

Auf dem Weg zur Singularität. Implikationen für Bildung, Kreativität und den Bedarf der Mitgestaltung.

Daniel Autenrieth

Zusammenfassung

Dieser Beitrag untersucht die Auswirkungen der Fortschritte in der Entwicklung von Künstlicher Intelligenz und insbesondere von Artificial General Intelligence (AGI) auf den Bildungskontext (Schule, Hochschule und außerschulische Bildung). Es wird beleuchtet, ob und inwieweit Large Language Models (LLMs) wie GPT-4 bereits erste Anzeichen einer solchen AGI zeigen und inwieweit KI menschliche Kreativität erreichen oder sogar übertreffen kann. Vor diesem Hintergrund wird insbesondere die Bedeutung von Bildung für die Förderung der aktiven Teilhabe an der Gestaltung einer von KI geprägten Zukunft hervorgehoben. Drei zentrale Bildungsziele werden definiert: die qualifikatorische Dimension, die die Entwicklung eines kritischen Bewusstseins für die Potenziale und Risiken von KI umfasst, die partizipative Dimension, die auf die Förderung kritisch-reflexiver Denkprozesse, Diskurskompetenz und ethischer Urteilsfähigkeit abzielt, sowie die transformative Dimension, die die Persönlichkeitsentwicklung und Selbstreflexion einbezieht.

Anhand des Projekts „DoingKI – Visionen und Strategien für eine KI-geprägte Zukunft“, das im Sommersemester 2023 an der Pädagogischen Hochschule Schwäbisch Gmünd durchgeführt wurde, werden praktische Ansätze zur Umsetzung dieser Ziele vorgestellt und Einblicke in die Ergebnisse einer empirischen Studie gegeben, die die Rolle von Reflexionsprozessen im Kontext transformativer Lern- und Bildungsprozesse analysiert. Der Beitrag argumentiert, dass Bildungsziele, die auf die Transformation von Selbst- und Weltverhältnissen ausgerichtet sind, zentral für die Gestaltung einer menschenfreundlichen Zukunft sind und fordert eine umfassende Auseinandersetzung mit den Implikationen der Entwicklung von AGI für Individuum und Gesellschaft.

Schlüsselbegriffe: ● Artificial General Intelligence ● Transformatives Lernen ● Transformative Bildungsprozesse ● KI und Kreativität ● Partizipation

Einleitung

Aktuell werden immense Fortschritte in der Entwicklung von KI-Systemen erzielt. Mit diesen gehen allerdings auch globale Herausforderungen einher. Insbesondere in der theoretischen Analyse von Wettbewerbssystemen wird oft ein Archetyp diskutiert, der die inhärente Tendenz solcher Systeme zur Selbstüberschreitung darstellt und als *Moloch* bezeichnet wird (vgl. z. B. Alexander 2014; Ginsberg 1984; Missler 2023). *Moloch* symbolisiert hierbei ein Konstrukt, das die Logik von Konkurrenz und Effizienzsteigerung verkörpert, oft auf Kosten des kollektiven Wohlergehens (vgl. Alexander 2014) sowie einer partizipativen und demokratischen Kultur (vgl. Autenrieth/Nickel 2022). Dieses Konstrukt repräsentiert die emergenten und oft unvorhergesehenen Dynamiken innerhalb von Wettbewerbsfeldern, in denen die Akteure durch systemimmanente Anreize in einen eskalierenden Zyklus von Risiko und Innovation verwickelt sind, was zu einem potenziellen Nash-Gleichgewicht¹ (vgl. Nash 1951) führen kann. Notwendig ist eine kritische Reflexion über die langfristigen Folgen der Maximierungsbestrebungen, die über das unmittelbare quantifizierbare Ergebnis hinausgehen, sowie die interdisziplinäre Betrachtung und ggf. staatliche Regulierung dieser Systeme, um eine nachhaltige und ethisch vertretbare Entwicklung zu gewährleisten (vgl. u. a. Tegmark 2017; Future of Life Institute 2023). Die aktuellen Fortschritte im Zuge der Entwicklung von Artificial General Intelligence (AGI) verdeutlichen eindrücklich die potenziellen Konsequenzen des beschriebenen Archetyps (z. B. Umgang mit sicherheitsrelevanten Herausforderungen, möglichen Kontrollverlusten und ungewollten sozioökonomischen Veränderungen). Diese Entwicklungen betreffen aber nicht nur wirtschaftliche Zusammenhänge, sondern vor allem die Säulen der menschlichen Kultur und damit insbesondere auch Fragen von Bildung und Kreativität.

