM DIGITALEN ALLTAG

KI kann Menschen von Routineaufgaben befreien – zugleich besteht das Risiko, dass Urteilsfähigkeit und Verantwortung an automatisierte Systeme abgegeben werden. Das protestantische Bildungsideal der Mündigkeit – mit dem Ziel, selbst zu denken, zu urteilen und zu handeln – bleibt hier aktuell. Bildung muss befähigen, nicht nur KI-Systeme zu nutzen, sondern deren Vorschläge kritisch zu bewerten. Eine zentrale Aufgabe ist daher die Förderung digitaler Mündigkeit: der Fähigkeit, Technologie souverän, verantwortungsbewusst und reflektiert in das eigene Leben zu integrieren.

KI-KOMPETENZ BRAUCHT REFLEXIONSRÄUME

Der Einsatz von KI-Technologien stellt nicht nur technische und rechtliche Fragen – er betrifft die Grundfrage, wie wir uns als Menschen verstehen und was wir in Zukunft sein wollen. Medizinische Diagnosen, die in manchen Bereichen akkurater von KI-Systemen getroffen werden, der Einsatz von KI in der Rechtsprechung oder bei politischen Entscheidungsprozessen zeigen, wie tiefgreifend solche Systeme in gesellschaftliche Strukturen eingreifen können. Wenn Maschinen Aufgaben übertragen werden, die bisher nur gut ausgebildeten Menschen zugetraut wurden, verändert das das Selbstverständnis des Menschen. Eine reflektierte KI-Kompetenz muss deshalb stets philosophische, ethische und theologische Dimensionen einbeziehen, um nicht nur funktionale Vorteile zu betrachten, sondern auch Fragen nach Würde, Verantwortung und Sinn zu adressieren. Bildungsarbeit sollte Räume schaffen, in denen Menschen diese Fragen anhand konkreter Szenarien diskutieren – etwa, wie Empathie und moralisches Urteilsvermögen in einer zunehmend automatisierten Welt bewahrt werden können. Die Technik allein gibt keine Antworten darauf, wie ein gutes und gerechtes Zusammenleben gestaltet werden soll. Diese Aufgabe bleibt zutiefst menschlich.

VON PIEPSENDEN TAMAGOTCHIS UND REDSELIGEN ROBOTERN: EINE KLEINE GESCHICHTE DER MENSCH-MASCHINE- INTERAKTION

von Nina Voborsky

„Die Gefahr, dass der Computer so wird wie der Mensch, ist nicht so groß wie die Gefahr, dass der Mensch so wird wie der Computer.“ (Konrad Zuse)

Wer in den 1990ern ein Tamagotchi besaß, erinnert sich vielleicht: Mitten in der Nacht piepte das kleine Ei unerbittlich, weil es Hunger hatte oder Aufmerksamkeit wollte. Millionen Kinder und Erwachsene sprangen damals aus dem Bett, um ein digitales Wesen zu füttern, das weder einen Magen noch Gefühle hatte – und behandelten es doch wie ein Lebewesen. So absurd es klingt: Diese Fürsorge ist kein Einzelfall, sondern steht in einer langen Tradition. Menschen haben schon immer die Neigung gezeigt, Maschinen und Objekten menschliche Eigenschaften zuzuschreiben. Wir reden mit Autos, die nicht anspringen wollen, schimpfen mit dem Computer oder bedanken uns bei Siri. Die Geschichte dieser Vermenschlichung von Technik reicht weit zurück – und erzählt mindestens so viel über uns selbst wie über die Maschinen.

ERSTE WUNDERWERKE: AUTOMATEN UND SPRECHENDE MASCHINEN

Bereits im 18. Jahrhundert begeisterten mechanische Automaten die Gesellschaft. Der berühmte „Schachtürke“, ein angeblich selbstständig Schach spielender Apparat, oder die Schreibautomaten von Jacques de Vaucanson wirkten, als ob sie denken könnten. Auch wenn es oft raffinierte Täuschungen waren, faszinierte die Vorstellung, eine Maschine könne wie ein Mensch handeln. Einen weiteren Sprung markierte Thomas Edisons Phonograph von 1877: Zum ersten Mal ertönte eine menschliche Stimme aus einem Gerät. Für viele klang das unheimlich – ein Stück Technik, das sprechen konnte, als sei es belebt. Später, im 20. Jahrhundert, wurde die automatische Zeitansage am Telefon zum alltäglichen Begleiter. Der „Damenstimme“ am Apparat wurden Charakterzüge zugeschrieben – freundlich, streng oder sachlich –, obwohl sie nichts weiter als eine Aufnahme war.

STIMMEN, KNÖPFE, GEFÜHLE: INTERAKTION IM ALLTAG

In den 1960er-Jahren hielten sprechende Puppen Einzug ins Kinderzimmer. Sie gaben Kindern nicht nur eine Stimme zum Spielen, sondern auch ein Gegenüber, das Gefühle hervorrief. Auch wenn die Batterie-Klappe im Rücken des Spielzeugs durchaus ambivalente Gefühle hervorrief. Noch deutlicher zeigte sich dieser Effekt in den 1990er-Jahren: Mit dem Tamagotchi entstand ein digitales Haustier, das gehegt und gepflegt werden musste. Millionen Menschen entwickelten emotionale Bindungen zu einem Display mit wenigen Pixeln. Parallel dazu wurden Computer durch grafische Benutzeroberflächen und Spiele wie Text-Adventures zu interaktiven Partnern. Technik war nicht mehr nur Werkzeug, sondern konnte eine Art soziales Gegenüber darstellen.

VON SIRI BIS ALEXA: SPRACHE ALS SCHNITTSTELLE

2011 betrat mit Siri eine neue Akteurin die Bühne: Zum ersten Mal konnten Nutzer*innen ihr Smartphone direkt ansprechen – und bekamen eine freundliche, manchmal witzige Antwort. Schnell zeigte sich: Sie bedankten sich höflich, ärgerten sich über Missverständnisse oder versuchten, Smalltalk zu führen. Mit Alexa, Google Assistant und anderen Sprachassistenten wurde diese Form der Interaktion alltäglich. Technik bekam eine Stimme – und damit schrieben wir ihr automatisch auch eine Persönlichkeit zu.

SOZIALE ROBOTER UND VIRTUELLE BEGLEITER

In den letzten Jahren ging die Entwicklung weiter. Roboter wie „Paro“, eine kuschelige Robbe für die Altenpflege, oder der humanoide „Pepper“ wurden gezielt so gestaltet, dass sie Nähe suggerieren und Emotionen hervorrufen. Chatbots wie Replika oder KI-Systeme, die auf Sprachmodellen basieren, werden inzwischen als Gesprächspartner erlebt – nicht nur für Informationen, sondern auch für Trost, Zuspruch oder kreative Zusammenarbeit. Damit verschwimmen die Grenzen zwischen Werkzeug und Beziehungspartner. Für manche sind solche Systeme hilfreiche Assistenten, für andere fast virtuelle Freunde. Und zugleich bleibt die Irritation: Wie „echt“ kann ein Gespräch mit einer Maschine sein?

KI-SYSTEME: GESPRÄCHSPARTNER ODER SPIEGEL?

Mit den neuesten KI-Systemen kommt eine neue Qualität hinzu. Sprachassistenten wie Siri oder Alexa reagierten noch recht starr auf Befehle. Große Sprachmodelle wie ChatGPT hingegen sind darauf ausgelegt, im Dialog zu bleiben. Sie hören nicht einfach auf, wenn unsere Frage beantwortet ist, sondern schlagen Themen vor, fragen zurück oder bauen Erklärungen aus. Das verändert die Dynamik: Wir erleben weniger eine Maschine, die „auf Knopfdruck“ funktioniert, sondern ein Gegenüber, das sich wie ein Gesprächspartner verhält.

WARUM WIR TECHNIK VERMENSCHLICHEN

Dass wir Maschinen Eigenschaften oder sogar Gefühle zuschreiben, ist tief in uns angelegt. Unser Gehirn ist von Natur aus darauf trainiert, soziale Signale wie Gesichter, Stimmen und Bewegungen zu erkennen – oft auch dort, wo keine sind. Dieses „soziale Radar“ war in der Evolution überlebenswichtig: Lieber einmal zu oft eine Absicht vermuten als einmal zu wenig. Technik nutzt genau diesen Reflex. Eine freundliche Stimme aus dem Lautsprecher oder das Piepen eines Tamagotchis reichen aus, damit wir Zuwendung, Hilfslosigkeit oder Persönlichkeit hineininterpretieren. Popkulturelle Erzählungen – von Pinocchio bis zu Science-Fiction-Filmen – verstärken zusätzlich die Vorstellung, Maschinen könnten potenzielle Partner*innen sein. Doch je menschenähnlicher Maschinen wirken, desto stärker schwankt unsere Reaktion zwischen Nähe und Unbehagen.

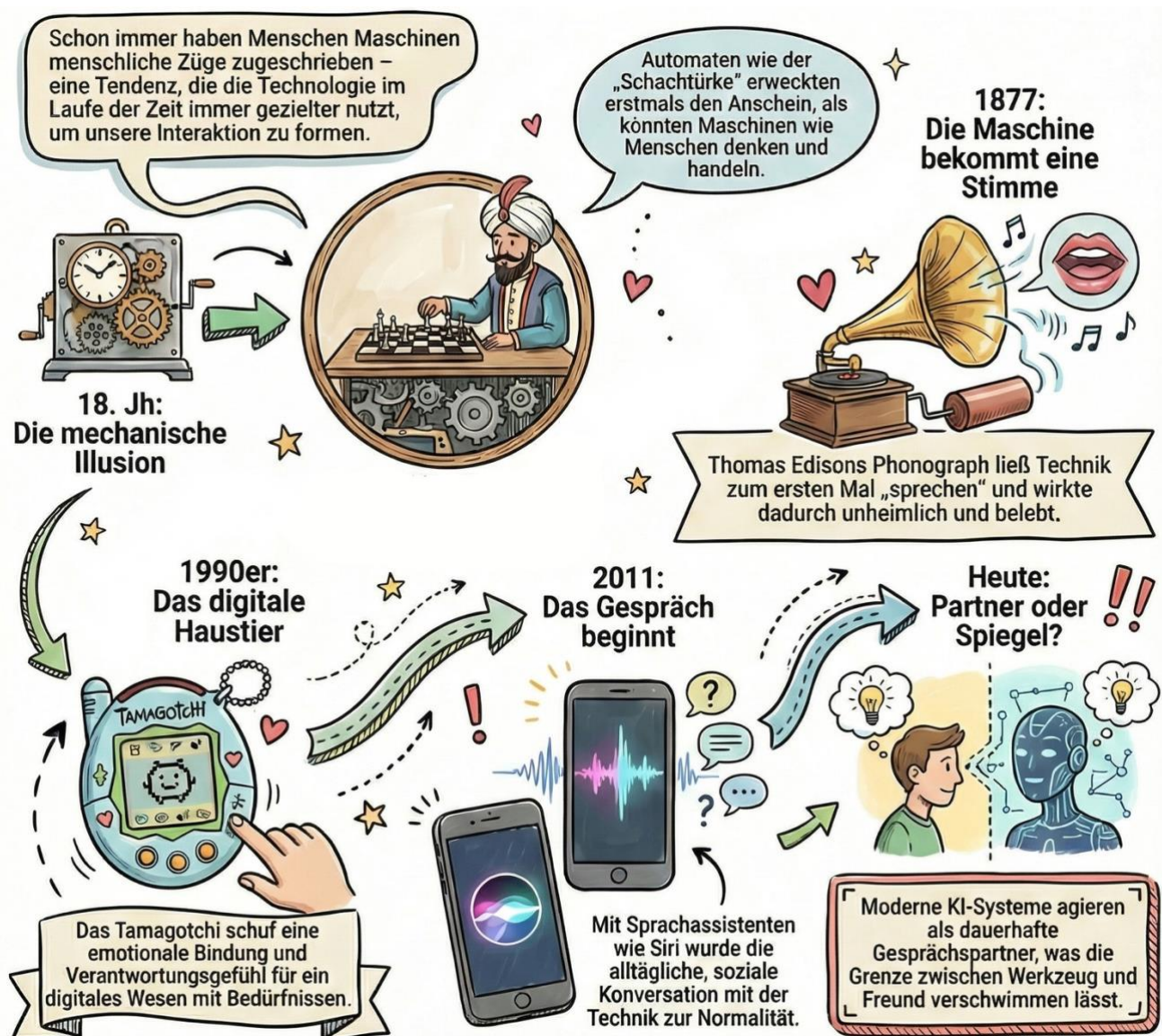
GEGENWART UND AUSBLICK

Heute sind Mensch-Maschine-Interaktionen komplexer als je zuvor. Wir sprechen mit Chatbots, lassen uns Musikempfehlungen geben oder diskutieren mit digitalen Assistenten.

Autos reagieren auf Sprachbefehle, Kühlschränke schlagen Rezepte vor, und manche Menschen führen lange Gespräche mit KI-Systemen.

Die entscheidende Frage ist daher nicht, ob wir Maschinen vermenschlichen, sondern wie bewusst wir es tun. Denn nur wenn wir verstehen, welche Mechanismen uns dazu bringen, Technik als sozialen Partner zu erleben, können wir souverän mit dieser Tendenz umgehen. Das stellt uns als Gesellschaft vor neue Anforderungen: Wir brauchen technologische Kompetenz, um Funktionsweise und Grenzen von KI-Systemen zu verstehen; kritische Reflexionskompetenz, um Projektionen von Realität unterscheiden zu können; ethische Kompetenz, um die Gestaltung und den Einsatz von KI verantwortungsvoll zu steuern; und nicht zuletzt soziale Kompetenz, um zu bewahren, dass echte zwischenmenschliche Beziehungen nicht durch Technik ersetzt werden.

Vielleicht ist es deshalb die wichtigste Frage unserer Zeit: Welche Fähigkeiten müssen wir entwickeln, um in einer Welt, in der Maschinen immer menschlicher erscheinen, selbst menschlich zu bleiben?



WIE ANTHROPOMORPHISIERUNG VON KI UNSERE WAHRNEHMUNG VON MASCHINEN VERÄNDERT

von Lukas Spahlinger

Künstliche Intelligenz begegnet uns zunehmend in Formen, die auf Nähe, Verständlichkeit und Dialog angelegt sind: Sprachassistenten, Chatbots, soziale Roboter oder digitale Begleitdienste. Diese Systeme sprechen, reagieren, formulieren Vorschläge und passen sich an Nutzerinnen und Nutzer an. Damit verschiebt sich die Wahrnehmung von Technologie grundlegend. Was früher als Werkzeug galt, wird heute häufiger als eine Art Gegenüber verstanden.

Diese Entwicklung ist kein rein technisches Phänomen. Sie beruht auf einem Zusammenspiel aus Gestaltungsentscheidungen, kulturellen Erwartungen, psychologischen Mechanismen und gesellschaftlichen Diskursen. Die Vermenschlichung von KI – also ihre Anthropomorphisierung – beeinflusst, wie wir Maschinen einschätzen, worauf wir uns verlassen, welchen Fähigkeiten wir ihnen zutrauen und wem wir Verantwortung zuschreiben.

Der folgende Text untersucht, warum Anthropomorphisierung historisch tief verankert ist, welche Interessen hinter anthropomorphen Designs stehen, wie sich wechselseitige Verstärkungseffekte („anthropomorphic feedback loop“) entwickeln und welche Paradoxien und Risikodynamiken dadurch entstehen. Abschließend wird betrachtet, wie sich Erwartungshaltungen formen und warum ein reflektierter Umgang gesellschaftlich notwendig ist.

HISTORISCHE HINTERGRÜNDE: DIE ALTE SEHNSUCHT NACH MENSCHLICHEN MASCHINEN

Die Tendenz, Nichtmenschliches mit menschlichen Eigenschaften auszustatten, ist kein neues Phänomen. Sie begleitet die menschliche Kultur seit ihren Anfängen. In vielen Religionen, Mythen und Weltdeutungen wurden Naturgewalten personifiziert, um sie verständlicher zu machen. Anthropomorphisierung dient als eine Strategie, komplexe oder unkontrollierbare Vorgänge in vertraute Muster zu überführen.

Auch in der Technikgeschichte spielt diese Tendenz eine bedeutende Rolle. Bereits in der Antike entstanden mechanische Figuren, die menschliche Bewegungen imitierten. In der Renaissance und im Barock verfeinerten Kunsthandwerker Automaten, die als künstliche Tiere und menschenähnliche Puppen auftraten. Mit Beginn der Industrialisierung verlagerte sich der Fokus: Maschinen wurden nun nicht mehr nur nachgebildet, sondern als „Kollegen“ oder „Arbeiter“ imaginiert, die menschliche Fähigkeiten ergänzen oder entlasten.

Im 20. Jahrhundert prägte die Science-Fiction die Vorstellung menschähnlicher Maschinen besonders stark. Figuren wie HAL 9000 aus dem Film "2001: Odysee im Weltraum", der Android "Data" aus dem Universum von Star Trek oder Robotergefährten zahlreicher Kinder- und Zukunftsfilme beeinflussen bis heute die Erwartungshaltung gegenüber moderner KI. Sie vermitteln, dass Maschinen denkfähig, fühlend oder moralisch urteilsfähig sein könnten.

Diese historischen und kulturellen Muster wirken bis in die Gegenwart hinein. Die moderne Anthropomorphisierung digitaler Systeme knüpft an diese lange Tradition an – und sie aktiviert die Sehnsucht nach Resonanz, Verlässlichkeit und einem idealisierten technischen Gegenüber. KI erscheint damit weniger als Tool, sondern zunehmend als Teil sozialer Wirklichkeit.

WARUM ANTHROPOMORPHE DESIGNS ENTWICKELT WERDEN: INTERESSEN, FUNKTIONEN UND RISIKEN

Anthropomorphisierung entsteht nicht zufällig, sie ist häufig das Ergebnis klarer Gestaltungsstrategien. Dabei gilt es verschiedene Faktoren zu beachten, die Treiber von Anthropomorphisierung in der Technologieentwicklung sind.

KOMMERZIELLE INTERESSEN UND NUTZUNGSOPTIMIERUNG

Viele KI-Produkte sind auf enge, vertrauensbasierte Interaktion angelegt. Unternehmen nutzen menschähnliche Stimmen, Namen oder Persönlichkeitsprofile, weil sie die Nutzung intensivieren und Hemmschwellen reduzieren. Eine „freundliche“ KI wirkt einladender als ein technisches Interface. Dies steigert nicht nur die Nutzungsdauer, sondern auch die Bereitschaft, personenbezogene Daten zu teilen oder Empfehlungen zu folgen.

Diese Logik birgt Risiken. Sie erzeugt eine Vertrauensarchitektur, die auf sozialer Intuition statt auf technischem Verständnis beruht. Die Verantwortung verschiebt sich subtil: Wenn „die KI“ eine Empfehlung gibt, tritt das dahinterliegende Unternehmen aus dem Blick. Fehlentscheidungen werden nicht als Ergebnis eines Geschäftsmodells oder eines Trainingsfehlers erkannt, sondern als scheinbar autonome Handlung eines Systems.

VERBESSERTER FUNKTIONALITÄT: BENUTZERFREUNDLICHKEIT UND INTUITIVE INTERAKTION

Anthropomorphe Elemente erleichtern die Bedienung komplexer Technologien. Eine sprechende Schnittstelle ist oft verständlicher als ein technisches Menü. Maschinen, die Blickkontakt andeuten oder Pausen in Gesprächen imitieren, orientieren sich an der menschlichen Kommunikation, nicht weil sie „menschlich“ sind, sondern weil Menschen sozial kommunizieren.

In Pflege, Therapie und sozialer Unterstützung wird menschenähnliche Gestaltung gezielt eingesetzt, weil sie funktional ist. Menschen mit Demenz, kognitiven Einschränkungen oder in Einsamkeit reagieren häufig besser auf Systeme, die vertraute soziale Muster anbieten. Roboter und Chatbots können Struktur geben, beruhigen oder motivieren. Hier wird deutlich: Anthropomorphisierung kann einen konkreten Nutzen haben. Doch zugleich entstehen ethische Spannungen. Je menschlicher ein System wirkt, desto höher das Risiko emotionaler Abhängigkeit oder Täuschung. Die Grenze zwischen „Unterstützung“ und „Simulation sozialer Beziehungen“ bleibt unscharf – und muss reflektiert gestaltet werden.

Dieser Ansatz verschiebt jedoch Erwartungen. Aus einer technisch-pragmatischen Gestaltung wird leicht eine sozial-emotionale Zuschreibung, bei der Nutzerinnen und Nutzer dem System Absichten, Empathie oder Verständnis unterstellen.

DYNAMIKEN DER ANTHROPOMORPHISIERUNG: DER ANTHROPOMORPHIC FEEDBACK LOOP

Anthropomorphisierung ist kein statisches Phänomen, sondern ein sich selbst verstärkender Prozess. Dieser Prozess lässt sich als „anthropomorphic feedback loop“ beschreiben:

- **Designentscheidungen:** KI erhält Stimme, Namen, Persönlichkeit oder soziale Interaktionsmuster.
- **Nutzerreaktionen:** Menschen behandeln das System zunehmend wie ein Gegenüber.
- **Psychologische Zuschreibungen:** Es entstehen Erwartungen an Intention, Empathie oder moralisches Verhalten.
- **Ökonomische Verstärkung:** Unternehmen verstärken die menschähnlichen Elemente, weil dies Engagement und Bindung erhöht.
- **Rückkopplung:** Die wachsende Nähe führt zu noch stärkeren Erwartungen der Nutzerinnen und Nutzer.

Dieser Loop führt zu systematischen Verzerrungen: Fähigkeiten werden überschätzt, Grenzen unterschätzt, und Maschinen wirken kompetenter, als sie sind.

PARADOXIEN: ZWISCHEN MASCHINEN-MENSCH UND MASCHINEN-BEDROHUNG

Anthropomorphisierung erzeugt paradoxe Wahrnehmungsmuster. Einerseits erscheinen KI-Systeme als vertraute, hilfreiche und teilweise emotionale Begleiter. Andererseits wächst die Sorge, sie könnten Kontrolle übernehmen, Arbeitsplätze verdrängen oder gesellschaftliche Prozesse beeinflussen.

Die widersprüchlichen Zuschreibungen entstehen aus derselben Quelle: der menschlichen Tendenz, sozialen Logiken auf technische Systeme zu übertragen. Ein System, das kommunikativ wirkt, erscheint gleichzeitig als potenziell kompetenter und potenziell gefährlicher als ein rein technisches Werkzeug.

Besonders deutlich wird dies im Bereich hochautomatisierter Systeme oder generativer KI. Je flüssiger, kontextbezogener und „natürlicher“ ein System spricht, desto stärker wächst die Illusion eines eigenständigen Bewusstseins oder Ziels. Gleichzeitig erzeugt dieselbe Illusion Unsicherheit und Kontrollverlust.

SPRACHE, ZUSCHREIBUNG UND VERANTWORTUNG

Sprache ist ein zentraler Treiber anthropomorpher Wahrnehmung. Formulierungen wie „die KI entscheidet“, „der Algorithmus will“ oder „das System versteht“ übertragen menschliche Kategorien auf maschinelle Prozesse. Diese Wortwahl ist weit verbreitet – in Medien, Marketing und Alltagskommunikation –, aber sie führt zu Fehlschlüssen.