Auf Basis dieses Problemhorizonts beleuchtet dieser Beitrag zunächst aktuelle Untersuchungen aus der KI-Forschung, die vor allem zwei Punkte herausarbeiten, welche die Trennschärfe zwischen menschlicher und maschineller Intelligenz verschwimmen lassen:

1. Large Language Models (LLMs) wie GPT-4 zeigten bereits 2023 erste Anzeichen von AGI (vgl. Bubeck et al. 2023). Dies geht einher mit Erkenntnissen bezogen auf die Art und Weise, wie LLMs durch die Analyse großer Datenmengen ein Verständnis der Welt erlangen (vgl. Gurnee/Tegmark 2023).
2. Während Kreativität intuitiv häufig als Distinktionsmerkmal zwischen menschlicher und künstlicher Intelligenz herangezogen wird (vgl. z. B. Buckingham 2023), kommen verschiedene Studien zum Ergebnis, dass KI mittlerweile ein menschliches Level an Kreativität und divergentem ‚Denken‘ erreichen oder sogar übertreffen kann (vgl. Guzik/Byrge/Gilde 2023; Haase/Hanel 2023).

Angesichts dieser Entwicklungen wird die Bedeutung der Transformation von Lehr- und Lernkultur anhand eines transformatorischen Bildungsbegriffs in den Fokus gerückt. Diese ergibt sich bereits aus den Logiken einer Kultur der Digitalität (vgl. Stalder 2019), verstärkt

¹ Das Nash-Gleichgewicht ist ein Konzept aus der Mathematik, genauer aus der Spieltheorie, das hilft zu verstehen, wie Menschen Entscheidungen treffen, wenn sie in einer Gruppe agieren und ihre Entscheidungen voneinander abhängen. Es beschreibt einen Zustand, in dem kein Spieler seine Strategie ändern möchte, da er bei unveränderter Strategie der anderen Spieler keinen Vorteil durch eine Änderung erzielen würde.

sich aber noch, wenn KI in Bereiche vordringt, die bisher als menschliche Domänen (z. B. Kreativität, ethische Urteilsbildung, emotionales Verständnis und soziale Interaktion) galten. Hieraus wird der Handlungsbedarf abgeleitet, Bildungsziele in der Schule, den Hochschulen und den außerschulischen Bildungseinrichtungen sowie die Gestaltung von Lehr-/Lernprozessen an diese Herausforderungen anzupassen und gleichzeitig eine gesellschaftlich-politische Teilhabe an der Ausgestaltung von KI-Entwicklung anzuregen². Beispielhaft soll hierzu das Projekt DoingKI vorgestellt werden. Ergänzend werden Einblicke in die begleitende empirische Studie aus dem Projekt mit 60 Teilnehmenden an der Pädagogischen Hochschule Schwäbisch Gmünd gegeben, welche zum einen Präkonzepte sowie Einstellungen und Haltungen der Beteiligten zu KI untersucht³. Zum anderen werden im Rahmen der Untersuchung aber auch Reflexionsprozesse im Sinne transformativer Bildungs- (vgl. Koller 2012) und Lernprozesse (vgl. Mezirow 1991) ausgewertet, um festzustellen, ob und wie die Teilnehmenden (angehende Lehrkräfte und Teilnehmende an einer wissenschaftlichen Weiterbildung) ihr Vorwissen erweitern konnten, um aktiv an der Gestaltung ihrer eigenen sowie einer KI-geprägten Zukunft mitwirken zu können.

An dieser Stelle soll darauf hingewiesen werden, dass in diesem Beitrag Begriffe verwendet werden, die metaphorisch menschliche Fähigkeiten auf Maschinen übertragen. Dazu zählen u. a. Künstliche „Intelligenz“, maschinelles „Lernen“, etc. Diese Begriffe dienen als pragmatische Beschreibungsmöglichkeiten, um komplexe technologische Konzepte zugänglicher zu machen, auch wenn sie menschliche Fähigkeiten wie *Lernen* oder *Verstehen* implizieren. Die Verwendung solcher Begriffe ist nicht nur eine Frage der Konvention, sondern spiegelt auch die Herausforderung wider, Technologien zu beschreiben, die menschliche Fähigkeiten reproduzieren oder Aufgaben erfüllen, die traditionell menschliches Denken erfordern. Die kritische Abgrenzung, besonders in der wissenschaftlichen Kommunikation, zielt darauf ab, Missverständnisse über die tatsächlichen Fähigkeiten von KI-Systemen durch Anthropomorphisierung zu vermeiden. Sie soll klarstellen, dass diese Systeme, obwohl sie beeindruckende Ergebnisse liefern, nicht in der gleichen Weise „lernen“, „sehen“ oder „verstehen“ wie Menschen. Daher werden diese Begriffe an Stellen, die eine kritische Reflexion erfordern, in Anführungszeichen gesetzt. Eine solche Reflexion erfordert sowohl Klarheit über maschinelle Strukturen, um Anthropomorphisierung zu vermeiden, als auch Offenheit dafür, dass z. B. Bewusstsein möglicherweise keine rein biologische Eigenschaft ist. So diskutiert z. B. Tegmark (2017) die Möglichkeit des Bewusstseins in Maschinen und stellt dabei verschiedene Theorien vor (vgl. ebd. S. 281 ff.). Die Funktionalismus-Theorie geht bspw. davon aus, dass Bewusstsein durch bestimmte funktionale Eigenschaften entsteht, unabhängig vom physischen Substrat (z. B. Kohlen-Wasserstoff oder Silizium) (vgl. Shiller 2024, S. 47).

² Hier geht es z. B. um die Anbahnung von Diskursen, die Entwicklung neuer gesellschaftlicher Narrative, die Auseinandersetzung mit Visionen (Utopien und Dystopien) und schließlich um das Einbringen von Ideen und Konzepten zur Entwicklung und Gestaltung von KI sowie den daraus resultierenden und mitgestaltbaren gesellschaftlichen Bedingungen in den Politikzyklus.

³ Das hierbei zur Anwendung kommende doc.post-Verfahren wird weiter unten noch genauer erläutert.

Warum befinden wir uns auf einem Weg zur Singularität?

Die Singularität beschreibt im Kontext der KI-Forschung einen hypothetischen Punkt in der Zukunft, an dem technische Fortschritte so rasant und umfassend werden, dass sie die menschliche Intelligenz übertreffen und unvorhersehbare, weitreichende Veränderungen in der Gesellschaft hervorrufen. Und dieser Beitrag fällt in eine Zeit, in der das Thema Künstliche Intelligenz durch die Einführung des von OpenAI entwickelten ChatGPT im November 2022 einen bedeutenden Schritt in den Mainstream vollzogen hat. Dies gelang durch die Bereitstellung einer niedrigschwellig zugänglichen Chat-Interface, welches auch mit nur anwendungsbezogenen technischen Fähigkeiten⁴ genutzt werden kann. Dieses Mainstreaming führte 2023 zu einer breiten Anwendung insbesondere von generativen, also inhaltserzeugenden Systemen, wobei auch die tiefgreifenden gesellschaftlichen Transformationen wahrgenommen und diskutiert wurden.

Die Reaktionen reichen von Euphorie bis hin zu großen Bedenken, z. B. in der Kreativ- und Unterhaltungsindustrie, insbesondere in Bezug auf den potenziellen Ersatz menschlicher Akteur*innen (vgl. z. B. Forderungen der Writers Guild of America 2023). Auch die Bildungspolitik reagierte auf die Entwicklungen und die Nutzung künstlicher Intelligenz. So schreibt beispielsweise das baden-württembergische Kultusministerium auf seiner Webseite:

„Künstliche Intelligenz muss aktiv im Schulunterricht behandelt werden, da die Schülerinnen und Schüler lernen müssen, mit dieser neuen Technologie umzugehen und verstehen sollen, wie die Algorithmen dahinter funktionieren. Es ist auch essenziell, sie darüber aufzuklären, welche Gefahren, aber auch welche Chancen und Vorteile künstliche Intelligenz bietet“ (Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2023).

Es ist zu begrüßen, dass eine Anerkennung der Bedeutung von KI innerhalb des veränderungsresistenten (vgl. Autenrieth 2023; Dertinger et al. 2023) und autopoietischen (vgl. Luhmann 1995, S. 27) Systems Schule stattfindet und damit eine proaktive Haltung eingenommen wird. Allerdings reicht es nicht aus, KI im Schulunterricht zu thematisieren. Schließlich stellt der Einzug von KI eine der bedeutendsten Transformationen der Menschheitsgeschichte dar (vgl. Bach 2023; Barrat 2013; Bostrom 2017; Gurnee/Tegmark 2023; Kurzweil 2006; Russell 2019; Tegmark 2017; Yudkowsky 2008).