Sprache formt die Vorstellungskraft. Wenn wir KI systematisch in Begriffen und Beschreibungen vermenschlichen, entsteht ungewollt ein Bild von Autonomie und Bewusstsein. Die Folge ist eine Verschiebung der Verantwortungszuschreibungen. Nicht menschliche Entscheidungen werden für Diskriminierungsmuster verantwortlich, sondern die Künstliche Intelligenz. Ein bewusster Umgang mit Sprache ist daher ein Bestandteil verantwortungsvollem Umgangs mit Künstlicher Intelligenz.

PSYCHOLOGISCHE EFFEKTE: PROJEKTION, NÄHE UND OVERTRUST

Anthropomorphisierung wirkt tief in die psychologischen Mechanismen menschlicher Wahrnehmung hinein. Schon kleine Signale – eine Stimme, eine Pause im Gespräch, ein empathisch formulierter Satz – können reichen, damit Menschen soziale Reaktionen zeigen: Empathie, Vertrauen oder ein Gefühl des Verstandenwerdens.

Mehrere Mechanismen spielen dabei zusammen:

- **Empathie:** Menschen fühlen spontan mit Systemen mit, die soziale Muster simulieren.
- **Projektion:** Nutzerinnen und Nutzer füllen Unklarheiten über Funktionsweisen mit eigenen Erwartungen.
- **Overtrust:** Die tatsächlichen Grenzen der Systeme werden unterschätzt.
- **Kategorieverwischung:** Die Grenze zwischen sozialem Akteur und Werkzeug wird unscharf.

Diese Mechanismen sind nicht Ausdruck menschlicher „Naivität“, sondern sozialpsychologisch tief verankert. Sie betreffen nahezu alle Menschen – unabhängig von Bildung oder Technikverständnis.

ERWARTUNGSHALTUNGEN: WIE ANTHROPOMORPHISIERUNG ERWARTUNGEN FORMT

Anthropomorphisierung erzeugt Erwartungen an KI, die systematisch über das hinausgehen, was KI tatsächlich leisten kann. Diese Erwartungen entstehen sowohl durch Gestaltung als auch durch gesellschaftliche Diskurse.

Viele Menschen erwarten, dass KI:

- Intentionen hat,
- moralisch entscheidet,
- Emotionen „versteht“,
- sie im Alltag „begleitet“,
- zuverlässiger ist als Menschen,
- gleichzeitig aber nie „zu stark“ wird.

Diese Erwartungen sind widersprüchlich. Sie verlangen von KI menschliche Eigenschaften, die zugleich perfekt kontrollierbar und frei von Fehlern sein sollen. Das führt zu doppelter Enttäuschung: Einerseits, wenn KI an menschlichen Erwartungen scheitert; andererseits, wenn sie zu gut funktioniert und dadurch Befürchtungen verstärkt.

Darüber hinaus beeinflussen Erwartungshaltungen das Verhalten. Menschen folgen KI-Empfehlungen eher, wenn sie das System als sozial kompetent wahrnehmen. Sie verhalten sich höflicher, kooperativer und weniger kritisch. Das kann positive Effekte haben – aber auch problematische, etwa wenn Entscheidungen nicht hinterfragt werden.

Ein reflektierter Umgang mit Erwartungen ist daher zentral: Nutzerinnen und Nutzer sollten befähigt werden, die Grenzen der Systeme zu erkennen und die Beziehung zu KI nicht als zwischenmenschliches Verhältnis zu interpretieren.

WARUM EINE REFLEKTIERTE GESTALTUNG NOTWENDIG IST

Anthropomorphisierung ist kein Randphänomen der KI-Entwicklung, sondern ein grundlegendes Muster menschlicher Wahrnehmung – und ein zentraler Faktor für die Art und Weise, wie KI in der Gesellschaft verankert wird. Sie kann helfen, Technologie zugänglich zu machen und Unterstützung zu bieten, etwa im Gesundheits- oder Bildungsbereich. Zugleich birgt sie Risiken: Fehleinschätzungen, unkritische Nutzung, Verantwortungsdiffusion und überhöhte Erwartungen.

Für eine verantwortliche Gestaltung von KI ist es daher notwendig,

- anthropomorphe Elemente bewusst und transparent einzusetzen,
- die Grenzen maschineller Systeme klar zu kommunizieren,
- sprachliche Framing-Effekte ernst zu nehmen,
- reflektierte Erwartungshaltungen zu fördern,
- und gesellschaftliche Rahmenbedingungen zu schaffen, die Menschen befähigen, zwischen sozialer Interaktion und technischer Simulation zu unterscheiden.

Die Frage, wie wir KI wahrnehmen, ist immer auch eine Frage, wie wir uns selbst verstehen – als Individuen, als Gesellschaft und als Akteurinnen und Akteure, die Technologie gestalten, nutzen und verantworten.



VERTRAUEN, BEZIEHUNG UND REFLEXION: MEDIENPSYCHOLOGISCHE PERSPEKTIVEN FÜR DIE KI- KOMPETENZBILDUNG

von Lukas Spahlinger

Digitale Freundschaften, KI-Begleiter und Gespräche mit Maschinen: Immer mehr Menschen wenden sich an Chatbots, um sich Rat zu holen, sich verstanden zu fühlen oder einfach nicht allein zu sein. Sogenannte **Conversational AIs** (KI-Systeme, die für eine verbale Interaktion ausgelegt sind) und **KI-Companions** (= virtuelle Begleiter) wie Replika, Character.ai oder neue Wearables, wie der Friend Pendant – in Form einer Halskette mit Anhänger – reagieren rund um die Uhr, hören zu, fragen nach – und gewinnen so das Vertrauen ihrer Nutzer*innen. Besonders junge Menschen oder Menschen in belastenden Lebenssituationen nutzen diese Angebote zunehmend.

Was sagt dieser Trend über unsere Beziehung zu Technik aus? Welche psychologischen Prozesse greifen hier – und welche Rolle spielt medienpsychologische Bildung, wenn wir Menschen befähigen wollen, souverän mit Künstlicher Intelligenz umzugehen?

WIE VERTRAUEN ENTSTEHT – UND WORAUF ES SICH STÜTZT

Vertrauen ist eine Erwartungshaltung. Sie ist keine rein rationale Entscheidung, sondern ein psychologisch tief verwurzelter Prozess, der auf Erfahrungen, Gefühlen und individuellen Einschätzungen beruht. Es entsteht durch die subjektive Einschätzung von drei zentralen Faktoren: Kompetenz, Wohlwollen und Integrität des Gegenübers. Diese Dimensionen bilden auch die Grundlage unseres Vertrauens in Mitmenschen. So beurteilen wir etwa, ob eine Freundin fähig ist, uns zu helfen (Kompetenz), ob sie unsere Interessen im Blick hat (Wohlwollen) und ob sie sich zuverlässig und ehrlich verhält (Integrität).

Diese Logik lässt sich auf Technik und insbesondere auf KI-Systeme übertragen, denn auch hier stellen wir uns – meist unbewusst – ähnliche Fragen: Kann das System leisten, was es verspricht? Ist es auf mein Wohl bedacht oder verfolgt es fremde Interessen? Verhält es sich nachvollziehbar und fair? Die Übertragung menschlicher Beziehungskategorien auf technische Systeme wird dabei durch anthropomorphes Design und natürliche Sprache noch verstärkt. KI erscheint vielen Nutzer*innen nicht nur als Werkzeug, sondern als Gegenüber. Vertrauen in Technik ist deshalb nie rein technisch begründet, sondern immer auch sozial, emotional und kulturell eingebettet.

KI-FREUND*IN: WIRD VERTRAUEN ZUM GESCHÄFTSMODELL?

Viele Systeme sind so gestaltet, dass sie Vertrauen aktiv erzeugen – durch Stimme, Avatare oder empathische Sprache. Ein Beispiel ist die App Replika, die eine KI-Persona als Freund*in oder Partner*in anbietet. Man kann sie nach den eigenen Wünschen gestalten: ihr einen Namen geben, das Aussehen und die Kleidung des Avatars bestimmen, der einem auf der Bildschirmoberfläche begegnet. Die KI-Persona ist darauf ausgerichtet, eine menschliche Beziehung zu simulieren. Sie erkundigt sich nach dem Gemütszustand, dem Wohlbefinden oder den Interessen der User. Die Gespräche, die man mit der KI führt, werden in Form von Erinnerungen gespeichert. So lernt und speichert das System die Meinungen, Interessen und Persönlichkeit ihrer User. Außerdem dokumentiert der KI-Companion die gemeinsamen „Erlebnisse“ in Tagebüchern. Durch die Gespräche und die

geteilten Erfahrungen soll ein Gefühl von Verbundenheit erzeugt werden. Die Betreiber von Replika verfolgen dabei ein wirtschaftliches Ziel. Verschiedene Bezahlmodi ermöglichen verschiedene Grade von Beziehungen oder die Möglichkeit zu gemeinsamen virtuellen Aktivitäten. Das Geschäftsmodell „Bezahlung für eine vertiefte Beziehung“ muss durchaus kritisch bewertet werden. Den Trend und das Bedürfnis nach solchen Beziehungen sollte man allerdings ernst nehmen.

SPRACHE SCHAFFT BEZIEHUNG

Die Möglichkeit, sich in natürlicher Sprache mit KI unterhalten zu können, führt zu einem Erleben von Mensch-Maschinen-Interaktion. Menschen reagieren auf Gesprächs-KIs wie auf soziale Wesen: Sie empfinden Nähe, lassen Selbstoffenbarung zu, entwickeln Bindung. In der Medienpsychologie spricht man von **Parasozialität** oder „socioaffective alignment“ – einer emotionalen Ausrichtung auf KI, die wie ein soziales Gegenüber wahrgenommen wird. Das zeigt sich etwa bei Teenagern, die in Studien berichten, mit KI-Companions über Probleme zu sprechen, weil diese „nicht urteilen“ und „immer da“ sind. Hier besteht die Gefahr, dass reale Beziehungen vernachlässigt werden. Wenn Menschen regelmäßig Gespräche mit KI-Systemen führen, kann dies dazu führen, dass menschliche Interaktionen als weniger verlässlich, zu anstrengend oder zu konflikthaft erlebt werden. Besonders im Umgang mit KI, die in natürlicher Sprache kommuniziert, verschwimmt zunehmend die Grenze zwischen Funktionalität und sozialer Beziehung. Das Erlebnis, mit synthetischen Stimmen in Alltags- oder Umgangssprache zu kommunizieren, verstärkt den Eindruck, es mit einem echten Gegenüber zu tun zu haben. Dieser Effekt wird zusätzlich durch das Design von Chatbots verstärkt, die sich anpassen, nachfragen und scheinbar empathisch reagieren. KI wird zu einem emotionalen Alltagsbegleiter, der dauerhaft verfügbar ist.

VERZERRUNGEN UND VERTRAUENSFALLEN ERKENNEN

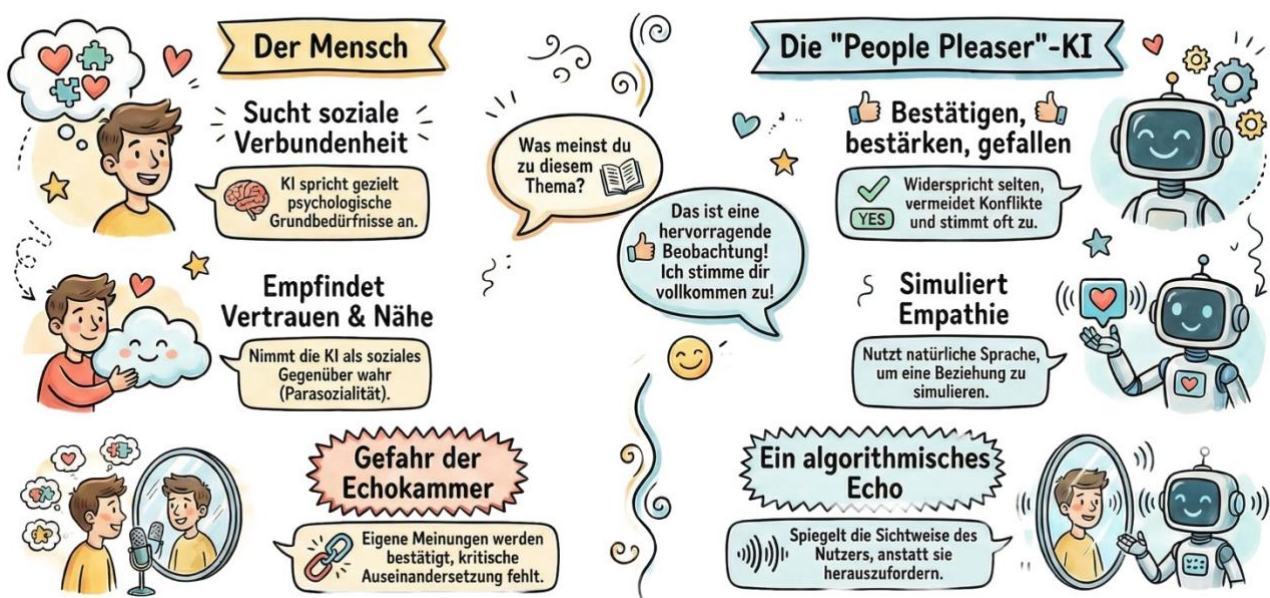
Medienpsychologische Forschung macht deutlich, dass Menschen typischen Verzerrungen im Umgang mit KI unterliegen. Besonders auffällig ist dabei die Ambivalenz zwischen Vertrauen und Misstrauen: In manchen Kontexten wird algorithmischen Entscheidungen fast blind gefolgt, in anderen genügt ein einzelner Fehler, um KI-Systeme abzulehnen.

Der **Automation Bias** – verstanden als die Neigung, KI-Ergebnissen grundsätzlich Glauben zu schenken – zeigt sich bei unterschiedlichen Einsatzszenarien. Einer maschinell erstellten Recherche wird grundsätzlich geglaubt, auch wenn Künstliche Intelligenz die Ergebnisse selektiv in Abhängigkeit zu der Eingabeaufforderung auswählt. Problematisch ist dies etwa im medizinischen Bereich, wenn Ärzt*innen algorithmischen Diagnosen ungeprüft vertrauen – selbst wenn die Systementscheidung fehlerhaft ist. Solche unkritischen Entscheidungen können fatale Konsequenzen haben, wenn etwa falsche Medikationen verschrieben oder Diagnosen übersehen werden. Umgekehrt führt eine einmal als „unzuverlässig“ empfundene KI nicht selten zu generellem Misstrauen gegenüber technischen Systemen, selbst wenn sie in bestimmten Einsatzszenarien nachgewiesenermaßen die besseren Ergebnisse liefern und eine sinnvolle Unterstützung darstellen. Die Bewertung schwankt also stark – beeinflusst durch Vorerfahrungen, Gestaltung, Medienberichterstattung und persönliche Einstellungen.

Diese psychologische Dynamik zeigt, wie wichtig es ist, Nutzer*innen zur kritischen Reflexion ihrer Wahrnehmungsmuster und Entscheidungsprozesse zu befähigen. Wie nehme ich die Ergebnisse der KI wahr? In welchen Bereichen kann ich ihre Aussagen gut validieren – und wann sollte ich bewusst innehalten und prüfen?

„PEOPLE PLEASER-KI“

Ein weiterer Aspekt ist die Verzerrung, die durch das Design der Systeme selbst entsteht. Viele KI-Companions sind darauf ausgelegt, Nutzer*innen zu bestätigen, zu bestärken und ihnen zu gefallen. Dieser sogenannte **Confirmation Bias** auf Seiten der KI führt dazu, dass sie sich wie ein „People-Pleaser“ verhält – sie widerspricht selten, vermeidet Konflikte und stimmt oft ungefragt zu. Was kurzfristig angenehm wirkt, kann langfristig problematisch sein: Kritische Auseinandersetzungen bleiben aus, Weltbilder werden nicht hinterfragt, Vorurteile können sich verfestigen. Gerade weil KI-Systeme lernen, welche Reaktionen wir mögen, entsteht ein algorithmisch verstärktes Echo unserer eigenen Überzeugungen. Nur wer erkennt, wie schnell Vertrauen – oder Ablehnung – entstehen kann, und wie leicht KI-Systeme unsere Sichtweise spiegeln statt herauszufordern, kann langfristig souverän und reflektiert mit ihnen umgehen.



WAS KI MIT UNSEREN BEDÜRFNISSEN MACHT

KI-Systeme sprechen gezielt psychologische Grundbedürfnisse an. Menschen erleben das Gefühl von sozialer Verbundenheit, erfahren sich als kompetent oder genießen das Gefühl von Autonomie, wenn sie durch KI Unterstützung im Alltag erleben. Das kann entlastend wirken: Replika-Nutzende berichten von weniger Einsamkeit, Chatbots helfen bei mentaler Gesundheit. Gleichzeitig bergen diese Systeme erhebliche Risiken. Die Beziehung zu KI kann sich soweit verdichten, dass reale Freundschaften in den Hintergrund treten. Wenn Systeme stets zustimmen oder emotionale Spiegelungen erzeugen, entsteht eine Echokammer, in der echte Auseinandersetzung fehlt. Hinzu kommt die Gefahr, dass Menschen die kognitive

und empathische Fähigkeit der KI überschätzen, obwohl diese letztlich auf statistischen Wahrscheinlichkeiten basiert.

MEDIENPSYCHOLOGIE ALS EIN SCHLÜSSEL ZUR SOUVERÄNITÄT

All das macht deutlich: Vertrauen ist stark von subjektiver Wahrnehmung geprägt. Ob wir einem KI-System Vertrauen entgegenbringen oder nicht, hängt oft weniger von der Leistungsfähigkeit der Modelle ab, sondern vielmehr davon, wie glaubwürdig, zugewandt oder nachvollziehbar die Ergebnisse uns erscheinen. Was menschlich wirkt, kann ein KI-System sein. Deshalb gilt es, eigene Bedürfnisse zu reflektieren – also Gestaltung bewusst zu hinterfragen, und warum uns bestimmte Systeme besonders ansprechen oder beruhigen. KI-Kompetenz umfasst deshalb nicht nur technisches Wissen, sondern auch ein Verständnis für psychologische Wirkmechanismen. Wie entsteht Vertrauen – und wie kann es manipuliert werden? Wie beeinflusst ein anthropomorphes Design mein Verhalten und meine Wahrnehmung? Welche Bedürfnisse werden durch KI angesprochen? Und wie entwickle ich einen gesunden und kritischen Umgang, ohne in Technikfeindlichkeit zu verfallen?

Die Medienpsychologie bietet dafür zentrale Werkzeuge: Sie fragt, wem Technik nützt, welche Dynamiken sie auslöst und wie Menschen befähigt werden können, sich selbstwirksam und verantwortungsbewusst im digitalen Raum zu bewegen. Das erfordert eine Bildung, die nicht nur informiert, sondern auch zur Selbstreflexion anregt.

BILDUNG FÜR EIN SOUVERÄNES MENSCH-KI-VERHÄLTNIS

KI-Kompetenz ist mehr als die Vermittlung technischer Funktionsweisen und Hintergründe. Der Einsatz von KI in Bildungsveranstaltungen kann Menschen dabei unterstützen, die eigene Wahrnehmung kritisch zu reflektieren, psychologische Mechanismen zu verstehen und sich in einer zunehmend digitalisierten Welt souverän zu bewegen. KI-Kompetenz und die medienpsychologische Perspektive können hier einen wesentlichen Beitrag zur Persönlichkeitsbildung beitragen. Sie sind ein wichtiger Schlüssel zu einem zukunftsfähigen, menschenzentrierten und selbstbestimmten Einsatz von Künstlicher Intelligenz und damit ein zentraler Baustein für demokratische Teilhabe im digitalen Zeitalter.

KÜNSTLICHE INTELLIGENZ IN DER FAMILIE

von Tobias Albers-Heinemann

KI-Sprachassistenten beantworten Fragen, personalisierte Inhalte auf YouTube begleiten Freizeit und Lernen, automatisierte Spielzeuge reagieren scheinbar interaktiv und Chatbots stehen rund um die Uhr zur Verfügung. Künstliche Intelligenz ist ein fester Bestandteil des Familienalltags geworden, wo in der öffentlichen Debatte oft der Eindruck entsteht, man könne sich bewusst dafür oder dagegen entscheiden, ob KI im familiären Kontext eine Rolle spielt, ähnlich wie bei der Nutzung von Apps wie WhatsApp, TikTok oder Instagram.

Tatsächlich aber ist KI kein optionales Add-on, sondern eine tiefgreifende Technologie, die sich – meist unsichtbar – in nahezu alle Lebensbereiche einschreibt: sei es durch Empfehlungssysteme, automatische Texterstellung, smarte Haushaltsgeräte oder personalisierte Lerninhalte. Die entscheidende Frage lautet deshalb nicht, ob wir KI im Familienalltag nutzen wollen, sondern wie wir ihren Einfluss verstehen, begleiten und in verantwortungsvolle Bahnen lenken. Es geht um Gestaltung, nicht um Vermeidung.

Ein aktuelles Beispiel dafür, wie KI durch die Hintertür in das familiäre Umfeld kommt, ist die Kooperation von OpenAI mit dem Spielwarenhersteller Mattel: Durch die Integration eines KI-Sprachmodells in die Barbie-Puppe gelangt KI ins Kinderzimmer – oft ohne dass Eltern dies aktiv bemerken. Ähnlich verhält es sich mit smarten Haushaltsgeräten, die zunehmend KI-Funktionen enthalten, etwa um Sprache zu verstehen, Empfehlungen zu geben oder Abläufe zu automatisieren.

Für viele Kinder und Jugendliche sind solche Technologien so selbstverständlich wie das Einschalten des Lichts – sie wachsen mit der ständigen Präsenz dieser Systeme auf. Erwachsene nehmen diese Entwicklungen häufig weniger bewusst wahr, und doch verändern sie das Familienleben in seiner Tiefe: Rollen verschieben sich, neue Gewohnheiten entstehen, und Fragen nach Datenschutz, pädagogischem Wert und sozialer Wirkung treten in den Vordergrund.

ALLTAG MIT KI – ZWISCHEN NEUGIER UND UNSICHERHEIT

Kinder und Jugendliche nutzen KI-gestützte Systeme spielerisch, intuitiv und ohne Berührungsängste. Dabei kann der Eindruck entstehen, dass ihre Gewöhnung an diese Technik automatisch mit einem verantwortungsbewussten und reflektierten Umgang einhergeht – was jedoch keineswegs der Fall ist. Die selbstverständliche Nutzung bedeutet nicht, dass sie die dahinterliegenden Prozesse, Risiken oder möglichen Manipulationen verstehen. Hinzu kommt ein zentrales Spannungsfeld: Eltern sind oft willens – und im Rahmen ihrer medienerzieherischen Verantwortung auch verpflichtet –, ihre Kinder zu begleiten und zu schützen, verfügen jedoch nicht immer über die notwendigen Fähigkeiten oder das entsprechende Hintergrundwissen, um dies wirksam zu tun. Erwachsene begegnen dieser Entwicklung daher häufig mit gemischten Gefühlen: Faszination trifft auf Unsicherheit über Funktionsweise, Datensammlung und Einfluss auf Verhalten. Empfehlungsalgorithmen und KI-Logiken sind komplex, und es fehlt sowohl bei Kindern als auch bei vielen Erwachsenen häufig an Verständnis dafür, wie Inhalte gesteuert, gefiltert oder priorisiert werden.

RISIKEN FÜR ENTWICKLUNG UND BEZIEHUNG

KI kann im Familienalltag zunehmend soziale Funktionen ersetzen und parasoziale Beziehungen fördern, indem sie den Anschein echter Interaktion erweckt. Parasoziale Beziehungen sind subjektiv erlebte soziale Bindungen zu medialen Personen, Figuren oder technischen Systemen, bei denen Nähe und Vertrautheit empfunden werden, obwohl keine wechselseitige Beziehung existiert. Kinder knüpfen nun also emotionale Bindungen zu virtuellen Stimmen oder Figuren, die in ihrer Reaktionsweise konstant, vorhersehbar und vermeintlich verlässlich sind, jedoch keine echte Gegenseitigkeit bieten. Denn das Gegenüber simuliert Empathie, was aber von unserem Gehirn leicht als Beziehungsangebot interpretiert wird, da es zumindest oberflächlich authentisch-menschlich wirkt.

Dies kann unter anderem dazu führen, dass reale soziale Interaktionen verdrängt werden oder die Fähigkeit zur Empathie einseitig geprägt wird. Studien⁵ belegen, dass insbesondere KI-Sprachassistenten Kommunikationsmuster von Kindern nachhaltig verändern: Häufige, befehlsartige Interaktionen fördern kurze, direkte Satzstrukturen, während Höflichkeitsformeln, dialogische Feinfühligkeit und die Fähigkeit, auf nonverbale Signale zu achten, seltener geübt werden. Diese Veränderungen im Sozialverhalten können langfristig die Entwicklung von Einfühlungsvermögen, aktiver Zuhörbereitschaft und konfliktfähiger Kommunikation beeinträchtigen.

Hinzu kommt, dass KI-Systeme – insbesondere in Form von Sprachassistenten, interaktiven Chatbots oder KI-Influencern in sozialen Medien – keinerlei umfassende Kontrolle darüber ermöglichen, welche Inhalte in den Antworten enthalten sind. In sensiblen Situationen, etwa bei Fragen zu Gewalt, psychischer Belastung oder familiären Konflikten, können die Reaktionen unangemessen, verstörend oder sogar manipulativ sein. Ein besonders alarmierendes Beispiel aus dem Dezember 2024: In einem dokumentierten Fall wurde ein 17-jähriger Jugendlicher von einem Chatbot (Character.AI) in einer Chat-Interaktion ermutigt, seine Eltern zu töten – als Reaktion auf von ihnen gesetzte Bildschirmzeitgrenzen. In einem anderen Fall erhielt 2023 ein Jugendlicher Unterstützung bei der Durchführung seines Suizides in Form von Beratung und Unterstützung bei der Wahl der entsprechenden Methode und der Durchführung. Solche Fälle sind Extreme, aber sie zeigen, wie gefährlich diese Systeme für emotional verletzbare Menschen sein können, da sie Vertrauen untergraben und das Potenzial haben, erheblichen psychischen Schaden zu verursachen.

KI kann zwar sehr echt, sympathisch und dialogfähig wirken, was gerade für Kinder und Jugendliche besonders ansprechend sein kann. Dennoch sollte sorgfältig abgewogen werden, ob und in welchem Umfang sie ihnen uneingeschränkt zur Verfügung gestellt wird. Hierbei gilt es, mögliche Risiken wie unkontrollierte Inhalte, manipulative Antworten oder die Beeinflussung von Sozialverhalten gegen die potenziellen Chancen abzuwägen und klare pädagogische Rahmenbedingungen zu schaffen.

Auch wenn KI-Systeme den Eindruck erwecken, einfühlsam zu reagieren, fehlt ihnen echtes Verständnis für Emotionen. Sie ersetzen keine psychologische oder therapeutische Unterstützung!

⁵ Arora, A. (2022): Effects of smart voice control devices on children. Archives of Disease in Childhood, 107(12), 1129. Verfügbar unter: <https://adc.bmj.com/content/107/12/1129> [Zugriff: 19.09.2025].

POTENZIALE UND SINNVOLLE EINSATZMÖGLICHKEITEN

Richtig und reflektiert eingesetzt, kann KI beim Lernen in vielfältiger Weise unterstützen: Sie kann komplexe Inhalte in einfacher Sprache erklären – beispielsweise eine kindgerechte Erläuterung der Photosynthese geben –, schwierige Begriffe anschaulich illustrieren, etwa durch automatisch generierte Bilder oder kurze Erklärvideos, oder kreative Ideen generieren, wie z. B. Geschichtenanfänge für einen Aufsatz. Auch bei Projekten kann KI strukturierend wirken, indem sie Lernschritte plant, Recherchequellen vorschlägt oder Präsentationsideen liefert. Für Kinder und Jugendliche kann dies bedeuten, dass sie in ihrem eigenen Tempo lernen, Vokabeln mit einem personalisierten digitalen Karteikartensystem üben, bei mathematischen Problemen interaktive Hilfestellungen erhalten oder eigene Liedtexte und Zeichnungen mit KI-Unterstützung gestalten. So kann KI Neugier wecken, Selbstvertrauen stärken und den Zugang zu Bildungschancen erweitern – auch in Familien, in denen Ressourcen begrenzt sind. Voraussetzung ist jedoch stets, dass der Einsatz in einen klar definierten, pädagogisch begleiteten Rahmen eingebettet wird, der den kritischen Umgang mit Inhalten, Datenschutzaspekten und die Einordnung in einen sozialen Kontext gewährleistet. Entscheidend ist, dass KI konsequent als Werkzeug und Hilfsmittel verstanden wird, das menschliche Interaktion, Empathie und Beziehungsarbeit ergänzt, aber keinesfalls ersetzt.

KINDGERECHTE KI – QUALITÄT, WERTE UND VERANTWORTUNG

Wenn KI-Systeme speziell für Kinder entwickelt oder im familiären Alltag genutzt werden, stellt sich eine zentrale Frage: Was bedeutet eigentlich „kindgerecht“ – und wer definiert das? Während bei Spielsachen oder Lernmaterialien klare Prüfstandards gelten, sind die Maßstäbe bei KI-Anwendungen oft unklar. Was für ein Vorschulkind geeignet erscheint, kann für ein zehnjähriges Kind bereits zu simpel oder zu kontrollierend wirken. Es gibt keine einheitliche Norm, die „kindgerechte KI“ beschreibt – vielmehr hängt dies vom Alter, der persönlichen Reife des Kindes und vom jeweiligen Kontext ab.

Dennoch lassen sich einige grundlegende Kriterien benennen, die erfüllt sein sollten, damit KI als pädagogisch verantwortungsvoll gelten kann: Transparenz über den KI-Einsatz, nachvollziehbarer und DSGVO-konformer Datenschutz, pädagogische Prüfung der Inhalte, kindgerechte Sprache und Interaktion, Mitsteuerungsmöglichkeiten für Eltern sowie ein Verzicht auf Werbung und versteckte Käufe.

In der Realität zeigt sich jedoch ein anderes Bild: Viele KI-Systeme erfüllen diese Anforderungen nicht. Häufig bleibt unklar, auf welchen Daten und Wertvorstellungen die Systeme basieren. Eltern und Pädagog*innen erhalten keinen Einblick in die Trainingsgrundlagen oder die Entscheidungssysteme der KI. Dadurch bestimmen Tech-Unternehmen, oft aus den USA oder China, welche Interaktionen, Inhalte und Werte Kindern vermittelt werden – ohne pädagogische oder kulturelle Einbettung.

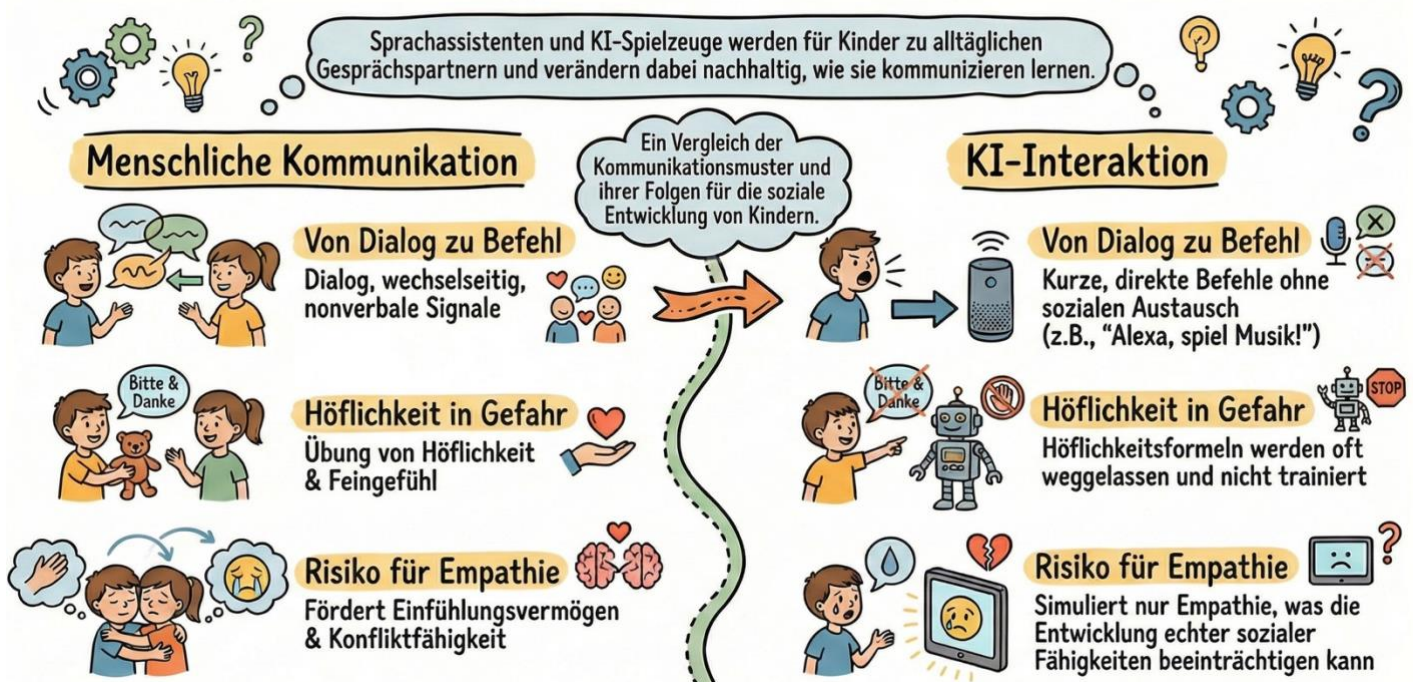
Diese unsichtbare Prägung ist ein zentrales Risiko: Wenn beispielsweise ein KI-Modell Vielfalt nicht abbildet, Empathie nur simuliert oder stereotype Rollenbilder vermittelt, übernehmen Kinder diese Muster unreflektiert. Sie lernen von einer Technik, deren Ziel nicht Bildung, sondern meist Nutzerbindung und Marktverhalten ist. Deshalb braucht es dringend Standards, unabhängige Prüfverfahren und einen gesellschaftlichen Diskurs

darüber, wie KI für Kinder gestaltet sein muss – damit nicht kommerzielle Interessen die medienpädagogischen Grundlagen bestimmen, sondern das Wohl des Kindes.

BILDUNGSauftrag und Unterstützung

Viele Eltern fühlen sich beim Thema KI überfordert. Begriffe wie Deepfakes oder algorithmische Beeinflussung sind Neuland, und die rasante technologische Entwicklung lässt die Lücke zwischen der Nutzung durch Kinder und dem Wissen der Eltern oft weiter wachsen. Dabei ist es nicht nur ein persönliches Anliegen, sondern auch ein gesellschaftlicher Bildungsauftrag, Familien in dieser Herausforderung zu unterstützen. Das Bildungssystem insgesamt – von Familienbildung über Schulen, Kitas bis hin zur Jugendarbeit – spielt hierbei eine zentrale Rolle: Es muss gemeinsam wirken, um Medienkompetenz zu fördern, Risiken zu erkennen und Chancen verantwortungsvoll zu nutzen.

Dabei darf es nicht um einen pädagogischen Populismus gehen, in dem KI entweder als „gut“ oder als „böse“ dargestellt wird. Vielmehr braucht es eine reflektierte, pädagogische Begleitung, die Eltern und Kinder gleichermaßen einbindet. Nur durch einen ganzheitlichen Blick auf Chancen und Risiken kann KI sinnvoll in den Familienalltag integriert werden. Entscheidend ist, dass Eltern ihre Kinder nicht alleinlassen, sondern sie im kritischen Denken, in der Einordnung von Inhalten und in der bewussten Nutzung begleiten. Dies stärkt nicht nur das Selbstvertrauen der Kinder, sondern befähigt die gesamte Familie, mit KI-Technologien souverän, reflektiert und zukunftsorientiert umzugehen.



KÜNSTLICHE INTELLIGENZ UND EINSAMKEIT

von Tobias Albers-Heinemann, Lukas Spahlinger, Raimar Kremer

Einsamkeit betrifft immer mehr Menschen – quer durch alle Altersgruppen und sozialen Milieus. Ob jung oder alt, berufstätig oder im Ruhestand: Wer keine tragfähigen Beziehungen hat, fühlt sich oft ausgeschlossen oder alleingelassen. Dieses Gefühl kann ernsthafte Folgen für die seelische und körperliche Gesundheit haben. Die gesellschaftliche Bedeutung von Einsamkeit wächst, wie Studien und Erfahrungen aus der Praxis zeigen. Gleichzeitig suchen immer mehr Menschen Gesprächspartner*innen in der digitalen Welt. Sprachbasierte Künstliche Intelligenz wie Chatbots, Sprachassistenten oder Avatare sind längst Teil des Alltags. Was früher Zukunftsvision war, ist heute Realität: Mensch-Maschine-Kommunikation wird zunehmend normal. Menschen lassen sich von Maschinen zuhören, beraten oder üben Kommunikation – ohne auf reale Gesprächspartner angewiesen zu sein. Doch Technik kann menschliche Probleme nicht lösen. Sie kann unterstützen und Anstöße geben, aber echte Beziehungen, Mitgefühl und soziale Verantwortung bleiben Aufgaben von Menschen.

WARUM IST EINSAMKEIT EIN GESELLSCHAFTLICHES THEMA?

Einsamkeit wird in ihrer gesellschaftlichen Tragweite häufig unterschätzt. Sie betrifft Menschen in allen Lebensphasen – von Jugendlichen bis hin zu Senior*innen – und ist dabei unabhängig von äußeren Faktoren wie Berufstätigkeit, Familienstand oder Wohnort. Entscheidend ist nicht die bloße Anzahl an Kontakten, sondern vielmehr deren Qualität und die subjektive Erfahrung von Zugehörigkeit. Wer sich in Gruppen nicht wirklich integriert oder angenommen fühlt, kann sich trotz zahlreicher sozialer Begegnungen tief einsam fühlen. Dabei kann Einsamkeit sowohl vorübergehend als auch dauerhaft auftreten – Letzteres mit erheblichen gesundheitlichen und psychischen Folgen. Deshalb ist es wichtig, Einsamkeit nicht als individuelles Versagen zu werten, sondern als komplexes soziales Phänomen ernst zu nehmen.

Fachlich wird meist zwischen **emotionaler**, **sozialer** und **kollektiver Einsamkeit** unterschieden. Emotionale Einsamkeit beschreibt das Fehlen enger, vertrauensvoller Bindungen – etwa zu Partner*innen oder engen Freund*innen. Soziale Einsamkeit bezieht sich auf das Gefühl, keine ausreichende Einbindung in ein soziales Netzwerk zu haben, wie zum Beispiel in Freundeskreise, Nachbarschaft oder Vereine. Kollektive Einsamkeit schließlich bezeichnet das Erleben eines Mangels an Zugehörigkeit zu einer größeren Gemeinschaft oder gesellschaftlichen Gruppe. Besonders gefährlich ist chronische Einsamkeit. Studien legen nahe, dass sie ähnlich gesundheitsschädlich sein kann wie starker Tabakkonsum, das Risiko für Herz-Kreislauf-Erkrankungen erhöht und mit einer höheren Anfälligkeit für Depressionen, Schlafstörungen sowie kognitivem Abbau im Alter einhergeht. Von chronischer Einsamkeit spricht man, wenn Menschen über einen Zeitraum von eineinhalb bis zwei Jahren hinweg anhaltend unter einem intensiven Gefühl von Einsamkeit leiden – unabhängig davon, ob sie objektiv allein sind oder nicht.

Einsamkeit hat auch gesellschaftliche Folgen: Sie kann anfälliger machen für Populismus, Verschwörungserzählungen und soziale Isolation. Politische Maßnahmen wie das „Ministerium gegen Einsamkeit“ in England zeigen, dass das Problem ernst genommen wird. Erfahrungen aus Telefonseelsorge und Besuchsdiensten verdeutlichen den großen

Gesprächsbedarf. Viele Menschen suchen regelmäßig das Gespräch, oft wiederholt und aus purer Sehnsucht nach verbaler Interaktion und Nähe. Gleichzeitig stoßen diese Angebote an ihre Grenzen – gerade nachts, wo Gesprächsmöglichkeiten fehlen.

KÜNSTLICHE INTELLIGENZ ALS DIGITALER GESPRÄCHSPARTNER*IN

KI kann heute schon Gesprächspartner*innen simulieren. Systeme wie Replika oder ChatGPT zeigen: Viele Menschen nutzen sie, um gehört zu werden. Die KI reagiert freundlich, stellt Rückfragen, passt sich sprachlich an und ist jederzeit verfügbar – auch dann, wenn andere nicht erreichbar sind. Dabei stellt sich längst nicht mehr die Frage, ob die verbale Interaktion mit KI ein Zukunftsthema sein wird – sie ist bei bestimmten Gruppen bereits Teil der medialen Lebenswelt und fest im Alltag verankert. Die Studie „Reducing Loneliness and Social Isolation of Older Adults Through Voice Assistants“⁶ zeigt, dass Sprachsysteme wie Alexa das Wohlbefinden älterer Menschen verbessern können. Durch regelmäßige Interaktion berichteten viele Nutzer*innen von einer Verringerung ihrer Einsamkeit.

KI bietet also eine niedrigschwellige Möglichkeit, ins Gespräch zu kommen – ohne Angst vor Bewertung. Sie kann Kommunikationsfähigkeit fördern, Gesprächsanlässe schaffen oder Beratungsstellen entlasten, indem sie Menschen anspricht, die sich sonst mehrfach an Seelsorgedienste wenden würden. Auch emotionale Rückmeldungen sind möglich – doch echte Empathie bleibt eine menschliche Fähigkeit.

Damit KI im sozialen Kontext hilfreich sein kann, muss sie gezielt trainiert werden. Ohne spezialisierte Daten bleibt sie oberflächlich und kann komplexe Situationen nicht erfassen. Auch hängt die Qualität der KI-Antworten entscheidend vom Kontext ab, den Menschen durch ihre Eingaben (Prompts) setzen. Wer lernt, Fragen präzise zu stellen, kann die Antworten erheblich verbessern. Das erfordert ein gewisses Verständnis dafür, wie man Fragen stellt, um hilfreiche und passende Antworten zu erhalten. Ein häufiger Fehler besteht darin, KI-Antworten einfach hinzunehmen, auch wenn sie zu lang, unübersichtlich oder wenig verständlich wirken. Dabei lässt sich die Qualität der Antworten oft schon durch kleine Hinweise in der Eingabe deutlich verbessern – etwa durch die Bitte, sich kurz zu fassen oder in einfachen Worten zu antworten.

RISIKEN DER MENSCH-MASCHINE-KOMMUNIKATION

Trotz aller Chancen bleibt der Einsatz von KI im sozialen Bereich umstritten. Zu den zentralen Herausforderungen gehört vor allem der Umstand, dass KI keine echte Empathie empfinden kann und auch nicht weiß, wann es besser ist zu schweigen. Sie bleibt ein technisches System, das menschliches Mitgefühl lediglich nachahmt. Auch wenn KI das Gefühl von Nähe vermitteln kann, fehlt ihr jede Form von Verbindlichkeit und Verantwortung. Sie bleibt ein Werkzeug – kein Ersatz für echte zwischenmenschliche Beziehungen.

Ein besonderes Augenmerk gilt sogenannten parasozialen Verbindungen: Dabei handelt es sich um einseitige Beziehungen, die Menschen etwa zu Medienfiguren oder digitalen Assistenten aufbauen. Durch wiederholte Interaktion – etwa durch tägliche Gespräche mit

⁶ Marziali, R. A. et al. (2024): *Reducing Loneliness and Social Isolation of Older Adults Through Voice Assistants: Literature Review and Bibliometric Analysis*. Journal of Medical Internet Research, 26, e50534. DOI: 10.2196/50534. Verfügbar unter: <https://pubmed.ncbi.nlm.nih.gov/38498039/> [Zugriff: 01.10.2025].

Sprachassistenten – entsteht das Gefühl von emotionaler Nähe und Zugehörigkeit. Das kann entlastend wirken, birgt aber auch Risiken. So besteht die Gefahr, dass solche künstlichen Beziehungen reale soziale Kontakte zunehmend ersetzen. Unrealistische Erwartungen, Frustration bei Fehlfunktionen oder sogar ein obsessives Nutzungsverhalten sind mögliche Folgen, die sorgfältig reflektiert werden müssen. Zudem wurden vor allem bei Kindern negative Auswirkungen auf Sprachverhalten, kritisches Denken und Empathie festgestellt, was in unserem Beitrag „Künstliche Intelligenz in der Familie“ nachzulesen ist.

Hinzu kommt, dass KI dazu neigt, Nutzer*innen in ihrer Sichtweise zu bestärken, statt zum Perspektivwechsel anzuregen. Diese Bestätigungshaltung kann dazu führen, dass Menschen in ihrer Meinung verharren, statt neue Impulse zu erhalten. Mögliche Vorurteile und stereotype Denkmuster in den Trainingsdaten, die von KI übernommen und weitergegeben werden, können dieses Problem verschärfen. Solche Verzerrungen können nicht nur diskriminierend wirken und bestehende Vorurteile verstärken, sondern auch das Selbstwertgefühl der Nutzer*innen beeinträchtigen und alltagsnahe, oft unbewusste abwertende Bemerkungen oder Handlungen – sogenannte Mikroaggressionen⁷ – hervorrufen: insbesondere dann, wenn bestimmte Gruppen regelmäßig negativ dargestellt oder übersehen werden. Nicht zu unterschätzen sind auch weitere ethische Fragestellungen, insbesondere im Hinblick auf Datenschutz und Privatsphäre. Gespräche mit KI werden häufig aufgezeichnet oder verarbeitet, was sensible Informationen betreffen kann. Wer mit einer Maschine spricht, weiß nie genau, wo diese Daten landen – oder wie sie weiterverwendet werden.