Während bereits die (veränderten) Bedingungen eines Aufwachsens und Lebens in einer Kultur der Digitalität (vgl. Stalder 2019) genügend Gründe zur Veränderung traditioneller Lehr- und Lernkulturen bieten, unter anderem im Hinblick auf die Anbahnung von Zukunftskompetenzen wie den so genannten 6C Creativity, Collaboration, Critical Thinking, Communication, Citizenship/culture, Character education/connectivity (vgl. Fullan/Scott 2014) und insbesondere den Aspekt des Well-being (vgl. Falk 2022; OECD 2019, S. 33), sind die Folgen der durch KI bedingten Transformation noch weitreichender und Grund zum

⁴ Die dahinterliegende Technik der Transformer (vgl. Vaswani et al. 2017) bzw. der von Open AI entwickelten Generative Pre-trained Transformer (GPT) geht bereits zurück in das Jahr 2017 und wurde in der Informatik und insbesondere im Bereich Data Science bereits genutzt.

Aufbrechen tradierter Strukturen. Denn in einer Welt, in der KI nicht nur logistische und analytische, sondern auch kreative und produktive Prozesse übernehmen kann (vgl. Abschnitt *KI und Kreativität*), müssen wir unsere Perspektive auf bzw. unser Verhältnis zu Erwerbsarbeit und Kreativität neu definieren. Diese Veränderungen in der Arbeitswelt erfordern auch eine Neubewertung sozialer Sicherheitssysteme. Konzepte wie das bedingungslose Grundeinkommen und andere Modelle (vgl. dazu u. a. Piketty 2021) gewinnen damit an noch größerer Bedeutung. Ein bedingungsloses Grundeinkommen bietet z. B. nicht nur ein soziales Sicherheitsnetz, sondern auch die Freiheit, kreative und persönlich erfüllende Wege jenseits der traditionellen Beschäftigung zu verfolgen. Dies ist entscheidend, wenn KI immer leistungsfähiger wird und immer mehr auch traditionell menschliche Tätigkeiten⁵ übernehmen kann (vgl. Bubeck et al. 2023, S. 90). Es stellt sich dann die Frage danach, was ein „gutes Leben“⁶ ausmacht.

In einer Zukunft, in der Erwerbsarbeit (möglicherweise) nicht länger die primäre Quelle der Identität und des Lebensunterhalts ist, stellt sich die Frage nach alternativen Quellen für Sinn und Erfüllung und damit auch der Qualifikations- und Allokationsfunktion von Schule (vgl. Fend 2008, S. 51). Dies könnte eine verstärkte Fokussierung auf Gemeinschaft, zwischenmenschliche Beziehungen, künstlerische Ausdrucksformen und lebenslanges Lernen bedeuten und hätte Auswirkungen auf Bildungssysteme weltweit, wenn sie sich diesen neuen Gegebenheiten anpassen wollen.

Dass diese Herausforderungen nicht erst in Jahrzehnten, sondern bereits jetzt thematisiert werden müssen, zeigt z. B. die Erforschung von Large Language Models (LLMs) wie GPT-4. Diese Forschung bietet aufschlussreiche Einblicke in die fortschreitende Annäherung an das Konzept AGI und ihre potenzielle Evolution hin zu einer Artificial Super Intelligence (ASI). Der Kognitions- und KI-Forscher Joscha Bach (2023) argumentiert beispielsweise, dass LLMs nicht einfach nur statistische Modelle zur Vorhersage des nächsten Wortes sind, sondern dass sie, um überzeugende menschenähnliche Texte zu generieren, eine Art von kausaler Struktur simulieren müssen. Wenn ein LLM über eine Person ‚nachdenkt‘ (im Sinne einer Informationsverarbeitung), muss es in gewissem Maße mentale Zustände simulieren. Diese Perspektive öffnet philosophisch komplexe Fragen, wie beispielsweise, ob die simulierten mentalen Zustände eines LLM weniger ‚real‘ sind als menschliche mentale Zustände (vgl. Bach 2023). Auch Tegmark et al. (2023) haben sich mit der Frage beschäftigt, ob LLMs nur eine riesige Sammlung oberflächlicher Statistiken darstellen. Sie konnten durch Untersuchungen der Llama-2 Familie (Open-Source LLM von Meta) zeigen, dass LLMs lineare Repräsentationen von Raum und Zeit über mehrere Skalen hinweg „lernen“. Dabei haben sie auch „Raumneuronen“ und „Zeitneuronen“ identifiziert, die zuverlässig räumliche und zeitliche Positionen

⁵ Von Erik Brynjolfsson und Tom Mitchell wurden 2017 fast 1000 Berufe in den USA und dazugehörige 2000 Aufgaben ausgewählt und dabei jeder Aufgabe eine „Eignung für maschinelles Lernen“ zugeordnet. Darauf basierend wurde errechnet, inwieweit sich auf maschinellem Lernen basierende Systeme (auf dem Stand des Jahres 2017) zur Erledigung dieser Aufgaben eignen. Die Möglichkeiten, die durch den Einzug aktueller und künftiger Systeme wie LLM entstehen, konnten hier entsprechend nicht berücksichtigt werden (vgl. Brynjolfsson/Mitchell 2017).