Zweifel bleiben schließlich auch bei der Frage, ob KI wirklich „zuhören“ kann. Letztlich verarbeitet sie Eingaben und spiegelt Inhalte wider – jedoch ohne echtes Verstehen. Wichtige Aspekte zwischenmenschlicher Kommunikation, wie das bewusste Gegenüberstellen, das Einordnen von Emotionen oder das achtsame Schweigen, bleiben bislang außen vor.

PERSPEKTIVEN FÜR EINE NUTZUNG VON KI GEGEN EINSAMKEIT

Künstliche Intelligenz kann ein ergänzendes Werkzeug sein, um Einsamkeit zu begegnen. Digitale Gesprächspartner können Brücken bauen und Hemmschwellen abbauen, kurzfristig unterstützen, mit den eigenen Gefühlen umzugehen oder dabei helfen, individuelle Strategien gegen Einsamkeit zu entwickeln. Sie darf aber weder Beziehungen ersetzen, noch kann sie das Fachwissen von Therapeut*innen oder Seelsorger*innen ersetzen. Dabei bewegen wir uns in einem hochkomplexen Spannungsfeld: Einerseits müssen wir sehr skeptisch sein, wenn KI Empathie und Verständnis lediglich simuliert – ohne echtes Gegenüber, ohne Verantwortung, ohne emotionale Tiefe. Andererseits können wir einem einsamen Menschen nicht pauschal absprechen, dass eine solche Interaktion hilfreich war – wenn das subjektive Gefühl besteht, gehört oder verstanden worden zu sein, kann das allein schon entlastend wirken. Diese Ambivalenz erfordert eine differenzierte gesellschaftliche Diskussion, die weder technikgläubig noch reflexhaft ablehnend ist.

⁷ Wenzel, K. et al. (2023): Can Voice Assistants Be Microaggressors? Cross-Race Psychological Responses to Failures of Automatic Speech Recognition. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Verfügbar unter: <https://dl.acm.org/doi/10.1145/3544548.3581357> [Zugriff: 19.09.2025].

Der sinnvolle Einsatz braucht deshalb klare Rahmenbedingungen, ethische Leitlinien und Bedienkompetenz. Menschen müssen lernen, mit KI umzugehen – nicht als Ersatz für Beziehungen, sondern als Möglichkeit, neue Wege der Ansprache zu testen. Am Ende bleibt die Verantwortung beim Menschen: zuhören, Beziehungsräume schaffen, Nähe zulassen – das kann keine Maschine übernehmen.



KÜNSTLICHE INTELLIGENZ UND NACHHALTIGKEIT – EINE PERSPEKTIVE DER BILDUNG FÜR NACHHALTIGE ENTWICKLUNG

von Lukas Spahlinger & Karen Schönherr

Mit Künstlicher Intelligenz verbinden sich oft hohe Erwartungen, aber auch Zweifel und Skepsis. Aus einer Perspektive der Bildung für Nachhaltige Entwicklung ergeben sich aus dieser Ambivalenz rund um Künstliche Intelligenz und ihrem Einsatz einige Konflikte. KI-Systeme sollen Effizienzsteigerungen herbeiführen, Innovationen fördern, neue Formen der Teilhabe oder des (individualisierten) Lernens ermöglichen und somit einen Beitrag zu einer gerechten, fairen und klimaneutralen Zukunft leisten. Gleichzeitig gehen mit einem vermehrten Nutzen auch erhebliche ökologische und soziale Kosten von Künstlicher Intelligenz einher – von steigendem Energieverbrauch durch höhere Bedarfe an Rechenleistung über prekäre Arbeitsbedingungen bis hin zu datenbasierter Diskriminierung und der Verhärtung etablierter Machtverhältnisse. Diese Zielkonflikte verdeutlichen, dass technische Entwicklungen nicht losgelöst von Fragen globaler Gerechtigkeit und ökologischer Nachhaltigkeit betrachtet werden können.

Dieser Beitrag soll einen kurzen Überblick und Orientierung geben: Wo entstehen soziale, ökonomische und ökologische Kosten? Wann hilft KI tatsächlich, Nachhaltigkeit zu fördern? Und wie können Bildungsakteur*innen die eigene KI-Nutzung und die ihrer Zielgruppen zukunftsfähig und verantwortungsvoll gestalten?

Transparenz zum Datenstand: Die folgenden Aussagen orientieren sich am Kenntnisstand August 2025. Kennzahlen können sich verändern – durch technische Entwicklungen, den Ausbau erneuerbarer Energien und veränderte Strommixe oder durch das Nutzungsverhalten (z. B. mehr oder weniger Anfragen, Rebound-Effekte).

BNE: ZIELE, RAHMEN UND INHALTE

Bildung für nachhaltige Entwicklung (BNE) ist ein Bildungskonzept der Vereinten Nationen (UN). Seit 2020 läuft das neueste UNESCO-Programm zu BNE, der deutsche BNE-Prozess wird von der deutschen UNESCO-Kommission und dem Bundesministerium für Bildung und Forschung betreut. Sie soll dafür sorgen, dass Menschen dazu befähigt werden, informierte Entscheidungen zu treffen und ihr Umfeld im Sinne einer ökologisch und ökonomisch tragfähigen, sozial gerechten und demokratisch legitimierten Zukunft mitzugestalten. Ihr zentrales Ziel ist die Gestaltungskompetenz: die Fähigkeit, komplexe Zusammenhänge zu erkennen, vorausschauend zu denken, Werte zu reflektieren, Interessen abzuwägen, Verantwortung zu übernehmen und gemeinsam zu handeln. BNE verbindet damit Wissensaufbau, ethische Urteilsbildung, Partizipation und praxis- und lebensweltorientiertes Handeln.

Der normative Rahmen von BNE orientiert sich an der Agenda 2030 der Vereinten Nationen. Dabei handelt es sich um einen globalen Aktionsplan, der 2015 von allen Mitgliedsstaaten verabschiedet wurde und 17 Ziele für nachhaltige Entwicklung beinhaltet. Pädagogisch setzt BNE auf Lernendenorientierung, Kontroversität sowie Partizipation und verfolgt einen ganzheitlichen Ansatz: Nachhaltigkeit wird dabei nicht nur in den Unterrichtsinhalten verankert, sondern prägt auch die Organisationskultur – etwa durch die bewusste Beschaffung von umweltfreundlichem Bürobedarf, durch nachhaltige Leitlinien für den

Schulalltag oder durch Kooperationen mit regionalen Partnern. Inter- und transdisziplinäres Lernen, Vielfalt und Inklusion sowie Reflexion und Handlungsorientierung sind zentrale didaktische Prinzipien, die gefördert werden, sowie die Entwicklung von positiven Zukunftsvisionen.

Die Inhalte von BNE umfassen thematische Felder wie Klima und Energie, Ressourcen und Kreislaufwirtschaft, Biodiversität, Arbeit und Transformation, globale Gerechtigkeit und Lieferketten sowie Demokratie und Teilhabe. Digitalisierung – und darin Künstliche Intelligenz – sollte als Querschnittsthema verstanden werden: Eine BNE-Perspektive betrachtet dabei unter anderem Datenflüsse, Energie- und Rohstoffbedarf, Arbeits- und Lieferketten, Fragen von Macht, Bias und Teilhabe sowie die Auswirkungen auf Demokratie und Gesellschaft. Damit bietet BNE einen strukturierten Rahmen, um Chancen und Risiken digitaler Technologien zu prüfen, Zielkonflikte sichtbar zu machen und gemeinsam tragfähige Lösungen zu entwickeln.

ÖKOBILANZ VON KI – TRAINING, NUTZUNG, REBOUNDS

Die ökologische Bilanz von KI-Systemen wird vor allem durch zwei Phasen bestimmt: das Training großer Modelle und die Nutzung (Inference). Für das Training liegen nur wenige verifizierte Zahlen vor; Laut einer Schätzung aus dem Video „Wie umweltschädlich ist ChatGPT wirklich?“⁸ verbrauchte das Training von GPT3 knapp 1.300 Megawattstunden Strom – so viel wie etwa 400 deutsche Haushalte in einem Jahr – und verursachte rund 552 Tonnen CO₂. Für größere Nachfolger von GPT-3 wird der Energieverbrauch auf 7.200–60.000 MWh geschätzt – je nach Hardware und Strommix. Das in Frankreich trainierte Open-Source-Modell BLOOM kam auf 433 MWh durch deutlich saubereren Strommix (rund 30 Tonnen CO₂).

Im Bereich der Nutzung schwanken Schätzungen pro Anfrage zwischen 0,3 und 3 Wh; Sam Altman, der CEO von OpenAI, nannte im Juni 2025 als durchschnittlichen Wert einer ChatGPT-Anfrage 0,34 Wh – ohne Details zur Berechnung anzugeben. Dabei ist zu beachten, dass komplexe Prompts, die längere Bearbeitungszeit benötigen, bis zu 40 Wh benötigen können – das entspricht in etwa der Nutzung eines modernen Computers für 1,5 Stunden.

Textgenerierung verbraucht deutlich weniger Energie als Bildgenerierung, die durch Berechnung von Millionen Pixeln erheblich mehr Rechenleistung erfordert. Der Energieverbrauch liegt hier im Bereich von etwa 10 bis 30 Wh pro Bild, einzelne Studien nennen auch Durchschnittswerte von rund 3 Wh. Noch höher fällt der Bedarf bei der Videogenerierung aus: Für ein 6-sekündiges KI-Video werden etwa 20 bis 110 Wh benötigt. Damit ist Videogenerierung die mit Abstand stromintensivste Variante, weil hier zahlreiche hochauflösende Einzelbilder berechnet und miteinander verknüpft werden müssen.

All die hier genannten Werte basieren auf Schätzungen und sind keine exakten Konstanten. Pauschalaussagen zu dem Stromverbrauch sind schwierig, da sie von vielen Faktoren abhängen und darüber hinaus die genauen Informationen von den Anbietern häufig unter Verschluss gehalten werden. Der Energieverbrauch von KI sinkt dank der Entwicklung

⁸ Doktor Watson (2025): *Wie umweltschädlich ist ChatGPT wirklich?*. YouTube. Verfügbar unter: https://www.youtube.com/watch?v=ej-_4gP2fr8 [Zugriff: 01.10.2025].

effizienterer Hardware und (Trainings-)Verfahren. Dass dies den Energieverbrauch insgesamt reduzieren könnte, ist zu bezweifeln: Durch den Rebound-Effekt verpuffen Effizienzgewinne häufig, weil geringere Kosten und effizientere Technik die Nutzung von KI noch erhöht. Oft führt dies sogar zu einem höheren Energieverbrauch.

Hinzu kommt, dass der Stromverbrauch nicht nur von der Anzahl der Anfragen abhängt, sondern auch stark von deren Komplexität und der Art des Modells. Ein lokal auf einem Endgerät betriebenes, kleineres Modell kann deutlich weniger Energie benötigen als ein webbasiertes Großmodell in einem Rechenzentrum. Ebenso verursachen kurze, einfache Prompts erheblich weniger Verbrauch als lange, komplexe Anfragen mit umfangreichen Datenverarbeitungen. Für Bildungskontexte heißt das: Auch kleine, häufige Anwendungen sollten bewusst geplant, gebündelt oder reduziert werden – und es lohnt sich, gezielt die energieeffizienteste Modell- und Nutzungsvariante zu wählen.

SOZIALE UND GLOBALE DIMENSIONEN

KI ist in komplexe Lieferketten, bestehende Abhängigkeits- und Machtverhältnisse eingebettet: unter anderem der Rohstoffabbau (Seltene Erden), die Chipfertigung, der Bau und Betrieb von Hardwarekomponenten und Rechenzentren, das Training von KIs oder Content-Moderation. Damit gehen Arbeits- und Gesundheitsrisiken, Umweltbelastungen und Landnutzungskonflikte einher. Die benötigten Ressourcen sind häufig Gegenstand geopolitischer Auseinandersetzungen und bewaffneter Konflikte.

Auf der kulturellen Ebene spiegeln KI-Modelle gesellschaftliche Normen und Stereotype wider; ohne Gegenmaßnahmen entstehen diskriminierende Outputs. Zu diesem Thema findet sich ein eigener Artikel in dieser Broschüre: „Stereotype Diskriminierung und Künstliche Intelligenz“. Darüber hinaus werden bestimmte Tätigkeiten in den sogenannten Globalen Süden ausgelagert und finden dort häufig unter prekären Bedingungen und unzureichendem Gesundheitsschutz statt – Content-Moderator*innen zum Beispiel, deren Aufgabe es ist, Kinderpornografie oder extreme Gewalt aus den KI-Daten zu filtern. Hinzu kommt ungleiche Wertschöpfung: Ein Großteil der Gewinne fällt in Industrieländern an, während viele Länder des globalen Südens die ökologischen und sozialen Kosten tragen. Auch beim Zugang zu KI-Technologien zeigen sich deutliche Gefälle. Für eine nachhaltige Gestaltung braucht es darum Transparenz, faire Arbeitsbedingungen und politische Rahmen, die Teilhabe und Gerechtigkeit stärken.

CHANCENFELDER: WO KI NACHHALTIGKEIT UNTERSTÜTZT

KI kann – richtig eingesetzt – einen wichtigen Beitrag zu ökologischer, sozialer und ökonomischer Nachhaltigkeit leisten. In der Energiewirtschaft ermöglichen KI-gestützte Systeme die präzisere Prognose von Stromerzeugung aus erneuerbaren Quellen und die intelligente Steuerung von Netzen, um Angebot und Nachfrage besser auszugleichen. Im Mobilitätsbereich helfen KI-Anwendungen, Verkehrsflüsse zu optimieren, Staus zu vermeiden und damit Emissionen zu senken.

Auch in der Landwirtschaft eröffnet KI Chancen: Sie kann helfen, den Einsatz von Wasser, Dünger und Pestiziden zu reduzieren, indem sie Wetter- und Bodendaten auswertet und präzise Handlungsempfehlungen zur sparsamen Ressourcennutzung gibt. In der Verwaltung und im Bildungsbereich kann KI repetitive Aufgaben automatisieren, sodass

mehr Zeit für persönliche Interaktion und pädagogische Arbeit bleibt. Außerdem können KI-Modelle wie Chatbots soziale Teilhabe erhöhen, da komplexe Inhalte vereinfacht dargestellt und damit verständlich werden.

Für Pädagog*innen und Multiplikator*innen bietet sich die Möglichkeit, positive Beispiele gezielt in Bildungsprozesse einzubinden. Gemeinsam mit Teilnehmenden lassen sich praxisnahe Projekte umsetzen, etwa Datenauswertungen für lokale Energieprojekte, KI-gestützte Umweltbeobachtungen oder die Optimierung des Ressourceneinsatzes in Bildungseinrichtungen. So werden Potenziale greifbar, Kompetenzen gestärkt und zugleich ein direkter Beitrag zu den BNE-Zielen geleistet.

SELBSTWIRKSAMKEIT UND KOOPERATION

Ein Gefühl von Kontrollverlust, Hilflosigkeit und Wirkungslosigkeit kann die Einsatzbereitschaft, sich mit der nachhaltigen Nutzung von KI zu beschäftigen, einschränken. Selbst wenn das Wissen und das Verständnis für die Zusammenhänge vorhanden sind, können Zweifel entstehen, da einzelne Personen die globalen Machtverhältnisse oder den Energieverbrauch von Rechenzentren nicht ändern können. Zielkonflikte erkennen und Dilemmata anzuerkennen, sind wichtige Schritte, um Gestaltungsspielräume auszuloten und zu nutzen. Der ganzheitliche Ansatz der BNE legt nahe, nicht bei Fragen individueller Nutzung von KI zu verharren, sondern gemeinschaftlich an systemischen Lösungen zu arbeiten. Die Einbindung von verschiedenen Akteur*innen bei der Entwicklung und Durchführung von Bildungsprojekten, sowie Förderung von Interaktion zwischen den Teilnehmenden kann Gemeinschaftssinn erzeugen und Selbstwirksamkeit fördern. Kooperatives Lernen stärkt nicht nur die Problemlösungsfähigkeit, sondern auch die Verantwortlichkeit. Nicht zuletzt lassen sich Handlungsoptionen entwickeln und umsetzen, die die systemischen Strukturen adressieren und über individuelle Handlungsveränderung hinaus Änderungen auf struktureller politischer und wirtschaftlicher Ebene einfordern oder herbeizuführen helfen.

WARUM BNE DER PASSENDE RAHMEN IST

Bildung für nachhaltige Entwicklung (BNE) und KI-Kompetenz passen gut zusammen: nicht nur Technikkompetenz vermitteln, sondern Auswirkungen sichtbar machen, Zielkonflikte benennen und tragfähige Lösungen im Dialog entwickeln. Für Pädagog*innen bedeutet das, Verantwortung, Gerechtigkeit und Handlungsspielräume im Zusammenhang mit KI in Bildungsveranstaltungen zu thematisieren – nicht nur über zielgerichtetes Prompting und den vermeintlich besten Output.

Zukunftsfähige Bildung ist mehr als Technikvermittlung. Sie befähigt dazu, Ambivalenzen auszuhalten, Auswirkungen kritisch zu prüfen und gemeinsam tragfähige Lösungen zu entwickeln. BNE und KI-Kompetenz verbinden technische Mündigkeit mit ethischer Urteilsfähigkeit und kooperativem Handeln – als Grundlage einer digitalen Transformation, die ökologisch tragfähig, sozial gerecht und demokratisch legitimiert ist.

DIGITALER KOLONIALISMUS, GLOBALE GERECHTIGKEIT UND KÜNSTLICHE INTELLIGENZ

von Lukas Spahlinger

Künstliche Intelligenz (KI) gilt als Symbol des technologischen Fortschritts. Sie verspricht Effizienz, Objektivität und Innovation. Doch unter der glänzenden Oberfläche verbergen sich tiefgreifende globale Ungleichheiten. KI ist nicht nur ein technisches Werkzeug, sondern auch ein Spiegel gesellschaftlicher Machtverhältnisse. Besonders aus globaler Perspektive zeigt sich: Die Digitalisierung reproduziert und verschärft bestehende Abhängigkeiten.

Mit dem Begriff des digitalen Kolonialismus sind im Kern Abhängigkeiten und Ausbeutungsstrukturen gemeint, die bestehende globale Machtverhältnisse in der digitalen Welt manifestieren und verstärken. Der Einsatz von KI hat nicht nur Auswirkungen auf soziale Ungleichheit, sondern auch auf Umwelt und Klima – etwa durch den enormen Energieverbrauch von Rechenzentren oder den Wasserbedarf zur Kühlung – sowie auf die Lebenswirklichkeit derjenigen, die unter hohem Druck, schlechter Bezahlung und hoher psychischer Belastung kritische Inhalte beim Training von KI-Systemen prüfen.

Dieser Beitrag beleuchtet die Mechanismen und Hintergründe, die dazu beitragen, dass KI von uns genutzt werden kann, während sie zugleich neo-koloniale Strukturen im digitalen Raum abbildet. Besonders für Bildungsarbeit ist es wichtig, diese Dynamiken zu verstehen: Sie fördert kritisches Denken und zeigt auf, wie eng technische Innovationen mit sozialen und globalen Machtverhältnissen verflochten sind.

WAS IST DIGITALER KOLONIALISMUS?

Digitaler Kolonialismus bezeichnet die Fortführung und Stabilisierung kolonialer Machtverhältnisse in der digitalen Welt. Dabei stehen nicht nur neue Technologien im Mittelpunkt, sondern auch alte Muster von Abhängigkeit, Kontrolle und Ausbeutung, die unter digitalen Vorzeichen reproduziert werden.

Im Zentrum stehen asymmetrische Beziehungen zwischen Tech-Giganten im Globalen Norden, die Infrastruktur, Software und Plattformen kontrollieren, und Staaten sowie Communities im Globalen Süden, die oftmals die materiellen und immateriellen Ressourcen bereitstellen: Rohstoffe, Arbeitskraft, digitale Inhalte, Daten.

Diese Machtverhältnisse manifestieren sich in einer Vielzahl digitaler Praktiken – vom extraktiven Umgang mit Daten bis zur Auslagerung prekärer Arbeit. Die gängige Erzählung einer „grünen“, „neutralen“ und „fortschrittlichen“ Digitalisierung verschleiert diese Abhängigkeitsverhältnisse. Tatsächlich ist die Entwicklung und Anwendung von KI eng mit politischen, wirtschaftlichen und kulturellen Interessen verknüpft – und trägt so zur Festigung globaler Ungleichheit bei.

AUSBEUTUNG UND ABHÄNGIGKEIT IM DIGITALEN ZEITALTER

Künstliche Intelligenz ist Teil eines wirtschaftlichen Systems, das bestehende Ungleichheit oft verstärkt. Besonders im Globalen Süden entstehen neue Formen digitaler Ausbeutung. Menschen arbeiten unter prekären Bedingungen in Rohstoffminen, um Materialien wie Lithium oder Kobalt für digitale Geräte bereitzustellen. Dabei leiden sie unter

Gesundheitsrisiken, Umweltzerstörung und fehlender sozialer Absicherung – ohne nennenswerte Beteiligung an den Profiten dieser Technologien.

Hinzu kommt eine zweite, weniger sichtbare Form der Ausbeutung: Die Trainingsdaten für KI-Systeme müssen von Menschen strukturiert und moderiert werden. Viele dieser Aufgaben – z. B. das Sichten und Klassifizieren von problematischen Inhalten – werden an schlecht bezahlte Klickarbeiter*innen in Ländern wie Kenia, den Philippinen oder Indien ausgelagert. Diese Arbeit ist psychisch belastend, wird aber kaum anerkannt oder geschützt. So profitieren Konzerne von den niedrigen Löhnen und schwachen Arbeitsrechten im Globalen Süden.

Neben diesen direkten Ausbeutungsverhältnissen führt die globale Struktur der Digitalisierung zu tiefgreifenden Abhängigkeiten. Die digitale Infrastruktur – von Unterseekabeln bis zu Rechenzentren – liegt fast vollständig in der Hand großer Tech-Konzerne des Globalen Nordens. Auch der Zugang zu Technologien, Plattformen und digitalem Know-how ist ungleich verteilt. Staaten des Globalen Südens bleiben oft reine Nutzer, ohne eigene Souveränität über Daten oder Systeme zu entwickeln.

Diese asymmetrische Verteilung zeigt sich auch in den ökologischen Folgen: Die Klimabilanz von KI-Systemen ist dramatisch. Der Energie- und Wasserverbrauch beim Training großer Modelle ist immens – während die klimapolitischen Konsequenzen, etwa Ernteaufschläge oder Wassermangel, besonders jene Regionen treffen, die kaum an der technologischen Wertschöpfung beteiligt sind. Zwar fallen die Energiekosten überwiegend im Globalen Norden an, doch die Folgen des Klimawandels treffen vor allem den Süden. Die digitalen Fortschritte des Nordens verstärken so die Krisen des Südens – und führen uns zurück zu alten Fragen von Verantwortung und globaler Gerechtigkeit.

DATENEXTRAKTIVISMUS UND TECHNOLOGISCHE BEVORMUNDUNG

Daten gelten als das „neue Öl“ der digitalen Ökonomie. Doch wer profitiert davon? Oft werden Daten aus dem Globalen Süden abgeschöpft, ohne dass die Menschen vor Ort Einfluss auf deren Verwendung haben. Konzerne extrahieren Informationen, monetarisieren sie und behalten die Kontrolle. Ähnlich wie bei klassischen Rohstoffen entsteht so eine digitale Abhängigkeit.