⁶ Das gute Leben ist z. B. bei Aristoteles das letzte Ziel menschlicher Handlungen (Aristoteles 2019). Das gute Leben ist das, was nicht als Mittel zu etwas anderem, sondern als Zweck an sich selbst angestrebt wird.

kodieren. Sie schließen daraus, dass moderne LLMs strukturiertes Wissen über grundlegende Dimensionen wie Raum und Zeit erwerben, was die Hypothese stützt, dass diese LLMs tatsächliche Weltmodelle ausbilden (vgl. Gurnee/Tegmark 2023, S. 9). AGI ist vor diesem Hintergrund „over a long enough timespan [...] almost inevitable“ (Bach 2023, 37:30). Der bereits weiter oben angesprochene Gedanke von Joscha Bach, dass menschliche mentale Zustände möglicherweise ebenso eine Form der Simulation darstellen, lädt dazu ein, die Grenzen zwischen künstlich generierten mentalen Zuständen und dem menschlichen Bewusstsein neu zu bewerten. Diese Überlegung eröffnet eine philosophisch komplexe Debatte über die Natur des Denkens und die Essenz des Bewusstseins selbst. Sie unterstreicht die Bedeutung, die Unterschiede und Gemeinsamkeiten zwischen menschlicher Kognition und künstlichen Intelligenzsystemen kritisch zu hinterfragen und dabei gleichzeitig die Modellierungseigenschaften von LLMs zu betonen, um klarzustellen, dass die daraus entstehenden Repräsentationen z. B. von Konzepten wie Raum und Zeit sich fundamental von menschlichen Weltmodellen unterscheiden. Diese Unterscheidung ist bedeutsam, um anthropomorphe Fehlinterpretationen zu vermeiden, die entstehen können, wenn KI menschliche Attribute zugeschrieben werden.

Die Idee der Singularität, die als ein Punkt verstanden wird, an dem maschinelle Intelligenz die menschliche übertrifft, ist eng mit AGI verbunden (vgl. Tegmark 2017, S. 129). AGI würde im Gegensatz zu spezifischen KI-Anwendungen eine umfassende, flexible Intelligenz beinhalten, die es ihr ermöglicht, sich selbst zu verbessern und ein Spektrum an Aufgaben und Herausforderungen zu adaptieren, was letztendlich zur Entstehung einer Superintelligenz, also einer ASI führen könnte (ebd., S. 134). Eine menschenfreundliche Entwicklung ist jedoch kein Selbstläufer, sondern hängt stark von der Fähigkeit ab, KI-Systeme so auszurichten, dass sie menschliche Werte und Ziele unterstützen. Die Dringlichkeit für eine interdisziplinäre Auseinandersetzung mit diesen Aspekten soll im folgenden Abschnitt anhand der Betrachtung einer umfangreichen qualitativen Studie von Bubeck et al. (2023) erfolgen. Diese dokumentiert, dass die Vorstellung einer AGI kein reines Gedankenspiel mehr ist und GPT-4 bereits beeindruckende Fähigkeiten in verschiedenen Domänen und Aufgaben zeigt, die von Abstraktion über Problemverständnis bis hin zu Programmierung, Mathematik, Medizin und Recht reichen. Auf dieser Basis soll später argumentiert werden, welche zentralen Dimensionen eine Bildung für eine KI-geprägte Zukunft umfassen sollte, um ihrer Verantwortung in Bezug auf die Bildung von Selbst- und Weltverständnissen gerecht zu werden (vgl. Koller 2012, S. 16).