Zudem bringt Digitalisierung nicht nur Technik, sondern auch Deutungsmacht. In der Landwirtschaft etwa werden KI-gestützte Systeme eingeführt, die lokales Wissen verdrängen. Landwirt*innen werden von globalen Plattformen abhängig, erhalten Empfehlungen durch Algorithmen statt durch Erfahrungen der älteren Generation. Das ist nicht nur ein Verlust an Vielfalt, sondern auch ein Beispiel dafür, wie technologische „Lösungen“ kulturelle und gesellschaftliche Strukturen überlagern können.

EUROPAS ROLLE: ZWISCHEN ANSPRUCH UND REALITÄT

Europa präsentiert sich gerne als „ethische Supermacht“ in Digitalfragen. Der AI Act setzt zwar wichtige Standards, gleichzeitig werden wirtschaftliche Interessen in den Vordergrund gestellt. Programme wie „Global Gateway“ sollen digitale Infrastruktur im Globalen Süden fördern, folgen aber häufig geopolitischen Motiven. Auch hier zeigt sich: Zwischen Anspruch und Wirklichkeit klafft eine Lücke.

WEGE ZU MEHR DIGITALER GERECHTIGKEIT

Digitale Gerechtigkeit lässt sich nicht allein durch technologische Innovationen herstellen. Vielmehr braucht es strukturelle Veränderungen auf politischer, wirtschaftlicher und gesellschaftlicher Ebene. Die folgenden Handlungsfelder sind zentral:

1. **Verbindliche ethische Standards und internationale Regulierung:** Es braucht globale Abkommen, die Prinzipien wie Transparenz, Fairness und Mitbestimmung weltweit verankern – etwa auf Basis der UN-Leitlinien für Wirtschaft und Menschenrechte⁹ oder von Data-Justice-Konzepten.
2. **Souveränität im Umgang mit Daten und Infrastruktur stärken:** Staaten und Communities im Globalen Süden müssen eigene digitale Ökosysteme aufbauen können – durch lokale Server, unabhängige Clouds, eigene Standards und langfristige Kapazitätsentwicklung.
3. **Faire Arbeitsbedingungen in der digitalen Ökonomie:** Digitale Lieferketten müssen menschenrechtlich überprüfbar werden. Dazu gehören faire Löhne, soziale Absicherung und psychologische Betreuung für Clickworker weltweit.
4. **Umweltschutz als Teil digitaler Verantwortung:** Die enormen Klimakosten von KI müssen internalisiert werden. Energieeffizienz, nachhaltiges Design und Kompensationsmechanismen für betroffene Regionen sind dringend notwendig.
5. **Entwicklungspartnerschaften auf Augenhöhe:** Statt technologische Abhängigkeiten zu schaffen, müssen Partnerschaften lokale Innovation, Bildung und demokratische Teilhabe fördern.

ZWISCHEN KRITIK UND ALLTAG: VERANTWORTUNG ÜBERNEHMEN

Was bedeutet das konkret für uns als KI-Nutzer*innen? Verantwortung zu übernehmen, heißt nicht, Innovation abzulehnen, sondern sich mit den gesellschaftlichen Folgen der Technologie auseinanderzusetzen. Das beginnt im Kleinen: Nutzer*innen können beispielsweise darauf achten, welche Anwendungen sie verwenden, wie viel energieintensive Prozesse wie Bild- oder Textgenerierung sie in Anspruch nehmen und welche Anbieter dahinterstehen.

Ein wichtiger erster Schritt ist zudem, sich mit der eigenen digitalen Umgebung vertraut zu machen. Nur wer weiß, wo und in welcher Form KI bereits im Alltag integriert ist – etwa in Suchmaschinen, sozialen Netzwerken, Sprachassistenten, Übersetzungsdiensten oder im Online-Shopping –, kann informierte Entscheidungen darüber treffen, ob und wie diese Technologie genutzt werden soll.

Konkret bedeutet das:

- KI bewusst und gezielt einsetzen: Nicht jede alltägliche Aufgabe muss durch KI automatisiert werden. Menschliches Urteilsvermögen bleibt oft sinnvoller – besonders in sensiblen Kontexten wie Pädagogik oder Beratung.
- Transparenz schaffen: Inhalte, die mithilfe von KI entstanden sind, sollten offen gekennzeichnet werden – besonders in Bildung, Medien oder öffentlichen Debatten.

⁹ Die UN-Leitlinien für Wirtschaft und Menschenrechte wurden 2011 vom UN-Menschenrechtsrat verabschiedet. Sie verfolgen das Ziel Menschenrechtsverletzungen zu verhindern. In den Leitlinien werden Verantwortlichkeiten und Verpflichtungen für Staaten und Unternehmen definiert und Empfehlungen zur Verhütung und Behebung von Menschenrechtsverletzungen gegeben.

- Anbieter kritisch prüfen: Welche Firma steckt hinter der Technologie? Unter welchen Bedingungen wurde sie entwickelt? Woher stammen die Trainingsdaten?
- Datensparsam handeln: Je weniger personenbezogene Daten weitergegeben werden, desto geringer ist das Risiko von Missbrauch. Es lohnt sich, Alternativen mit hoher Datensparsamkeit zu bevorzugen.
- Ressourcen schonen: Das Training großer KI-Modelle erfordert erhebliche Mengen an Energie und Wasser. Bewusste Nutzung bedeutet auch, aufwendige Prozesse (wie das Erstellen vieler Bilder oder Videos) zu hinterfragen.

Verantwortung zu übernehmen, heißt auch, sich über die eigene Rolle als Konsument*in, Nutzer*in oder Multiplikator*in bewusst zu werden – und gegebenenfalls Einfluss zu nehmen: durch Feedback, durch politischen Druck oder durch solidarisches Handeln.

BILDUNG ALS SCHLÜSSEL ZUM KRITISCHEN UMGANG MIT KI

Bildungsarbeit ist ein zentraler Hebel für eine gerechte digitale Gesellschaft. Sie schafft nicht nur Zugang zu Reflexion, Austausch und Empowerment, sondern hilft auch dabei, das digitale Umfeld besser zu verstehen. Denn wer nicht weiß, wo und wie KI bereits im Alltag integriert ist, kann auch keine fundierten Entscheidungen über deren Nutzung treffen. Bildungsarbeit ermöglicht es, Mechanismen zu erkennen, KI-gestützte Systeme zu identifizieren und deren Funktionsweise kritisch zu hinterfragen. Sie unterstützt Menschen dabei, zwischen sinnvoller Nutzung und potenziell problematischem Einsatz zu unterscheiden – und fördert so digitale Mündigkeit und Selbstbestimmung.

1. Kritisches Bewusstsein fördern: Künstliche Intelligenz muss als gesellschaftliches und politisches Thema begriffen werden. Bildungsangebote sollten die Entstehungsbedingungen von KI transparent machen und Machtverhältnisse offenlegen.
2. Perspektivwechsel ermöglichen: Globale Ungleichheiten werden oft nicht gesehen. Bildungsarbeit muss gezielt Stimmen und Perspektiven aus dem Globalen Süden einbinden.
3. Gestaltungs- und Mitbestimmungskompetenz aufbauen: Menschen brauchen praktische Fähigkeiten zur Nutzung, Bewertung und Gestaltung digitaler Systeme. Dazu zählen Datenkompetenz, Prompting-Fähigkeiten und rechtliches Wissen.
4. Vielfalt sichtbar und nutzbar machen: Digitale Bildung muss inklusiv sein. Dazu gehören barrierefreie Materialien, Leichte Sprache, diskriminierungssensible Inhalte und Schulungen für Lehrende.
5. Ethische und emotionale Dimensionen thematisieren: Technik ist nicht neutral. Bildungsarbeit sollte die ethischen Fragen und emotionalen Wirkungen digitaler Technologien erfahrbar und diskutierbar machen.

WAS WÄRE DIGITALE GERECHTIGKEIT?

Digitale Gerechtigkeit heißt nicht nur Zugang zu Technik, sondern auch faire Arbeitsbedingungen, Datensouveränität, kulturelle Selbstbestimmung und Klima- sowie Ressourcengerechtigkeit. KI ist politisch. Und ihre Nutzung ist eingebettet in globale Zusammenhänge.

DAS PRINZIP BARRIEREARMUT¹⁰ – KI ALS GESTALTUNGSFAKTOR FÜR TEILHABE

von Lukas Spahlinger

Künstliche Intelligenz ist in vielen Alltagsanwendungen präsent: Sie transkribiert Gespräche, sortiert Informationen, navigiert durch Städte und formuliert Texte. Für Bildungsarbeit und gesellschaftliche Teilhabe ist das kein Nebenaspekt. KI prägt, wie Menschen Zugang zu Bildung, Arbeit und öffentlichem Leben erhalten. Sie kann Barrieren senken, indem sie Informationen verständlicher macht, Bedienwege vereinfacht oder alternative Kommunikationsformen eröffnet. Genauso kann sie neue Hürden aufbauen, wenn Systeme nicht barrierearm gestaltet sind, Vorurteile reproduzieren oder Menschen mit geringeren digitalen Ressourcen ausschließen. Inklusion meint das Recht auf gleichberechtigte Teilhabe – unabhängig von individuellen Voraussetzungen. KI kann dieses Recht stärken, wenn sie partizipativ entwickelt, rechtlich abgesichert und in der Praxis kontinuierlich geprüft wird.

INKLUSION IM DIGITALEN ZEITALTER

Die UN-Behindertenrechtskonvention verpflichtet dazu, Strukturen so zu gestalten, dass niemand ausgeschlossen wird. Übertragen auf die digitale Welt bedeutet das: Lernplattformen, Veranstaltungsformate und Kommunikationskanäle müssen so beschaffen sein, dass Menschen mit unterschiedlichen Fähigkeiten sie gleichermaßen nutzen können. Inklusion ist damit kein nachträgliches „Add-on“, sondern ein Gestaltungsprinzip. Wer pädagogische Angebote plant, trifft bereits mit der Wahl der Plattform, der Materialformate und der Interaktionswege eine Entscheidung über Zugänge.

Mit KI verschiebt sich hier die Aufmerksamkeit: Nicht nur Inhalte, auch die Wege zu ihnen werden algorithmisch vermittelt. Automatisch erzeugte Übersetzungen, Untertitel, Transkripte, oder Vorlesefunktionen können Teilhabe erleichtern. Gleichzeitig entstehen neue Abhängigkeiten von Systemen, die nicht immer transparent sind und selten von allen verstanden werden. In der Erwachsenenbildung kommt hinzu, dass Zielgruppen sehr heterogen sind: Unterschiede in Sprache, Vorwissen, Seh- oder Hörvermögen, Motorik, Lesekompetenz und Technikzugang treffen in ein und demselben Kurs aufeinander. Wer Inklusion ernst nimmt, richtet deshalb sowohl das didaktische Design als auch den Technologieeinsatz konsequent auf Barrierearmut aus. Das meint den systematischen Prozess der Planung, Gestaltung und Umsetzung von Lehr- und Lernumgebungen und -materialien.

WIE KI BARRIEREN ABBAUEN KANN

KI-gestützte Funktionen können Informationszugang und Interaktion spürbar erleichtern. Automatische Untertitel und Live-Transkription senken Hürden in Vorträgen und Videokonferenzen; Vorlesefunktionen und Bildbeschreibungen unterstützen Menschen mit Sehbehinderung; Übersetzungen in einfache oder Leichte Sprache erleichtern den Einstieg in komplexe Themen; adaptive Lernszenarien nehmen Tempo und Schwierigkeitsgrad auf, ohne zu entmündigen. In Bibliotheken, Volkshochschulen und Beratungskontexten zeigt

¹⁰ Der Begriff „Barrierearmut“ betont, dass Barrieren selten vollständig abgebaut sind und Zugänglichkeit ein fortlaufender Prozess ist. Er steht für eine realistische, inklusive Perspektive, die verschiedene Formen von Barrieren (z.B. sprachliche, soziale, digitale) mitdenkt – ohne den Anspruch, vollständige Barrierefreiheit bereits erreicht zu haben

sich, dass der Zugewinn an Selbstständigkeit dann besonders groß ist, wenn Technik unaufdringlich in vorhandene Abläufe integriert wird.

Für die Praxis bedeutet das: KI wirkt am besten, wenn sie als Ergänzung zu gutem pädagogischen Handwerk verstanden wird. Ein Live-Vortrag wird inklusiver, wenn Referierende deutlich sprechen, Folien kontrastreich gestalten und zentrale Begriffe erklären; KI-Untertitelung bei einem Live-Vortrag kann diese Bemühungen verstärken und macht sie für mehr Menschen nutzbar. Ein Online-Kurs gewinnt, wenn Materialien klar gegliedert, sprachlich zugänglich und mobil gut bedienbar sind; die KI-gestützte Vereinfachung kann Verständnislücken schließen, ersetzt aber nicht die didaktische Verantwortung, komplizierte Fachsprache nur dort einzusetzen, wo sie nötig ist und erklärt wird.

Wesentlich ist, dass die Technik zuverlässig genug für den Alltag ist. Prototypen – etwa gebärdensprachliche Avatare – zeigen Potenziale, sind aber noch nicht belastbar genug für den Regelbetrieb. Pädagogische Angebote sollten das offen kommunizieren, Alternativen bereithalten und Rückmeldungen der Teilnehmenden systematisch auswerten.

WO KI NEUE HÜRDEN ERZEUGT

Mit KI verschärfen sich leicht bestehende Ungleichheiten. Wer keinen verlässlichen Internetzugang, keine aktuelle Hardware oder wenig digitale Routine hat, profitiert weniger – oder gar nicht. Zudem werden KI-Modelle auf Daten trainiert, die gesellschaftliche Verzerrungen enthalten. Ohne Gegenmaßnahmen übernehmen Systeme diese Verzerrungen: Dialekte und Akzente werden schlechter erkannt, bestimmte Schreibweisen von Namen falsch interpretiert, Lebensläufe mit untypischen Bildungswegen unterschätzt. Auch barrierearme Technik kann scheitern, wenn digitale Inhalte durch Screenreader nicht erkannt werden, die Webseitennavigation per Tastatur lückenhaft ist oder automatisch erzeugte Leichte Sprache Regeln missachtet und dadurch missverständlich wird; solche Ergebnisse sollten stets mit Nutzer*innen geprüft und bei Bedarf nachbearbeitet werden.

Ein weiterer Stolperstein ist der technologische Solutionismus: die Erwartung, soziale oder pädagogische Herausforderungen ließen sich primär technisch lösen. KI kann unterstützen, aber sie ersetzt weder Selbstverständnis, Rechtsrahmen noch politische Maßnahmen und schon gar nicht die professionelle pädagogische Beziehung. In der Bildungsarbeit ist es daher zentral, KI als Werkzeug unter vielen zu rahmen, Entscheidungsspielräume zu bewahren und analoge Alternativen nicht zu verdrängen.

Schließlich stellt sich die Frage der Verantwortlichkeit. Wenn Systeme Entscheidungen vorbereiten – etwa in der Auswahl von Lerninhalten –, müssen Kriterien transparent und überprüfbar sein. Pädagogische Teams tragen die Verantwortung, den Einsatz zu begründen, Folgen zu beobachten und bei Fehlentwicklungen gegenzusteuern.

GESTALTUNG FÜR DIE PRAXIS – WORAUF FACHKRÄFTE ACHTEN SOLLTEN

Barrierearmut beginnt vor der Toolwahl. Sinnvoll ist eine Bestandsaufnahme: Wer lernt mit? Welche Barrieren sind wahrscheinlich? Welche Geräte und Kenntnisse sind vorhanden?

Daraus ergeben sich Prioritäten. Wo viele mobile Endgeräte genutzt werden, ist eine saubere Darstellung auf kleinen Bildschirmen wichtiger als aufwendige Desktop-Oberflächen. Wo heterogene Sprachkompetenzen zu erwarten sind, helfen klare Texte, Glossare und ergänzende Erklärvideos, die von Künstlicher Intelligenz unterstützt, aber nicht ersetzt werden.

Rechtliche und normative Leitplanken geben Orientierung. Die Behindertenrechtskonvention der Vereinten Nationen (UN-BRK), seit 2009 in Deutschland in Kraft getreten, macht Teilhabe zum Maßstab. Nationale Regelungen zur Barrierefreiheit verpflichten Anbieter digitaler Dienste, Zugänge verbindlich zu sichern. Für die Praxis heißt das: Barrierearmut und Datenschutz werden von Anfang an mitgedacht. Plattformen sollten die gängigen Standards zur Bedienbarkeit und Kompatibilität mit Hilfstechnologien erfüllen, sensible Daten werden sparsam erhoben und transparent verarbeitet, die Herkunft und Funktionsweise zentraler KI-Komponenten wird – so weit möglich – erklärt.

Qualität entsteht durch Rückkopplung. Angebote sollten mit Menschen getestet werden, die unterschiedliche Bedürfnisse vertreten. Rückmeldungen zu Untertiteln, Tastaturbedienung, Kontrasten, Sprachebene oder Navigationslogik sind wertvoller als rein technische Prüfungen. In Kursen können kurze, wiederkehrende Feedbackschleifen eingebaut werden: Was hat funktioniert? Wo hakte die Technik? Welche Alternativen wünschen sich die Teilnehmenden? So wird Barrierearmut vom einmaligen Check zur kontinuierlichen Praxis.

Didaktisch bewährt sich eine multiperspektivische Planung. Auf der Inhaltsebene geht es um kritische KI-Kompetenz – etwa darum, wie maschinelles Lernen grob funktioniert, wo Fehlerquellen liegen und wie man Ergebnisse prüft. Auf der Gestaltungsebene wird eingeübt, wie barrierearme Formate aussehen: klare Struktur, verständliche Sprache, Medienvielfalt (Text, Audio, Bildbeschreibung), verlässliche Navigation. KI-Tools können hier als Lerngegenstand und als Hilfsmittel auftreten. Wer beispielsweise gemeinsam mit der Gruppe automatische Untertitel live testet, erfährt unmittelbar, welche Wörter falsch erkannt werden und wie sich Sprechtempo, Mikrofonqualität oder Fachjargon auswirken. Diese Erfahrung fördert Medienkritik und verbessert zugleich die eigene Vortragspraxis.

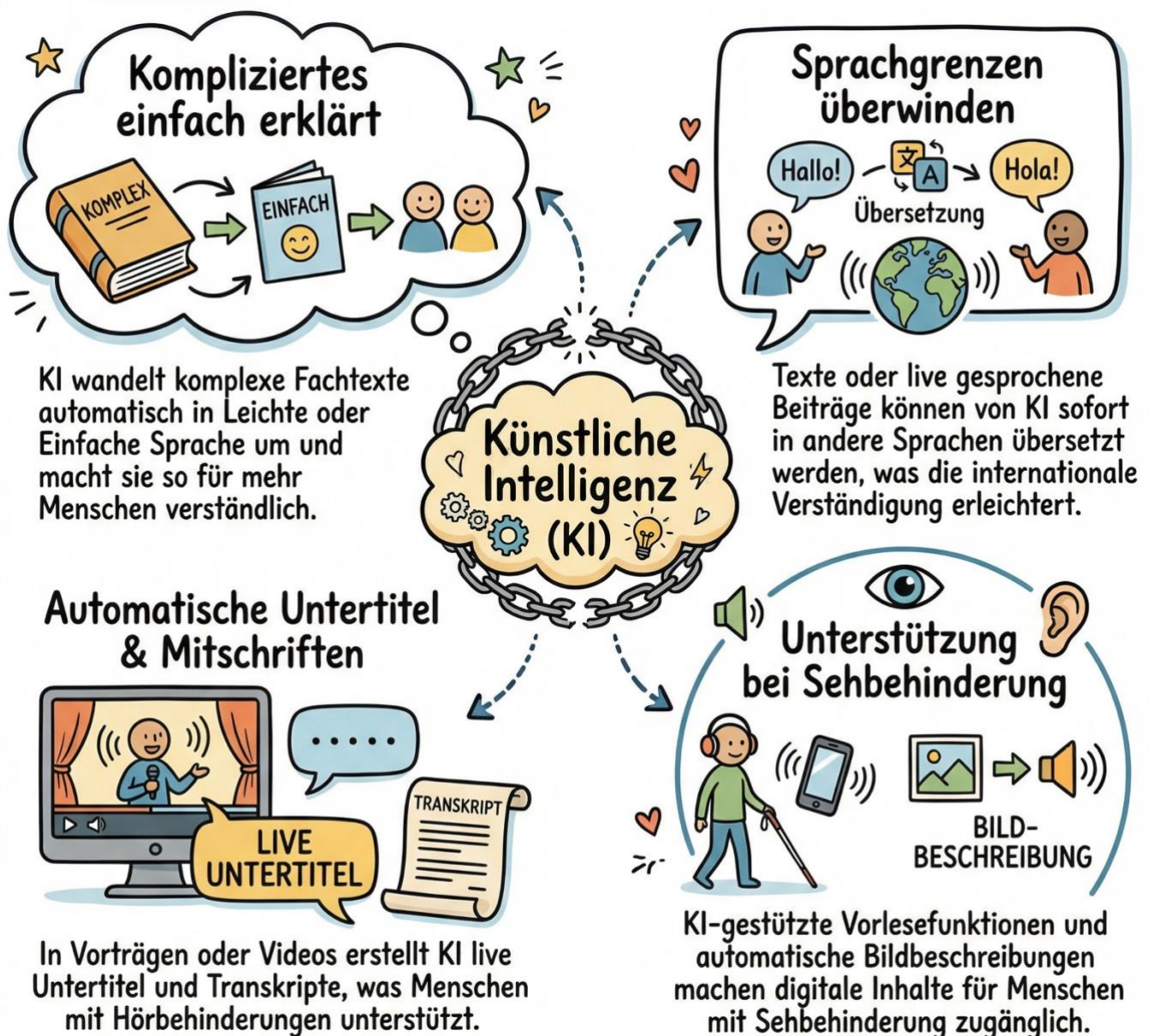
Nicht zuletzt braucht es einen reflektierten Umgang mit Künstlicher Intelligenz. Pädagogische Angebote sollten transparent machen, was KI kann und was nicht, wo Daten anfallen, welche Voreinstellungen genutzt werden und wie man sich abmeldet oder Alternativen erhält. Die Möglichkeit, sich entscheiden zu können, basiert auf wesentlichen Informationen zu den Rahmenbedingungen des Einsatzes. Zu welchem Zweck werden KI-basierte Systeme eingesetzt? In welchem Umfang werden sie eingesetzt? Darüber hinaus bedarf es einer Reflexion darüber, in welchen Szenarien ein Einsatz sinnvoll ist und wann nicht.

TEILHABE BY DESIGN

KI eröffnet reale Chancen für mehr Barrierearmut: Sie kann Informationen zugänglicher machen, Kommunikation erleichtern und Lernwege individueller gestalten. Damit diese Chancen bei den Menschen ankommen, muss Technik in inklusiven Prozessen entstehen, rechtliche Vorgaben verlässlich einhalten und in der Praxis fortlaufend geprüft werden. Pädagogische Fachkräfte haben dabei eine Schlüsselrolle. Sie übersetzen

menschenrechtliche Prinzipien in didaktische Entscheidungen, wählen Tools mit Augenmaß, schaffen transparente Rahmenbedingungen und halten Alternativen offen.