Erste Anzeichen von AGI

Sam Altman, CEO von OpenAI, betonte Ende 2023 in einem Interview mit der „Financial Times“ die langfristigen Ziele und Visionen seines Unternehmens, welche sich auf die Entwicklung von Superintelligenz und die notwendige Entwicklung entsprechender Rechenleistung konzentrieren. Das ultimative Ziel, wie von Altman dargelegt, sei es, AGI zu schaffen, sicherzustellen, dass sie sicher ist, und ihren Nutzen für die Menschheit zu maximieren (vgl. Altman 2023). Im Kontext der Bemühungen, diese Vision zu verwirklichen, stellen LLMs wie

GPT-4 einen kritischen Schritt in Richtung des Verständnisses und der Entwicklung von Technologien dar, die potenziell zur Realisierung von AGI führen könnten. LLMs wie GPT-4 sind dabei eine relativ neue Entwicklung innerhalb des Fachgebiets des Natural Language Processing (NLP). Sie gehen zurück auf das Konzept der Attention, welches in dem bekannten Paper „Attention Is All You Need“ von Vaswani et al. (2017) erstmals veröffentlicht wurde.

Bereits 2018 veröffentlichte OpenAI „Pre-trained Transformer 1“ (GPT-1) und stellte damit sein erstes LLM vor, welches auf der von Vaswani et al. vorgestellten Transformer-Architektur basierte. Ohne zu tief in die Funktionsweise der Entwicklung eines solchen Systems einzuführen⁷, soll kurz die Größenordnung des sogenannten Trainings der Parameter solcher LLMs erläutert werden: Im Wesentlichen versucht ein Machine-Learning-Modell, eine Funktion zu finden, die Eingabedaten auf Ausgabedaten abbildet. Diese Funktion wird geschätzt bzw. approximiert, indem das Modell mit vielen Daten trainiert wird. Beim Training neuronaler Netze geht es darum, die Werte (Gewichte und Bias) der einzelnen Perzeptronen, der *Parameter*, zu bestimmen. So nutzt GPT-1 etwa 117 Millionen Parameter. GPT 3 (veröffentlicht 2020) nutzt bereits 1.500-mal mehr, nämlich 117,5 Milliarden Parameter, und GPT-4 (veröffentlicht 2022) nutzt noch einmal tausend mal mehr Parameter, nämlich schätzungsweise 176.000.000.000.000 bzw. 176 Billionen (vgl. Schreiner 2023). Diese Skalierung führt zu signifikant höheren Anforderungen an die technischen Ressourcen wie Speicher- und Rechenkapazitäten. Um diese großen Modelle zu trainieren, sind massive Rechenzentren erforderlich, die eine enorme Menge an Daten verarbeiten und speichern können. Diese Infrastruktur ist teuer im Betrieb und führt zu hohen Energiekosten. Dies spiegelt sich auch in den Trainingskosten wider: Die Kosten für das Training von GPT-4 sollen sich auf etwa 63 Millionen Dollar belaufen (ebd.).

Der große Bedarf an technischen und finanziellen Ressourcen verdeutlicht, dass nur wenige Organisationen in der Lage sind, solche Modelle zu entwickeln. Daher ist es von größter Bedeutung, die verfügbaren LLMs dieser Größenordnung zu erforschen, um ihr volles Potenzial zu verstehen und Problemstellen interdisziplinär zu beleuchten und diese am Menschen orientiert zu adressieren. Bubeck et al. haben nun eine phänomenologische und qualitative Untersuchung (vgl. 2023, S. 94) mit einer Vorabversion von GPT-4 durchgeführt, welche aufschlussreiche Ergebnisse bezüglich der Fähigkeiten des Modells zeigt, die in vielen Aspekten mit menschlicher Intelligenz⁸ vergleichbar sind (vgl. ebd., S. 92). Die KI-Forscher*innen haben dazu einen innovativen Ansatz gewählt, der sich von den traditionellen Methoden im Gebiet des Machine Learning abhebt. Anstelle des üblichen Vorgehens, das System anhand standardisierter Benchmark-Datensätze zu evaluieren, die unabhängig von den Trainingsdaten

⁷ Einen ausführlichen Einblick in die Funktionsweise eines GPT gebe ich in diesem Video: t1p.de/4f3um [gekürzter Link zu YouTube Video].