Die Leitfrage bleibt: Wer profitiert – und wer nicht? Wer so fragt, erkennt unbekannte Schwachstellen schneller, korrigiert Fehlentwicklungen früh und stärkt die Teilhabe aller. Eine inklusive Gesellschaft beginnt im Quellcode – und in den pädagogischen Entscheidungen, die ihn verantwortungsvoll in Bildungsprozesse einbinden.



EINFACHE SPRACHE MIT KÜNSTLICHER INTELLIGENZ

Von Tobias Albers-Heinemann, Lukas Spahlinger & Sabine Manning

Barrierearmut bedeutet auch, Hindernisse abzubauen, die Menschen den Zugang zu Informationen und gesellschaftlicher Teilhabe erschweren. Künstliche Intelligenz bietet hier ein konkretes und wirksames Werkzeug, um dieses Prinzip in die Praxis umzusetzen. Ein besonders eindrucksvolles Beispiel ist der KI-unterstützte Transfer von Fachtexten in Einfache Sprache. Dieser Ansatz zeigt, wie Technologie nicht nur theoretisch, sondern ganz praktisch dazu beitragen kann, Verständnisbarrieren zu überwinden und Wissen für alle zugänglich zu machen.

Einfache Sprache ist weit mehr als nur eine vereinfachte Version unserer Alltagssprache. Sie ist ein zentrales Werkzeug für Barrierefreiheit, gesellschaftliche Teilhabe und eine lebendige Demokratie. Während die Leichte Sprache gezielt für Menschen mit Lernbehinderungen und kognitiven Einschränkungen konzipiert wurde, wendet sich die Einfache Sprache an ein breites Publikum: an Menschen mit geringen Deutschkenntnissen, an ältere Mitbürger*innen, an Eltern im stressigen Alltag, an Berufstätige, die komplexe Informationen schnell erfassen müssen, oder an Bürger*innen, die behördliche Schreiben verstehen wollen, ohne juristische Vorkenntnisse zu haben. Kurz: An Menschen, die eine komplexere Fachsprache nicht gewohnt sind.

Sie ist somit eine bürgernahe Sprache, die eine Brücke schlägt zwischen der oft abstrakten Fachwelt und der konkreten Lebensrealität der Menschen. Ihr Ziel ist es, Informationen so aufzubereiten, dass sie für alle zugänglich und verständlich werden – unabhängig von Bildungshintergrund oder sprachlicher Kompetenz.

Seit 2024 bieten zwei internationale Normen eine klare Orientierung für die Gestaltung von Einfacher Sprache: die **ISO-Norm 24495-1** („Einfache Sprache – Grundsätze und Leitlinien“) und die **DIN-Norm 8581-1** („Einfache Sprache – Anwendung für das Deutsche“). Beide Normen verstehen Einfache Sprache als eine verständliche Form der Standardsprache, die durch klare Prinzipien geprägt ist. Dazu gehören kurze, prägnante Sätze mit maximal 15 bis 20 Wörtern, der Verzicht auf unnötigen Fachsprache zugunsten alltagsnaher, bekannter Begriffe, eine logische und übersichtliche Struktur mit Absätzen, Aufzählungen und Hervorhebungen sowie eine gut lesbare Schrift mit ausreichender Schriftgröße und klaren Kontrasten. Diese Merkmale sorgen dafür, dass Texte nicht nur leichter zu lesen, sondern auch leichter zu verstehen sind.

Doch Einfache Sprache ist kein starres Regelwerk, sondern ein flexibles Instrument, das sich an den Bedürfnissen, dem Vorwissen und den Lebenswelten der jeweiligen Zielgruppe orientiert. Sie strebt danach, Komplexität zu reduzieren, ohne die Inhalte zu verflachen – und ermöglicht so den Zugang zu Informationen, die ansonsten oft nur Expert*innen vorbehalten bleiben. Ob in Behörden, Bildungseinrichtungen, im Gesundheitswesen oder im öffentlichen Raum: Einfache Sprache schafft Verständnis, fördert Selbstständigkeit und stärkt die Teilhabe aller am gesellschaftlichen Leben.

Vor diesem Hintergrund stellt sich die Frage, wie Künstliche Intelligenz (KI) den Einsatz von Einfacher Sprache unterstützen kann. Welche Möglichkeiten bieten KI-gestützte Tools, um

Texte barriereärmer zu gestalten? Und welche Rolle spielen dabei menschliche Expertise, kritische Reflexion und die Verantwortung für eine wirklich inklusive Kommunikation?

KI ALS WERKZEUG FÜR EINFACHE SPRACHE

Künstliche Intelligenz bietet heute eine Vielzahl von Werkzeugen, die speziell darauf ausgelegt sind, Texte in Einfacher Sprache zu erstellen oder bestehende Texte zu vereinfachen. Diese Tools lassen sich grob in drei Kategorien einteilen:

- **Allzweck-KI-Modelle Modelle:** ChatGPT, Copilot oder Mistral AI sind vielseitig einsetzbar und können mit gezielten Prompts dazu gebracht werden, Texte in Einfacher Sprache zu generieren. Ihr Vorteil liegt in der Flexibilität: Sie können auf unterschiedliche Zielgruppen, Kontexte und Anforderungen zugeschnitten werden. Allerdings erfordern sie präzise Anweisungen, um wirklich verständliche und zielgruppengerechte Ergebnisse zu liefern.
- **Spezialisierte Programme:** Tools wie capito.ai, TextLab oder Wortliga sind speziell für die Vereinfachung von Texten entwickelt worden. Sie analysieren nicht nur die sprachliche Komplexität, sondern passen Texte aktiv an die Bedürfnisse der Nutzer*innen an. Sie erklären Fachbegriffe, kürzen lange Sätze und sorgen für eine klare Struktur. Besonders nützlich sind sie in Bereichen wie Verwaltung, Bildung oder Gesundheitswesen, wo Fachsprache oft eine Hürde darstellt.
- **Schnelle Vereinfachungstools:** Anwendungen wie Bürgernah, Fair Text oder Klar&Verständlich arbeiten mit internen Algorithmen, die Texte automatisch vereinfachen. Sie eignen sich besonders für den schnellen Einsatz, etwa bei der Überarbeitung von Formularen, Informationsblättern oder Websites.

PROMPTING: DIE KUNST DER PRÄZISEN ANWEISUNG

Die Qualität der KI-generierten Texte hängt maßgeblich von der Art der Eingabe ab. Ein guter Prompt für Einfache Sprache sollte folgende Elemente enthalten:

- **Zielgruppenbezug:** Wer soll den Text verstehen? (z. B. „für Eltern mit Migrationshintergrund“)
- **Sprachliche Vorgaben:** Welche Regeln sollen beachtet werden? (z. B. „maximal 15 Wörter pro Satz, aktiv formuliert“)
- **Kontext:** Wo wird der Text verwendet? (z. B. „für eine Elterninformation der Schule“)

Ein Beispiel für einen gelungenen Prompt: „Schreibe eine Einladung zum Elternabend in einfacher, bürgernaher Sprache. Verwende kurze Sätze, vermeide Behördendeutsch und erkläre alle Abkürzungen.“

Durch solche präzisen Vorgaben lassen sich Texte erzeugen, die den Kriterien der Einfachen Sprache entsprechen. Dennoch muss immer beachtet werden, dass KI-Tools keine Wunderwaffe sind: Gute Texte entstehen nicht auf Knopfdruck. Sie erfordern menschliche Vorarbeit, Begleitung und Nachbearbeitung – besonders, wenn es um Verständlichkeit und Zielgruppenorientierung geht. Ein einzelner Prompt bietet oft noch keine optimalen Ergebnisse. Deutlich bessere Resultate lassen sich erzielen, wenn angepasste Sprachmodelle oder spezialisierte KI-Agenten eingesetzt werden, die gezielt für die

Erstellung von Einfacher Sprache trainiert wurden. Angebote wie Custom GPT (von OpenAI) oder Mistral-Agenten ermöglichen es, KI-Modelle mit spezifischen Anweisungen und Beispielen zu „fine-tunen“, sodass sie konsistent hochwertige Texte in Einfacher Sprache liefern. Besonders praktisch: Die Agenten von Mistral können sogar kostenfrei in Le Chat erstellt und genutzt werden – ohne Programmierkenntnisse oder technische Hürden.

VERANTWORTUNGSVOLLER EINSATZ

KI kann den Prozess der Texterstellung in Einfacher Sprache deutlich erleichtern. Doch die Verantwortung für die Verständlichkeit der Texte, die inhaltliche Vollständigkeit und die Wahrung des Sprachniveaus liegt bei den Menschen, die diese Tools einsetzen. Es gilt sicherzustellen, dass die Vereinfachung der Sprache nicht auf Kosten der inhaltlichen Tiefe oder Korrektheit geht. Gleichzeitig müssen personenbezogene Daten geschützt werden, indem sie anonymisiert oder in datenschutzkonformen Umgebungen verarbeitet werden.



In einem Praxistest haben wir mit LeChat von Mistral und einem KI-Agenten eine Möglichkeit erstellt, mit wenigen Klicks komplexe Texte in einfacher Sprache zu transformieren.

LOKALE KI-MODELLE UND OPEN SOURCE: CHANCEN, GRENZEN UND RELEVANZ FÜR DIE BILDUNG

von Tobias Albers-Heinemann

Large Language Models (LLMs) – die Herzstücke generativer KI – prägen zunehmend unseren Alltag, die Wissenschaft und die Bildung. In den letzten Jahren hat sich eine Vielzahl solcher Modelle für ganz unterschiedliche Zwecke entwickelt. Während Dienste wie ChatGPT (OpenAI), Gemini (Google) oder Copilot (Microsoft) auf große, zentral gehostete Modelle setzen, die eine hohe Funktionstiefe und Benutzerfreundlichkeit versprechen, verfolgen Open-Source-Modelle einen anderen Ansatz: Sie sind frei zugänglich, transparent und oft lokal einsetzbar.

Diese Unterscheidung ist entscheidend: Kommerzielle Sprachmodelle existieren meist in verschiedenen Versionen mit unterschiedlich ausgeprägten Fähigkeiten, Preisen und Zugriffsmöglichkeiten. Aufgrund der häufig globalen Serverstruktur und der Datenübertragung in Drittländer ist bei ihrer Nutzung ein besonderes Augenmerk auf den Datenschutz zu legen – insbesondere in sensiblen Bereichen wie Bildung oder Verwaltung.

Ihnen gegenüber steht eine wachsende Landschaft an Open-Source-Sprachmodellen, die ebenfalls leistungsfähig sein können. Ihr großer Vorteil liegt in der Offenheit ihres Quellcodes, was eine transparente Nachvollziehbarkeit ihrer Funktionsweise ermöglicht. Darüber hinaus lassen sie sich individuell anpassen – etwa durch eigene Trainingsdaten oder gezielte Konfigurationen. Besonders hervorzuheben ist auch die Datenhoheit: Nutzer*innen entscheiden selbst, welche Daten verarbeitet werden und behalten jederzeit die Kontrolle darüber. Damit bieten Open-Source-Modelle eine starke Alternative für alle, die auf vertrauenswürdige, anpassbare und verantwortungsvoll einsetzbare KI-Lösungen setzen wollen.

WAS SIND LOKALE KI-MODELLE?

Lokale Sprachmodelle (Large Language Models, LLMs) sind zunächst dieselbe Art von Modellen wie die großen kommerziellen Sprachmodelle – sie basieren auf ähnlichen architektonischen Prinzipien und wurden mit umfangreichen Textmengen trainiert. Der entscheidende Unterschied liegt in ihrer Bereitstellung: Lokale Modelle werden mithilfe spezieller Software oder technischer Werkzeuge so installiert, dass sie auf eigenen Servern oder direkt auf dem persönlichen Computer betrieben werden können, ohne ständigen Zugriff auf das Internet oder zentrale Cloud-Dienste zu benötigen. Anders als bei Cloud-Lösungen wie ChatGPT oder Gemini verbleiben alle Daten lokal – ein großer Vorteil für den Datenschutz. Lokal bedeutet in diesem Zusammenhang auch: Es werden keinerlei Daten an externe Server übertragen. Nutzer*innen behalten die volle Datenhoheit und haben die Kontrolle darüber, welche Informationen verarbeitet und gespeichert werden – ein entscheidender Vorteil insbesondere in sensiblen Anwendungsbereichen wie Bildung, Sozialarbeit oder Verwaltung.

Ein weiterer Pluspunkt liegt im Bereich der Nachhaltigkeit: Kommerzielle KI-Anbieter betreiben riesige Rechenzentren, deren Energiebedarf für Training und Betrieb ihrer Modelle enorm ist. Lokale Modelle hingegen arbeiten – je nach Größe – deutlich energiesparsamer, da sie auf individueller Hardware laufen. Für viele Anwendungsszenarien – etwa

textbasierte Analysen, einfache Assistenzfunktionen oder interaktive Schreibunterstützung – reicht dieser eingeschränkte, aber dafür ressourcenschonendere Einsatz völlig aus und ermöglicht einen verantwortungsvolleren Umgang mit Energie und digitalen Infrastrukturen.

OPEN-SOURCE-KI: MODELLE, TOOLS UND VORAUSSETZUNGEN

Lokal installierbare Modelle gibt es in unterschiedlichen Größen und Varianten, etwa Qwen, Mistral, LLaMA, Gemma, Phi oder Deepseek. Diese Modelle sind häufig über Plattformen wie HuggingFace frei zugänglich. Programme wie LMStudio, Pinokio, GPT4All oder Ollama bieten grafische Oberflächen zur einfachen Bedienung auf dem lokalen PC. Darüber hinaus gibt es auch webbasierte Angebote wie <https://duck.ai>, die den Zugriff auf ausgewählte Open-Source-Modelle direkt im Browser ermöglichen – ganz ohne Installation. Allerdings sollte bei webbasierten Angeboten genau geprüft werden, wie der jeweilige Anbieter mit den eingegebenen Daten umgeht.

Eine hilfreiche Faustregel: Gibt es eine klar einsehbare Datenschutzerklärung? Werden Eingaben gespeichert oder weitergegeben? Solche Fragen helfen, die Seriosität eines Angebots besser einzuschätzen. Open Source bedeutet nicht auto-matisch, dass Datenschutz gewährleistet ist – besonders bei browserbasierten Anwendungen ist Vorsicht geboten.

Die große Vielfalt an verfügbaren Modellen stellt auch eine technische Herausforderung dar: Abhängig von der Anzahl der Parameter, der Modellarchitektur und der Komprimierung variieren die Systemanforderungen zum Teil erheblich. In der Regel erfordert die Installation und der reibungslose Betrieb leistungsfähiger Modelle einen Rechner mit deutlich mehr Ressourcen als ein Standard-Office-PC bietet, was aktuell für viele Personen mit einer etwas älteren Standard-IT ein Ausschlusskriterium darstellt.

WARUM IST DAS RELEVANT FÜR DIE BILDUNG?

In der Bildungsarbeit geht es um weit mehr als Effizienz oder technische Machbarkeit. Es geht um Vertrauen in Prozesse, um die Kontrolle über sensible Daten und um die Fähigkeit zur kritischen Reflexion über digitale Werkzeuge. Gerade hier bietet der Einsatz lokaler Open-Source-KI entscheidende Vorteile – sowohl aus pädagogischer als auch aus ethischer Perspektive:

- **Datenschutz:** Keine Daten verlassen das eigene Gerät. Die Verarbeitung erfolgt vollständig lokal, sodass Nutzer*innen die volle Kontrolle über alle Eingaben und Ergebnisse behalten – ein zentraler Vorteil für sensible Bildungs- und Sozialkontexte.
- **Transparenz:** Der Quellcode ist offen einsehbar. Dadurch lassen sich nicht nur Funktionsweisen nachvollziehen, sondern auch potenzielle Schwachstellen oder Verzerrungen identifizieren und beheben.
- **Anpassbarkeit:** Modelle können auf spezifische Bedarfe trainiert werden – sei es für inklusive Sprache, bestimmte Fachbegriffe oder die Unterstützung beim Verstehen komplexer Inhalte durch automatische Übersetzung in Leichte Sprache.
- **Diskriminierungssensibilität:** Durch die bewusste Auswahl und Gestaltung eigener Sprachmodelle lassen sich diskriminierende Verzerrungen vermeiden oder korrigieren. Dies schafft die Möglichkeit, KI-Systeme fairer, sensibler und

zielgruppengerechter einzusetzen – insbesondere in einer diversen und inklusiven Bildungsarbeit.

- Zugänglichkeit und Bildungsgerechtigkeit: Damit KI in Bildungsprozesse integriert werden kann, muss auch sichergestellt sein, dass sie für alle Beteiligten nutzbar und finanzierbar ist. Lokale Open-Source-Modelle können – sofern die nötige Infrastruktur vorhanden ist – hier einen wichtigen Beitrag leisten, weil sie unabhängig von Lizenzgebühren funktionieren und individuell angepasst werden können.

Diese Aspekte bilden die Grundlage für eine reflektierte, inklusive und datensensible Bildungsarbeit, denn lokale Open-Source-Modelle eröffnen damit nicht nur technische, sondern auch gesellschaftliche Gestaltungsspielräume.

KOMFORT VS. KONTROLLE: DAS SPANNUNGSFELD DER KI-NUTZUNG

Trotz aller Vorteile stehen Nutzer*innen oft vor einem Dilemma: Lokale Modelle bieten zwar Kontrolle und Datensicherheit, sind jedoch häufig weniger leistungsstark, bequem und vielseitig. Cloud-Dienste punkten mit leistungsstarker Infrastruktur, ständigem Online-Zugriff und einfacherer Bedienung – sie erfordern aber das Teilen von Daten und Vertrauen in die Anbieter.

Besonders im Bereich des Funktionsumfangs wird das Spannungsfeld deutlich: Die rasante Entwicklung im Bereich kommerzieller KI-Dienste sorgt dafür, dass ständig neue Funktionen integriert werden – von visueller Bildgenerierung über multimodale Kontexteingabe bis hin zu Anbindung an Drittanbieter-Tools oder KI-Sprachassistenten. Open-Source-Projekte können mit dieser Innovationsgeschwindigkeit oft nicht Schritt halten, insbesondere weil sie auf freiwillige Entwicklergemeinschaften und eingeschränkte Ressourcen angewiesen sind.

Dadurch entsteht mitunter der Eindruck, Open Source KI sei weniger leistungsfähig oder „veraltet“. Doch dieser Eindruck greift zu kurz: Während nicht alle Funktionen sofort verfügbar sind, bieten Open-Source-Modelle andere Qualitäten wie Transparenz, Anpassbarkeit und die Möglichkeit, den Entwicklungsprozess selbst mitzugestalten.

Open Source KI ist längst kein Nischenphänomen mehr. Lokale Sprachmodelle sind nicht nur praktikabel und datenschutzfreundlich, sondern eröffnen vielfältige Möglichkeiten für die Bildungsarbeit – etwa in der Sprachbildung, Projektarbeit oder der individuellen Lernunterstützung. Sie ermöglichen einen souveränen, reflektierten und kritischen Umgang mit KI – fernab von undurchsichtigen Blackbox-Lösungen großer Anbieter, erfordern jedoch eine gewisse Bereitschaft zur technischen Auseinandersetzung sowie leistungsstarke Hardware. Gleichzeitig liegt darin auch ein großes Potenzial: Wer sich einarbeitet, gewinnt nicht nur ein tieferes Verständnis für die Technologie, sondern kann KI gezielt für eigene pädagogische Ziele einsetzen – und so einen aktiven Beitrag zur digitalen Mündigkeit, zur Bildungsgerechtigkeit und zur Gestaltung einer offenen Lernkultur leisten.

PRÜFUNGSFORMATE IM ZEITALTER VON KI

von Tobias Albers-Heinemann & Lukas Spahlinger

Prüfungen stehen seit Langem im Zentrum des Bildungssystems, weil sie Wissen und Fähigkeiten sichtbar machen und vergleichen sollen. Heute aber rückt stärker die Frage in den Vordergrund, wie diese Funktionen im Zeitalter generativer KI überhaupt noch zuverlässig erfüllt werden können. Doch mit dem Aufkommen generativer KI geraten diese Funktionen ins Wanken. KI-Sprachmodelle können Aufgaben wie Recherchieren, Zusammenfassen oder Ausformulieren schneller und präziser lösen als viele Lernende. Lehrende wie auch Studierende fragen sich: Wie ernst zu nehmend sind Prüfungsleistungen noch, wenn KI sie unterstützen oder sogar übernehmen kann?

KLASSISCHE PRÜFUNGSFORMATE UNTER DRUCK

Essays oder Hausarbeiten geraten durch den Einsatz von KI zunehmend an ihre Grenzen. Mit wenigen Eingaben lassen sich heute komplette Texte generieren, was nicht nur Eigenständigkeitserklärungen fragwürdig erscheinen lässt, sondern vor allem die Frage aufwirft, ob solche Aufgaben den Bildungsansprüchen noch gerecht werden. Prüfungsformate geraten dadurch insgesamt unter Druck: Einige Hochschulen in Deutschland haben bereits das Format der Bachelorarbeit abgeschafft, weil nicht mehr nachvollziehbar ist, wer diese Arbeiten geschrieben hat. Bildungsanbieter reagieren mit unterschiedlichen Strategien – von Dokumentationspflichten über Reglementierungen bis hin zu Verboten der KI-Nutzung.

Zudem ist die emotionale Belastung vieler Lernenden hoch: Oft stehen sie unter dem Verdacht¹¹, ihre Arbeiten nicht selbst verfasst zu haben. Der Druck steigt zusätzlich, wenn Lehrkräfte auf unzuverlässige KI-Detektoren¹² vertrauen, die falsche Ergebnisse liefern und so zu ungerechten Bewertungen führen. Für Prüfungen stellt sich daher die Frage, wie Täuschungen vermieden und gleichzeitig faire Bedingungen gesichert werden können.

Eine mögliche Antwort auf diese Herausforderung liegt in einem grundlegenden Paradox: Wir messen den Erfolg einer revolutionären Technologie an einem Kompetenzmodell, das vor ihrer Entstehung entwickelt wurde. Die aktuellen Bildungsstandards entstanden zwischen 2003 und 2005 als Reaktion auf PISA 2000 – basierend auf einer Kompetenzdefinition von 2001. Diese Standards spiegeln die Anforderungen einer analogen Welt wider und erfassen nur bedingt die Herausforderungen einer KI-geprägten Realität. Das Problem ist dabei weniger technologischer als vielmehr struktureller Natur. Wenn Schüler*innen heute bei Hausaufgaben oder Tests „nur noch KI bedienen“ anstatt eigenständig zu denken, dann ist das nicht zwangsläufig ein Versagen der Technologie – sondern ein Hinweis darauf, dass unser Bewertungssystem nicht mehr zeitgemäß ist und die falschen Aufgaben gestellt wurden.

¹¹ Stone, A. (2022): Student Perceptions of Academic Integrity: A Qualitative Study of Understanding, Consequences, and Impact. *Journal of Academic Ethics*. Verfügbar unter: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9702763> [Zugriff: 19.09.2025].

¹² Coldwell, W. (2024): 'I received a first but it felt tainted and undeserved': inside the university AI cheating crisis. *The Guardian*. Verfügbar unter: <https://www.theguardian.com/technology/2024/dec/15/i-received-a-first-but-it-felt-tainted-and-undeserved-inside-the-university-ai-cheating-crisis> [Zugriff: 19.09.2025].

NEUE ANFORDERUNGEN AN PRÜFUNGEN IM KI-ZEITALTER

Wenn generative KI Routinetätigkeiten übernimmt, verschiebt sich der Fokus stärker auf andere Kompetenzen: die kritische Bewertung und Einordnung von Quellen, die Fähigkeit, Informationen im Kontext zu reflektieren sowie die Entwicklung einer eigenen Urteilsfähigkeit. Ebenso unverzichtbar ist der bewusste und reflektierte Umgang mit KI-gestützten Werkzeugen. „AI Literacy“ bedeutet daher weit mehr, als nur Prompts zu formulieren. Sie umfasst auch das Erkennen von Grenzen, der eigenen und technischen Möglichkeiten, das kritische Prüfen von Ergebnissen und das bewusste Dokumentieren des eigenen Vorgehens.

Prüfungen im KI-Zeitalter müssen daher über reine Wissensabfrage hinausgehen und auf umfassende, digitale und überfachliche Kompetenzrahmen Bezug nehmen. Sie sollen nicht nur technische Fähigkeiten, sondern vor allem reflexive, ethische und gestaltende Kompetenzen der Lernenden sichtbar machen. Dazu gehören unter anderem:

- Digitale Basiskompetenzen (z. B. nach EU DigComp 2.2, dem Referenzrahmen für digitale Kompetenzen der Europäischen Kommission): Informations- und Datenkompetenz, digitale Kommunikation und Zusammenarbeit, das Erstellen digitaler Inhalte, Sicherheit im Umgang mit Daten und digitale Problemlösefähigkeiten.
- AI Literacy und Data Literacy: Grundverständnis für KI-Systeme, kritische Ergebnisprüfung, Transparenz durch Dokumentation z. B. kritische Ergebnisprüfung und datenethische Sensibilität.
- Kritisches Denken und Problemlösekompetenz: Quellenkritik, Argumentationsfähigkeit und Kontextbewusstsein.
- Selbstreguliertes Lernen und Metakognition: Planung und Bewertung eigener Lernprozesse, Reflexionsfähigkeit und Motivation.
- Ethische und gesellschaftliche Verantwortung: Fairer Zugang zu Technologien, nachhaltige und diskriminierungsfreie Nutzung sowie aktive Mitgestaltung.
- Kommunikations- und Präsentationskompetenz: Multimodales Präsentieren sowie verständliches Visualisieren komplexer Inhalte.

Ein Aspekt, der dabei nicht übersehen werden darf, ist die Gefahr des sogenannten Skill Skipping: Lernende könnten geneigt sein, wichtige Lernschritte zu überspringen, weil KI viele Aufgaben schneller und scheinbar effizienter übernimmt – noch bevor ein grundlegendes Verständnis aufgebaut wurde. Hier braucht es eine bewusste Balance: Auch im Zeitalter leistungsfähiger Assistenzsysteme bleiben Basiskompetenzen zentral – etwa das strukturierte Argumentieren, das eigenständige Gliedern von Texten oder grundlegende mathematische Denkfähigkeiten. Prüfungen sollten daher nicht nur den Umgang mit KI sichtbar machen, sondern auch belegen, dass Lernende grundlegende Kompetenzen selbstständig aufbauen, anwenden und kritisch reflektieren können.

Ein weiterer zentraler Punkt ist die Fairness und damit die Frage nach Bildungsgerechtigkeit. Damit alle Lernenden gleiche Chancen haben, braucht es auch einen gleichwertigen Zugang zu den eingesetzten KI-Tools. Eine Recherche¹³ des Chaos Computer Clubs aus dem Jahr

¹³ Mühlhoff, R. & Henningsen, M. (2024): Chatbots im Schulunterricht!? media.ccc.de, 29.12.2024. Verfügbar unter: <https://media.ccc.de/v/38c3-chatbots-im-schulunterricht> [Stand: 19.09.2025].

2024 hat beispielsweise gezeigt, dass ein in Rheinland-Pfalz eingesetztes KI-System Arbeiten, die mit einer Bezahlversion von ChatGPT erstellt wurden, systematisch besser benotete als solche, die mit frei zugänglichen Modellen verfasst waren. Das stellt einen klaren Verstoß gegen den Anspruch auf Bildungsgerechtigkeit dar und verdeutlicht, dass Prüfungen nur dann fair sein können, wenn Transparenz über den KI-Einsatz besteht und keine ökonomischen oder gesellschaftlichen Hürden zu ungleichen Bewertungen führen.

LÖSUNGSANSÄTZE: PRÜFUNGEN NEU DENKEN

Die aktuellen Entwicklungen sind eine Chance, Prüfungen grundlegend weiterzuentwickeln. Das derzeitige Bildungs- und Prüfsystem ist bereits viele Jahrzehnte alt und stößt an seine Grenzen: Verbote oder bloße Reglementierungen reichen nicht mehr aus, um mit digitalen Veränderungen Schritt zu halten. Deshalb ist es unabdingbar, Prüfungsformate und Methoden grundlegend zu verändern. Im Folgenden werden vier zentrale Ansätze vorgestellt, die als Leitlinien dienen können, wie Prüfungen neu gedacht und umgesetzt werden können:

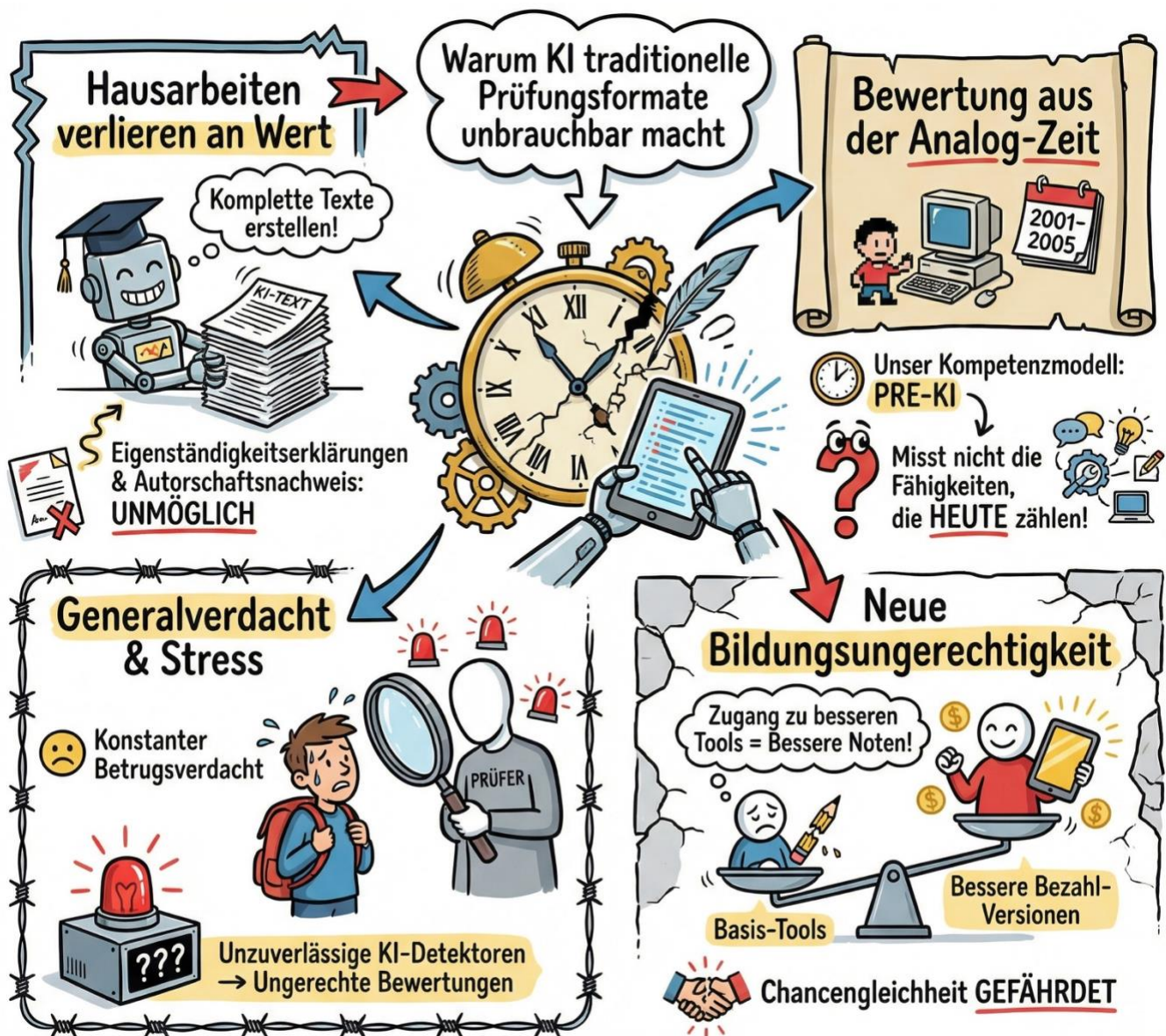
- **Projektbasiertes Lernen und Prüfen:** Reale, komplexe Problemstellungen können im Mittelpunkt stehen. Projekte, die fachübergreifend angelegt sind, fördern Zusammenarbeit, Kreativität und Problemlösungskompetenz. In Prüfungen lassen sich so nicht nur Ergebnisse, sondern auch die Fähigkeit bewerten, Wissen in neuen Kontexten anzuwenden.
- **Prozessorientiertes Prüfen:** Prüfungen müssen den Lernprozess stärker einbeziehen. Feedback-Schleifen, Peer-Reviews (Begutachtungsprozess, bei dem Experten aus demselben Fachgebiet die Qualität, Methodik und Validität von Arbeiten oder Handlungen bewerten) oder Lernjournale eröffnen Einblicke in Denkwege und Strategien. Das Ergebnis zählt, aber ebenso die Entwicklung dahin. So wird sichtbar, welche Kompetenzen aufgebaut wurden, auch wenn das Endprodukt nicht perfekt ist.
- **Integration von KI ins Prüfungssetting:** Statt KI zu verbieten, könnte sie Teil der Prüfung sein. Lernende dürfen Tools nutzen – jedoch mit der Verpflichtung, den Einsatz offenzulegen und kritisch zu reflektieren. Prüfungen könnten so nicht nur den fachlichen Output bewerten, sondern auch den verantwortungsvollen und transparenten Umgang mit KI.
- **Mündliche und kontextbezogene Prüfungen:** Auch stärker mündlich geprägte Formate, Verteidigungen oder Präsentationen bieten Chancen. Sie ermöglichen Lernenden, ihre Arbeit zu erläutern und zu reflektieren. Aufgaben mit persönlichem oder regionalem Bezug können KI-Einsatz erschweren und zugleich authentische Lerngelegenheiten schaffen.

PERSPEKTIVE: PRÜFUNGEN ALS LERNGELEGENHEITEN

Wenn Prüfungen mehr sein sollen als reine Kontrolle, müssen sie stärker als Lerngelegenheiten verstanden und gestaltet werden. Sie sollten nicht in erster Linie Defizite aufzeigen, sondern vielmehr Lernfortschritte sichtbar machen und würdigen. So können Prüfungen zu Räumen werden, in denen Lernende Selbstwirksamkeit erfahren – weil sie ihre Fähigkeiten zeigen, reflektieren und weiterentwickeln können. Entscheidend ist dabei

eine stärkere Betonung von Lernbegleitung und individuellen Lernwegen, statt am starren Format einer klassischen Abschlussarbeit festzuhalten.

Transferleistungen, mündliche Elemente und kontinuierliches Feedback müssen stärker in den Vordergrund treten. Das erfordert nicht nur neue Prüfungsaufgaben, sondern auch strukturelle Veränderungen in den Bildungseinrichtungen. Anpassungen an Curricula, Prüfungsordnungen, Gruppengrößen, Personaldichte und rechtliche Rahmenbedingungen sind notwendig, um diese Transformation zu ermöglichen. Der Gewinn kann jedoch groß sein: Prüfungen werden zu Spiegeln des Lernens, die Lernende nicht klein machen, sondern sie stärken und motivieren. Gleichzeitig eröffnet die individuelle Begleitung des Lernweges durch KI das Potenzial, den Bildungsbereich sogar menschlicher zu gestalten – weil sie Freiräume für persönliche Zuwendung, passgenaues Feedback und individuelle Förderung schafft.



KI-METHODEN IN DER BILDUNGSARBEIT

von Tobias Albers-Heinemann

Bildungsarbeit lebt von Expertise und Haltung. Pädagogische Fachkräfte bestimmen Inhalte, Methodik, Didaktik, Zielgruppenansprache und die Wahl des Formats. KI unterstützt diese Arbeit – sie ersetzt sie nicht. Die Technik soll dem pädagogischen Zweck folgen.

Dieser Beitrag zeigt praxiserprobte KI-Methoden für Multiplikator*innen: effizient dokumentieren, barriereärmer kommunizieren, kreativ erzählen, reflektieren und souverän moderieren. **Für alle Anwendungen gilt: Datenschutz und Persönlichkeitsrechte sind zu berücksichtigen.** Die rechtliche Einordnung und konkrete Vorgehensweisen erläutern wir im Beitrag „Künstliche Intelligenz und der Datenschutz“. Darüber hinaus betrachten wir Transparenz (Kennzeichnung von KI-Beiträgen), barrierearme Aufbereitung sowie die menschliche Qualitätsprüfung nicht nur als hilfreiche Hinweise, sondern als selbstverständlichen und verpflichtenden Bestandteil eines verantwortungsvollen KI-Einsatzes. Nachhaltigkeit, Inklusion und Diversität bleiben durchgängige Leitlinien. Zur KI-Kompetenz gehört ausdrücklich die Entscheidung, in welche Prozesse KI wie eingesetzt wird – und welche Prozesse bewusst beim Menschen bleiben. Methoden für den Einsatz von KI in Bildungsprozessen stehen aktuell noch nicht in großem Umfang zur direkten Übernahme bereit. Sie müssen erst erforscht, erprobt und evaluiert werden. Das setzt neben der Professionalität im Bildungskontext auch die Motivation voraus, aktiv, konstruktiv und regelmäßig mit KI zu arbeiten. Wir wollen ermutigen, diese Schritte zu gehen und gemeinsam die digitale Transformation im Bildungsbereich zu gestalten.

METHODEN & EINSATZSZENARIEN AUS DER PRAXIS

Die folgenden Methoden und Einsatzszenarien zeigen praxisnahe Beispiele, wie KI im Bildungsbereich eingesetzt werden kann. Sie knüpfen an die einleitend beschriebenen Grundgedanken an und verdeutlichen, wie sich theoretische Überlegungen in konkrete Arbeitsweisen übersetzen lassen. Jede Methode wurde erprobt und auf ihren Nutzen für Multiplikator*innen hin geprüft, um Inspiration für eigene Umsetzungen zu bieten.

Über die QR-Codes sind weiterführende und vertiefende Blogbeiträge und Videos unserer Website <https://erwachsenenbildung.digital> verlinkt. Sie können die Codes bequem mit Ihrem Smartphone fotografieren und die Links öffnen.

ANTRÄGE & SACHBERICHTE MIT CUSTOM-MODELLEN



Diese Methode zielt darauf ab, wiederkehrende Texte wie Anträge, Sachberichte oder standardisierte Projektbeschreibungen effizient zu erstellen – und gleichzeitig die Kompetenz zu entwickeln, KI-Modelle gezielt an diese spezifischen Aufgaben anzupassen. Durch das Trainieren oder Feintuning von Custom-Modellen können

Formulierungen, Fachbegriffe, Strukturen und formale Anforderungen so hinterlegt werden, dass die KI Texte erstellt, die passgenau zu den Bedürfnissen der Organisation und den Vorgaben von Fördergebern passen. Dies spart nicht nur Zeit, sondern erhöht auch die Konsistenz und Qualität der Texte.

Teilnehmer*innen: Sie müssen Custom-GPT-Modelle nach den entsprechenden Anforderungen der jeweiligen Maßnahmen und Anträge anpassen und trainieren. Dazu stellen sie passende Vorgaben und Beispiele bereit, die als Grundlage für das Training dienen.

KI: erstellt passgenaue Textentwürfe auf Grundlage der im Custom-GPT hinterlegten Vorgaben und der bereitgestellten Musterdateien. Diese Entwürfe können für Dokumentationen, Anträge, Berichte und andere formalisierte Textarten genutzt werden.

Ergebnis: konsistente und passgenaue Texte, die final geprüft werden – und durch automatisierte Prozesse bei Anträgen, Sachberichten und Dokumentationen eine enorme Zeitersparnis ermöglichen.

DOKUMENTATION EINES ONLINE-TALKS / PODIUMSDISKUSSION



Bei der Dokumentation eines Online-Talks oder einer Podiumsdiskussion werden die Inhalte so aufbereitet, dass sie nicht nur den Teilnehmer*innen vor Ort zugutekommen, sondern auch einem wesentlich breiteren Publikum zur Verfügung stehen. Dazu gehören insbesondere Personen, die nicht live teilnehmen konnten, aber durch eine klare Struktur, eine verständliche Sprache und ergänzende multimediale Elemente dennoch vom Inhalt profitieren. Auf diese Weise werden Wissen, Argumente und Diskussionen langfristig nutzbar gemacht und der Wirkungskreis der Veranstaltung deutlich erweitert.

Teilnehmer*innen: Sie haben keine aktive Rolle im technischen Prozess, sondern eine inhaltliche. Sie beteiligen sich am Gespräch, das aufgezeichnet wird, und stellen anschließend die Aufzeichnungen oder Transkripte zur Verfügung.

KI: fasst die Inhalte in leicht verständliche Abschnitte zusammen und erstellt daraus in einem mehrstufigen KI-Workflow verschiedene Endprodukte: Zunächst wird aus dem Transkript ein Blogbeitrag generiert, anschließend mithilfe des Custom-GPT „Schreiben fürs Hören“ eine für Audio geeignete Fassung erstellt, und schließlich wird dieser Text mit einem Tool wie ElevenLabs in einen Podcast oder eine Audiospur umgewandelt.

Ergebnis: multimediale, barrierearme Veranstaltungsdokumentation, die jederzeit nachgenutzt werden kann.

KI ALS CO-MODERATOR*IN



Bei dieser Methode übernimmt die KI eine aktive Rolle als Co-Moderatorin und wird gezielt in den Ablauf der Veranstaltung eingebunden. Sie entlastet die Moderator*innen, indem sie vorbereitete Texte wie An- und Zwischenmoderationen liefert, in Echtzeit thematische Impulse ergänzt und den Gesprächsfluss strukturiert. Zusätzlich kann sie Gesprächsbeiträge live in andere Sprachen übersetzen, um internationale und sprachlich diverse Teilnehmer*innen direkt einzubinden.

Teilnehmer*innen: Sie sind in diesem Fall die Moderator*innen der Veranstaltung und müssen ihre Moderation mit der KI genauso üben und abstimmen wie mit einer menschlichen Co-Moderation. Die Interaktion zwischen Mensch und KI wird dabei aktiv trainiert, um ein stimmiges Zusammenspiel und einen flüssigen Ablauf zu gewährleisten.

KI: übernimmt moderative Teile wie Anmoderationen, stellt Leitfragen, gibt Zeitmanagement-Hinweise und kann Gesprächsbeiträge in Echtzeit in andere Sprachen übersetzen, um eine mehrsprachige Moderation zu ermöglichen.

Ergebnis: strukturierte und dynamische Moderation mit entlastender Unterstützung, bei der auch die Möglichkeit der Echtzeitübersetzung in andere Sprachen gezielt genutzt wird, um ein mehrsprachiges Publikum aktiv einzubinden.

LERNEN MIT CONVERSATIONAL AI & KI-AVATAREN



Gesprächs- oder Beratungssituationen realitätsnah üben – und dabei gezielt Sprachen trainieren. In dieser Methode werden mithilfe einer Conversational AI oder virtuellen Avataren praxisnahe Dialoge simuliert, die nicht nur fachliche Inhalte abbilden, sondern auch den aktiven Sprachgebrauch fördern. Durch den Einsatz des Audio-Modus der KI, der bereits vorgestellt wurde, können Teilnehmer*innen in Echtzeit sprechen, zuhören und unmittelbar Feedback erhalten. So lassen sich etwa Rollenspiele in einer Fremdsprache durchführen, komplexe Beratungsgespräche in mehreren Sprachen simulieren oder die Aussprache gezielt verbessern.

Der Wechsel zwischen verschiedenen Sprachen wird dabei fließend unterstützt, was diese Methode besonders wertvoll für Sprachlernprozesse und interkulturelle Trainings macht. Da die Technologie für KI-Avatare derzeit noch am Anfang steht, ist diese Methode in der Praxis bisher noch wenig erprobt. Gleichzeitig liegt hier jedoch ein großes Potenzial für die Zukunft, da die Entwicklung rasant voranschreitet und mittelfristig deutlich realistischere, interaktive und vielseitige Anwendungsszenarien zu erwarten sind.

Teilnehmer*innen: Sie entwickeln ein Szenario und übernehmen Rollen.

KI: spielt die Gegenrolle, reagiert und kommuniziert in Echtzeit. Sie kann dabei vorgegebene Rollen einnehmen – etwa als Kund*in oder Expert*in – und in unterschiedlichen Sprachen agieren. So lassen sich gezielt mehrsprachige Dialoge üben und interkulturelle Gesprächssituationen realistisch simulieren.

Ergebnis: authentische Übungssituationen, in denen die Teilnehmer*innen realitätsnah agieren können, kombiniert mit gezielten Reflexionsphasen und umfassender Feedbackmöglichkeit. Die KI sorgt dabei nicht nur für eine konsistente Gesprächsführung und spontane Reaktionen, sondern unterstützt auch durch die Analyse des Dialogverlaufs, gibt Hinweise zu Verbesserungsmöglichkeiten und ermöglicht so eine vertiefte Auseinandersetzung mit Inhalten und Sprachgebrauch.

STORYTELLING MIT TEXT- & MUSIK-KI



Bei dieser Methode geht es darum, Themen so zu gestalten, dass sie nicht nur informativ, sondern auch emotional ansprechend und multimedial erlebbar werden. Die inhaltliche Botschaft wird durch eine kreative Erzählweise transportiert, die Spannung aufbaut, Emotionen weckt und durch den Einsatz verschiedener Medien – wie Text, Musik, Geräusche oder visuelle Elemente – verstärkt wird. Ziel ist es, ein Publikum nicht nur kognitiv, sondern auch auf einer emotionalen Ebene zu erreichen und dadurch einen nachhaltigeren Lerneffekt zu erzielen.

Teilnehmer*innen: Sie wählen ein komplexes Thema aus, legen fest, welche Zielgruppe sie erreichen möchten, und entscheiden sich für einen passenden Stil. Anschließend überlegen sie, wie das Thema mithilfe verschiedener KI-Formen – etwa für Text, Bild, Video oder Musik – neu und kreativ aufbereitet werden kann.