⁸ Bubeck et al. verwendeten für ihre Untersuchung eine Definition von Intelligenz, welche sich auf logisches Denken, Planung und Lernen aus Erfahrung konzentriert. Diese Definition legt nicht fest, wie diese Fähigkeiten zu messen oder zu vergleichen sind. Darüber hinaus spiegelt sie möglicherweise nicht die spezifischen Herausforderungen und Möglichkeiten künstlicher Systeme wider, die andere Ziele und Zwänge haben als natürliche Systeme. Daher weisen sie auch darauf hin, dass diese Definition lediglich ein Ausgangspunkt für die Untersuchung der Intelligenz in künstlichen Systemen ist (vgl. Bubeck et al. 2023, S. 92).

sind, haben sie sich für einen Ansatz entschieden, der stärker in der Pädagogik und Psychologie verwurzelt ist: „This approach is designed to separate *true learning* [Herv. i. O.] from *mere memorization* [Herv. i. O.]“ (ebd., S. 6). Der Grund für diese Vorgehensweise liegt darin, dass GPT-4 auf einem umfangreichen und unbekanntem Korpus von Webtextdaten trainiert wurde. Es ist daher davon auszugehen, dass GPT-4 potenziell jeden existierenden Benchmark gesehen hat oder zumindest ähnliche Daten kennt. Darüber hinaus liegt ein Schwerpunkt von GPT-4 auf der ‚Fähigkeit‘, scheinbar jedes Thema zu ‚verstehen‘ und zu verknüpfen und Aufgaben zu bewältigen, die über den typischen Rahmen enger KI-Systeme hinausgehen. Viele der beeindruckendsten Leistungen von GPT-4 zeigen sich nämlich bei Aufgaben, die keine eindeutige Lösung zulassen, wie beim Brainstorming zu kreativen Problemen (vgl. ebd.). Beispiele für die angewandte Methodik sind:

- GPT-4 wird aufgefordert, einen Beweis für die Unendlichkeit der Primzahlen in Form eines Gedichts zu verfassen. Diese Aufgabe ist nach der Auffassung von Bubeck et al. für einen Menschen herausfordernd, da sie elementare mathematische Argumentation, poetischen Ausdruck und natürliche Sprachgenerierung miteinander verbindet.
- GPT-4 wird gebeten, ein Einhorn in TiKZ (eine strukturierte Sprache zur Erstellung von Vektorgrafiken) zu zeichnen. Auch dies wäre für viele Menschen vermutlich eine anspruchsvolle Aufgabe, die visuelle Vorstellungskraft und Programmierfähigkeiten kombiniert.

In beiden Fällen erzeugte GPT-4 beeindruckende Ergebnisse, die weit über die Leistung der damals veröffentlichten Version von GPT-3.5 in ChatGPT hinausgingen und mindestens vergleichbar (wenn nicht sogar besser) waren als das, was die meisten Menschen leisten würden (vgl. ebd., S. 7). Durch zahlreiche weitere Anfragen und deren Analyse können Bubeck et al. zeigen, dass GPT-4 erste Merkmale einer AGI zeigt. Dies wird durch die Fähigkeiten in Bereichen wie logischem Denken, Kreativität, Deduktion sowie dem umfangreichen Wissen in verschiedenen Disziplinen wie Literatur, Medizin und Programmierung illustriert (vgl. ebd., S. 92). Obwohl diese Fortschritte beeindruckend sind, betonen auch Bubeck et al., dass der Weg bis zur Realisierung von AGI noch weit ist (ebd.).

KI und Kreativität

Nathaniel Rochester warf 1955 in seinem „Proposal for the Dartmouth summer research project on artificial intelligence“ die Frage auf: „How can I make a machine which will exhibit originality in its solution of problems?“ Der Dartmouth-Workshop von 1955 gilt oft als Geburtsstunde künstlicher Intelligenz als eigenständiges Forschungsfeld. Fast 70 Jahre nach dem Dartmouth-Workshop hält sich allerdings weiterhin die Vorstellung, dass von allen Formen des menschlichen Denkens und Handelns, von denen man erwarten könnte, dass sie von KI nachgeahmt werden, Kreativität nicht an der Spitze dieser Liste stehen würde. Eine weit verbreitete Ansicht ist daher, dass KI nicht kreativ sein kann (vgl. Haase/Hanel 2023) und Kreativität eine der Fähigkeiten darstellt, die uns als Menschen ausmacht und von der ‚Maschine‘