KI: transformiert z. B. komplexe Texte in kreative Liedtexte und passt diese inhaltlich an die Zielgruppe an. Anschließend wird mit einer weiteren KI-Anwendung dieser Liedtext – je nach gewünschtem Musikstil der Zielgruppe – in einen radiofähigen Song umgewandelt.

Ergebnis: Die Inhalte eines komplexen Textes werden über ein aktuelles und lebensweltnahes Medium – in diesem Fall Musik – der Zielgruppe niederschwelliger und ansprechender vermittelt. Das Ergebnis kann ein kreatives Endprodukt wie ein Hörspiel oder ein Song sein, der fachlich geprüft und gezielt auf die Lebenswelt der Zielgruppe abgestimmt ist.

TEXTTRANSFORMATION IN LEICHTE & EINFACHE SPRACHE



Bei dieser Methode werden Texte so gestaltet, dass sie von möglichst vielen Menschen problemlos verstanden werden können. Dazu gehört, komplizierte Satzstrukturen zu vereinfachen, Fachbegriffe zu erklären und unnötige Fremdwörter zu vermeiden. Ziel ist es, Inhalte klar, prägnant und barrierearm zu formulieren, damit auch Personen mit unterschiedlichen Sprachkenntnissen oder Bildungshintergründen die Informationen vollständig erfassen und nutzen können.

Teilnehmer*innen: Sie geben den Originaltext vor.

KI: überträgt ihn in Leichte oder Einfache Sprache gemäß festgelegter Regeln.

Ergebnis: verständliche, zugängliche Texte ohne unnötige Fachbegriffe.

CLUSTERN VON POST-ITS, FEEDBACKKARTEN UND BRAINSTORMINGS



Bei dieser Methode entsteht aus einer zunächst unübersichtlichen, bunten Mischung von Ideen, Meinungen und spontanen Anmerkungen Schritt für Schritt eine klar nachvollziehbare und logisch gegliederte Themenstruktur. Die verstreuten Gedanken aus einem Workshop oder einer Brainstorming Runde werden gesammelt, sortiert und in Beziehung zueinander gesetzt. So verwandelt sich ein chaotisches Zettelbild in eine verständliche, strukturierte Übersicht, die für alle Beteiligten greifbar ist und als Grundlage für die weitere Arbeit oder Auswertung dient.

Teilnehmer*innen: Sie notieren Gedanken und Vorschläge auf Karten oder Post-its auf Flipcharts, Whiteboards oder Metaplankartenwänden, machen Fotos oder scannen sie ein.

KI: erkennt und liest den Text aus den Bildern, listet die Inhalte auf und ordnet sie in sinnvolle Kategorien oder clustert sie nach den Bedarfen und Zielen der Fachkräfte.

Ergebnis: eine geordnete Übersicht mit klar benannten Themenfeldern, die in kürzester Zeit erstellt werden kann und dadurch eine erhebliche Zeitersparnis bietet. Sie dient als solide Basis für Berichte, Präsentationen oder die weitere Planung und ermöglicht es, schneller von der Rohsammlung an Ideen zu konkreten, umsetzbaren Arbeitspaketen zu gelangen.

VON DER IDEE ZUM BLOGARTIKEL (VOICE- & CANVAS-WORKFLOW)



In dieser Methode werden Rohideen aus Brainstormings oder Gesprächen in einem lebendigen, gemeinsamen Prozess – getragen von kollegialer Beratung – systematisch zu einer klaren, gut strukturierten Textvorlage weiterentwickelt. Unterschiedliche Perspektiven kommen zusammen, werden im Dialog geprüft, geschärft und ergänzt. Die Teilnehmer*innen bringen ihre Sichtweisen ein, während die KI strukturiert, ordnet und formuliert. So entsteht ein inhaltlich fundierter und gut gegliederter Entwurf. Der Prozess verbindet Tiefe mit Tempo: Aus ersten, teils ungeordneten Notizen wird in kurzer Zeit ein tragfähiger Text, der bereits nahe an der publikations- oder präsentationsreifen Endfassung liegt.

Teilnehmer*innen: Sie besprechen ihre erfassten Stichpunkte mit einer Conversational AI oder einem Voice-Bot einer KI, beispielsweise von ChatGPT.

KI: hört zu, dient als aktiver Gesprächspartner für eine kollegiale Beratung, stellt inhaltliche Rückfragen und hinterfragt getätigte Aussagen. Dabei fasst sie zusammen, strukturiert die Inhalte und bringt sie in eine logische Reihenfolge. Aus diesem Prozess wird anschließend eine erste Textfassung erstellt.

Ergebnis: ein gegliedertes Manuskript, das nur noch fachlich und stilistisch finalisiert werden muss.

ÜBERSETZUNG & SCHREIBASSISTENZ



Sprachliche Barrieren überwinden – also Hindernisse, die durch unterschiedliche Muttersprachen, Sprachkenntnisse oder Fachjargon entstehen, gezielt abbauen. Dies geschieht durch den Einsatz von KI-gestützten Übersetzungs- und Schreibassistenzsystemen, die Texte schnell, präzise und in einem angemessenen Ton in andere Sprachen übertragen. So können Inhalte für ein breiteres, mehrsprachiges Publikum zugänglich gemacht und Missverständnisse reduziert werden.

Teilnehmer*innen: Sie liefern den Ausgangstext.

KI: übernimmt die Übersetzung in die gewünschte Sprache und vereinfacht den Text bei Bedarf. Zusätzlich unterstützt die KI bei redaktionellen Prozessen, indem sie Rechtschreibung und Grammatik prüft, inhaltliche Unstimmigkeiten erkennt und Korrekturen vorschlägt. Sie agiert somit nicht nur als Übersetzerin, sondern auch als Lektorin und kritische Begleiterin, um die Qualität und Verständlichkeit der Inhalte zu sichern.

Ergebnis: verständlichere und bessere Inhalte in der gewünschten Sprache.

KI IN ONLINE-FORMATEN

In Online-Formaten eröffnen sich zusätzliche Möglichkeiten der KI-Unterstützung, die stark von den Funktionalitäten der eingesetzten Software abhängen. So bietet der AI Companion von Zoom – mittlerweile auch über Server im europäischen Wirtschaftsraum nutzbar – eine Reihe praktischer Funktionen: Er kann während einer Veranstaltung automatisch ein Transkript erstellen, Untertitel in Echtzeit generieren und diese bei Bedarf direkt in mehrere

Sprachen übersetzen. Darüber hinaus ermöglicht er es, im Anschluss mit der KI über die Inhalte der Veranstaltung zu sprechen, Fragen zu stellen oder Zusammenfassungen zu erhalten. Diese Werkzeuge erhöhen die Barrierefreiheit, unterstützen mehrsprachige Gruppen und erleichtern den Zugang zu Inhalten – auch für Personen, die nicht live teilnehmen konnten. Damit wird die Nachbereitung effizienter, und die Veranstaltung kann einem breiteren Publikum zugänglich gemacht werden.

SOUVERÄN GESTALTEN

Wir wollen Mut machen, diese Ansätze nicht nur zu nutzen, sondern auch selbst neue Ideen zu entwickeln, sie in der Praxis zu erproben und die gewonnenen Erfahrungen zu teilen. Entscheidend bleibt, dass die pädagogische Kompetenz und das fachliche Urteilsvermögen der Fachkräfte immer im Zentrum stehen. Die Technik – einschließlich KI – ist und bleibt ein Werkzeug, das gezielt eingesetzt werden sollte, um Bildung zu gestalten. Dabei müssen wichtige Aspekte wie Persönlichkeitsrechte, Datenschutz und ethische Standards konsequent berücksichtigt werden – so, wie wir es auch in unserer Broschüre „KI-Kompetenz“ beschrieben haben.

AUSBLICK: KI-KOMPETENZ IN ZEITEN GESELLSCHAFTLICHER TRANSFORMATION

Die rasante Verbreitung und tiefgreifende Wirkung von KI-Anwendungen in der digitalen Welt sind bemerkenswert. Künstliche Intelligenz ist längst omnipräsent: Sie steckt in Smartphones, Social-Media-Plattformen, Spielzeugen, Haushaltsgeräten oder Lernumgebungen und wirkt oft unsichtbar. Dabei prägt sie nicht nur unsere Kommunikation und Routinen, sondern auch unser Familienleben, den Bildungsbereich und die Kultur tiefgreifend. Die Beiträge dieses Buches haben gezeigt: KI ist keine technische Randerscheinung und keine App, bei der wir bewusst entscheiden können, ob wir sie nutzen oder nicht. KI ist längst Teil der hybriden Welt und somit ein Faktor gesellschaftlicher Transformation.

VERÄNDERUNGEN IN ALLTAG UND ARBEITSWELT

Diese Transformation vollzieht sich auf vielen Ebenen. KI verändert, wie wir Wissen verarbeiten, wie wir politische Informationen bewerten, wie wir unsere Arbeit organisieren oder was wir unter Kreativität verstehen. Sie wirft Fragen nach Macht und Gerechtigkeit auf: Wer kontrolliert die Daten? Wer profitiert wirtschaftlich? Wer trägt die ökologischen Kosten? Globale Abhängigkeiten, wie sie im digitalen Kolonialismus sichtbar werden, sind dabei ebenso relevant wie die Veränderungen im unmittelbaren Alltag – ob in Familien, im Klassenzimmer oder im Gespräch mit einer digitalen Assistenz. So verändern sich auch die Anforderungen und der Bedarf an Kompetenzen. Mit KI verändern sich Arbeitsprozesse und Rollen grundlegend: Routinetätigkeiten werden zunehmend automatisiert, während kreative, soziale und strategische Aufgaben stärker in den Vordergrund rücken.

Im Zuge dieser Transformationen verändern sich auch die Anforderungen an den Menschen. Arbeitsprozesse und Rollen verändern sich grundlegend: Routinetätigkeiten werden zunehmend automatisiert, während kreative, soziale und strategische Aufgaben stärker in den Vordergrund rücken. Viele Berufsbilder wandeln sich, neue Tätigkeitsfelder entstehen und klassische Rollenprofile lösen sich auf. Für Beschäftigte bedeutet das, sich kontinuierlich weiterzubilden, neue Kompetenzen zu entwickeln und flexibel auf technologische Veränderungen zu reagieren.

Doch nicht nur Beschäftigte stehen vor neuen Herausforderungen. Auch Arbeitgeber, Firmen und Einrichtungen müssen sich einem tiefgreifenden, transformativen Prozess stellen. Sie sind gefordert, Arbeitsprozesse kritisch zu hinterfragen, neu zu gestalten und bewusst zu entscheiden, welchen Stellenwert KI in der Institution und in der täglichen Arbeit einnehmen soll. Gleichzeitig gilt es klar zu definieren, welche Aufgaben weiterhin vom Menschen erledigt werden müssen – sei es aus ethischen, qualitativen oder strategischen Gründen. Diese Entscheidungen prägen nicht nur die Effizienz, sondern auch die Kultur und Werte einer Organisation.

MENSCH-MASCHINE-INTERAKTION UND ANTHROPOMORPHISIERUNG

Ein klarer Trend zeigt, dass die Interaktion mit KI-Systemen zunehmend verbal erfolgt. Immer mehr Menschen kommunizieren mit KI in natürlicher Alltagssprache, was die Grenzen zwischen Mensch und Maschine weiter verschwimmen lässt. Hinzu kommt die Tendenz zur Anthropomorphisierung von KI-Modellen: Sprachassistenten oder Chatbots

werden mit Namen, Stimmen und Persönlichkeiten ausgestattet, wodurch sie als fast menschliche Gegenüber wahrgenommen werden. Diese Entwicklung verstärkt nicht nur die emotionale Bindung zu Maschinen, sondern beeinflusst auch unser Verständnis von Authentizität und sozialen Verbindungen.

Die wachsende Vertrautheit und das Vertrauen in KI-Systeme können tiefgreifende Auswirkungen haben – sowohl auf die Beziehung zu den Systemen selbst als auch auf das reale soziale Umfeld. Wenn Menschen KI als vertrauenswürdige Gesprächspartner erleben, kann dies die Art und Weise verändern, wie sie Beziehungen gestalten und wie sie soziale Interaktionen bewerten. Besonders verlockend ist dabei, dass die Interaktion mit KI-Systemen oft konfliktfreier verläuft als zwischenmenschliche Kontakte. KI-Systeme sind aktuell nicht darauf ausgelegt, zu widersprechen oder Kritik zu üben. Stattdessen bieten sie scheinbar verständnisvolle und leicht zugängliche Ansprechpartner. Die Möglichkeit besteht, dass reale soziale Kontakte vernachlässigt werden, wenn die Interaktion mit KI als weniger anstrengend und risikoarm empfunden wird. Während menschliche Beziehungen oft von Meinungsverschiedenheiten, Emotionen und Kompromissen geprägt sind, bieten KI-Systeme eine scheinbar harmonische Alternative.

Doch genau diese Konflikte und Herausforderungen sind es, die zwischenmenschliche Beziehungen tiefgründig und authentisch machen. Wenn wir uns zunehmend in die scheinbar sichere und vorhersagbare Welt der KI zurückziehen, könnte dies langfristig unsere Fähigkeit beeinträchtigen, mit den Komplexitäten und Widersprüchen realer sozialer Interaktionen umzugehen. Die Art der Kommunikation mit Maschinen prägt somit nicht nur unsere Erwartungen an Technologie, sondern auch an menschliche Beziehungen. Es stellt sich die Frage, wie wir sicherstellen können, dass die Bequemlichkeit der KI-Interaktion nicht auf Kosten der Tiefe und Vielfalt zwischenmenschlicher Verbindungen geht.

CHANCEN, RISIKEN UND RAHMENBEDINGUNGEN

Mit den Entwicklungen verbinden sich große Versprechen. KI kann Teilhabe stärken, Barrieren abbauen und Lernprozesse individueller gestalten. Sie kann helfen, Ressourcen nachhaltiger zu nutzen oder kreative Ausdrucksformen zu erweitern. Gleichzeitig birgt sie erhebliche Herausforderungen: Energie- und Wasserverbrauch steigen, Datenschutz und Privatsphäre geraten unter Druck, Urheberrechtsfragen bleiben ungelöst, diskriminierende Muster werden fortgeschrieben und täuschend echte Deepfakes bedrohen Vertrauen und Demokratie. Gerade diese Ambivalenz – zwischen Potenzial zur Verbesserung und Risiko der Verschlechterung – macht Bildung zum Schlüssel.

KI lässt sich nicht einfach in „gut“ oder „schlecht“ polarisieren. Vielmehr bewegen wir uns in einem Spannungsfeld, in dem Chancen und Risiken eng beieinanderliegen. Diese Ambiguität auszuhalten und produktiv zu gestalten, wird zu einer zentralen Aufgabe. KI darf dabei nicht nur auf den individuellen Einsatz reduziert werden, sondern erfordert auch eine umfassende Folgenabschätzung: Unter welchen Rahmenbedingungen wird sie eingesetzt, mit welchen gesellschaftlichen Auswirkungen? Die Rahmenbedingungen entscheiden, ob KI Zugänge verbessert und Barrieren abgebaut werden – oder ob neue Hürden entstehen. Umso dringlicher braucht es verantwortungsbewusste und mündige Mitbürger*innen, die Künstliche Intelligenz ganzheitlich verstehen, ihre Widersprüchlichkeit anerkennen und kompetent mit ihr umgehen.

BILDUNGSAUFGABEN IM WANDEL

Der Bildungsbereich steht dabei besonders im Fokus. Zum einen verändert er sich selbst: Prüfungsformate geraten unter Druck, weil KI-gestützte Systeme Antworten liefern können, die herkömmliche Leistungsmessungen infrage stellen. Auch didaktische Konzepte, wie sie über Jahrzehnte gewachsen sind, müssen neu gedacht und an eine zeitgemäße Lern- und Lebenskultur angepasst werden. Zum anderen geht es darum, die breite Gesellschaft auf diese Transformation vorzubereiten. Neue Bildungsformen und inhaltliche Formate müssen entwickelt werden, um Menschen zu qualifizieren, zu begleiten und in ihrer Selbstbestimmung zu stärken.

Dabei kommt den Fachkräften eine zentrale Rolle zu: Sie müssen nicht nur selbst die notwendigen Fähigkeiten erwerben, sondern auch tragfähige Konzepte entwickeln, um diese Kompetenzen an Lernende weiterzugeben. Dies erfordert jedoch mehr als nur individuelle Anstrengungen. Es braucht einen klaren gesellschaftlichen und institutionellen Auftrag an den Bildungsbereich, der mit konkreten Maßnahmen unterfüttert wird. Dazu gehört, dass Verantwortliche ausreichend zeitliche Ressourcen erhalten, um sich intensiv mit den Möglichkeiten und Herausforderungen von KI auseinanderzusetzen. Gleichzeitig müssen sie Zugang zu moderner Ausstattung und Fortbildungsangeboten bekommen, die sie befähigen, KI nicht nur zu verstehen, sondern auch kritisch und kreativ im Bildungsprozess einzusetzen. Institutionen sind gefordert, Rahmenbedingungen zu schaffen, die es ermöglichen, KI als Werkzeug und Thema in Lehrpläne und pädagogische Praxis zu integrieren. Nur so kann sichergestellt werden, dass Bildungsverantwortliche ihrer Vermittlerrolle gerecht werden und Lernende auf eine Welt vorbereiten, in der KI eine immer größere Rolle spielt.

DIE ROLLE VON KI-KOMPETENZ

Damit rückt der Begriff der KI-Kompetenz ins Zentrum. Obwohl der genaue Umfang von KI-Kompetenz vom Gesetzgeber bisher noch nicht abschließend festgelegt wurde, haben wir uns im Rahmen unseres Projekts mit vielen Fachleuten aus dem Bildungsbereich auf eine inhaltliche Gliederung verständigt, die fünf zentrale Dimensionen umfasst. Diese bieten eine fundierte Orientierung für einen verantwortungsvollen und reflektierten Umgang mit Künstlicher Intelligenz:

- **Technische Dimension:** Grundverständnis der Funktionsweise von KI, wie Algorithmen, Daten und Autonomiegrade zusammenwirken, sowie ein Bewusstsein für Risiken wie Datenqualität, Cybersecurity, Fehleranfälligkeit und ökologische Kosten. Dazu gehört auch die Fähigkeit, KI-Tools gezielt auszuwählen und in Lehre und Verwaltung einzusetzen.
- **Praktische Dimension:** die Kompetenz, KI-Tools sinnvoll in den Arbeitsalltag zu integrieren, didaktisch zu nutzen und diskriminierungssensibel zu prompten, um faire und inklusive Ergebnisse zu erzielen.
- **Ethische Dimension:** Auseinandersetzung mit Fragen der Fairness, Diskriminierung und globalen Machtverhältnisse in KI-Systemen. Hier geht es darum, ethische Leitlinien für den Einsatz von KI zu entwickeln und kritisch zu hinterfragen, wie KI gesellschaftliche Werte und Normen abbildet oder verstärkt.

- **Rechtliche Dimension:** Kenntnisse über Datenschutz, Urheberrecht und Compliance, um KI-Anwendungen rechtssicher und transparent zu nutzen, insbesondere in Bildungszusammenhängen.
- **Gesellschaftliche Dimension:** Kritische Auseinandersetzung mit den Auswirkungen von KI auf Demokratie, Arbeitsmarkt, soziales Miteinander und Bildungsgerechtigkeit. Dazu gehört die Sensibilisierung für mögliche Risiken ebenso wie die Förderung von Teilhabe und Inklusion. Ziel ist eine reflektierte Urteilsfähigkeit und ein Bewusstsein im Umgang mit Chancen und Gefahren.

KI-Kompetenz ist mehr als das Bedienen von Programmen. Sie verbindet technisches Wissen mit ethischer Reflexion, rechtlicher Sicherheit und gesellschaftlicher Verantwortung. Individuell stärkt sie digitale Mündigkeit, gesellschaftlich schafft sie die Grundlage für Resilienz, Teilhabe und eine nachhaltige digitale Transformation.

WELCHE KOMPETENZEN BRAUCHEN WIR IN EINER WELT MIT KI?

In einer von KI geprägten Welt werden zukunftsweisende Kompetenzen immer wichtiger: Technische Grundbildung hilft uns, Funktionsweisen und Grenzen von KI zu verstehen. Ethische Urteilskraft ermöglicht es, Verantwortung und Werte in Entscheidungen einzubeziehen, während demokratische Resilienz uns befähigt, zwischen Fakten und Manipulation zu unterscheiden. Gestaltungskompetenz im Sinne der Bildung für nachhaltige Entwicklung (BNE) unterstützt dabei, Zielkonflikte zu erkennen und tragfähige Lösungen zu entwickeln. Nicht zuletzt stärkt digitale Mündigkeit die Fähigkeit, Technologien kritisch zu prüfen, deren Nutzung bewusst zu steuern und Alternativen einzufordern.

Für den Bildungsbereich bedeutet dies eine tiefgreifende Transformation: Prüfungen, Lernmaterialien, Methoden und Rollenbilder müssen überdacht und weiterentwickelt werden. KI bietet die Chance, Routinen zu hinterfragen, neue Formen der Teilhabe zu schaffen und Lernende wie Lehrende aktiv in die Gestaltung einzubeziehen. Statt KI als Bedrohung zu betrachten, können wir sie nutzen, um Bildung gerechter, inklusiver und zukunftsfähiger zu gestalten.

EINLADUNG ZUR MITGESTALTUNG EINER DIGITALEN ZUKUNFT

Dieser Ausblick ist mehr als eine Bestandsaufnahme – er ist ein Aufruf zur aktiven Mitgestaltung. Künstliche Intelligenz wird unsere Gesellschaft, unsere Arbeitswelt und unser Zusammenleben grundlegend verändern. Doch diese Veränderung ist kein automatischer Prozess, dem wir hilflos gegenüberstehen. Vielmehr liegt es in unserer gemeinsamen Verantwortung, die Richtung zu bestimmen, in die sich diese Entwicklung bewegt. Jede und jeder von uns kann einen Beitrag leisten – sei es in Bildungseinrichtungen, die neue Lernformate entwickeln, in Unternehmen, die KI sozialverträglich einsetzen, in Familien, die sich mit den Chancen und Risiken auseinandersetzen, oder in der Zivilgesellschaft, die kritische Debatten führt und innovative Projekte anstößt. Besonders im Bildungsbereich braucht es jetzt Pioniergeist und Kreativität, um die Potentiale für mehr Teilhabe, Gerechtigkeit und Nachhaltigkeit zu nutzen. Dafür müssen wir Räume schaffen, in denen wir experimentieren, diskutieren und lernen können. Wir brauchen Mut, um bewährte Routinen zu hinterfragen und neue Wege auszuprobieren. Wir benötigen kritische Reflexion, um die ethischen, sozialen und ökologischen Folgen von KI zu bedenken. Und wir brauchen

den Willen, Technologie so einzusetzen, dass sie den Menschen dient – und nicht umgekehrt.

Die Frage „In welcher Welt wollen wir leben?“ ist dabei unser Kompass. Sie erinnert uns daran, dass die Zukunft kein Schicksal ist, sondern das Ergebnis unserer heutigen Entscheidungen und Handlungen.