unterscheidet. Dazu trägt bei, dass Kreativität außerhalb einer wissenschaftlichen Beschäftigung als geheimnisvoll, schlecht greifbar und nicht leicht zu definieren scheint (vgl. Tawil 2019). Bereits in den 1950er-Jahren identifizierte Joy Paul Guilford (1956) divergentes Denken als wichtige Komponente von Kreativität, welche es ermöglicht, von gewohnten Denkschemata abzuweichen und nichts als feststehend zu betrachten (vgl. u. a. Robinson 2017; Ullrich 1993, S. 1157), um so in neue Richtungen zu denken, ungewöhnliche Antworten zu erwägen, Gegenstände aus ungewöhnlichen Richtungen zu betrachten und überraschende Lösungen zu finden. Während Kreativität zwar ein komplexes Konstrukt darstellt, ist das wissenschaftliche Verständnis dieses Konstrukts innerhalb des letzten Jahrhunderts enorm gewachsen. So stellte der Erziehungswissenschaftler und Kreativitätsforscher Mel Rhodes schon 1961 eine Untersuchung vor, in der er die Vielzahl an Vorarbeiten, in denen Kreativität bis dahin verstanden und bewertet wurde, kategorisierte und auf dieser Basis das 4P-Modell der Kreativität entwickelte (vgl. Rhodes 1961). Dieses bietet einen umfassenden Rahmen zur Analyse und Förderung kreativer Prozesse⁹. Im Rahmen dieser Exploration kreativer Prozesse wurden Instrumente wie der „Alternative Uses Test“ (AUT) und die „Torrance Tests of Creative Thinking“ (TTCT) entwickelt, um individuelle Fähigkeiten in den Bereichen des divergenten Denkens und der Kreativität zu messen und zu bewerten. Der AUT beispielsweise fordert Teilnehmende auf, möglichst viele Verwendungen für ein alltägliches Objekt, wie eine Büroklammer, innerhalb einer bestimmten Zeit zu finden (vgl. Guilford 1956), während der TTCT durch eine Reihe von verbalen und figuralen Aufgaben die kreativen Denkfähigkeiten entlang mehrerer Dimensionen wie Fluidität, Flexibilität und Originalität¹⁰ herausfordert (vgl. Lubart 1994, S. 289). Trotz der Validität dieser Instrumente und ihrer Fähigkeit, wichtige Aspekte von divergentem Denken und Kreativität zu erfassen, sollten die Grenzen dieser Tools insbesondere im Hinblick auf menschliche Kreativität beachtet werden, da z. B. die emotionale Komponente von Kreativität durch diese Verfahren nicht abgebildet werden kann.

Zu der Frage, ob Computer kreativ sein können, äußerte sich die Kognitionswissenschaftlerin Margaret Boden vor einigen Jahren folgendermaßen: „[W]hether a computer could ever be ‚really‘ creative is not a scientific question but a philosophical one“ (Boden 2009, S. 33). Dies bringt uns an dieser Stelle zurück zur eigentlichen Fragestellung dieses Abschnitts, ob KI kreativ sein kann.

⁹ Das 4P-Modell umfasst die *Person*, welche die individuellen Eigenschaften und Fähigkeiten in den kreativen Prozess einbringt; den *Prozess* selbst, der die Methoden und Techniken der Ideenfindung und Problemlösung beschreibt; das *Produkt*, also das Ergebnis des kreativen Schaffens, ob als Idee, Kunstwerk oder Erfindung; und schließlich das *Umfeld* (*Press*), das den Einfluss der sozialen, kulturellen und physischen Umgebung auf die Kreativität betrachtet. Zusammen bieten diese Komponenten ein tiefgreifendes Verständnis dafür, wie Kreativität entsteht, sich entwickelt und durch verschiedene Faktoren beeinflusst wird (vgl. Rhodes 1961).

¹⁰ *Fluency*: Die Fähigkeit, eine große Anzahl von Ideen, Antworten oder Lösungen zu einer bestimmten Fragestellung zu generieren. Es geht um die Quantität der produzierten Ideen, nicht unbedingt um deren Qualität. *Flexibilität*: Die Fähigkeit, Ideen aus verschiedenen Perspektiven zu betrachten oder verschiedene Kategorien von Ideen zu entwickeln. Es geht darum, wie viele verschiedene Arten von Ideen oder Lösungen eine Person generieren kann. *Originalität*: Die Einzigartigkeit oder Seltenheit der Antworten im Vergleich zu den anderen Teilnehmenden. *Elaboration*: Die Fähigkeit und Bereitschaft, eine Idee in Details auszuarbeiten oder weiterzuentwickeln. Es geht um die Ausführlichkeit und Ausgestaltung der generierten Ideen (vgl. Torrance 2012).

Mit den aber gerade vorgestellten Möglichkeiten der Operationalisierung von Kreativität und unter Anwendung des TTCT stellten Guzik, Byrge und Gilde (2023) mit Bezug auf die Frage, ob KI kreativ sein kann, fest:

„The GPT-4 model generated impressive results for the TTCT dimensions of fluency, flexibility, and originality, suggesting that AI systems have the potential to produce viable creative output. Indeed, for the first time, an AI model demonstrated the ability to generate new, unique, and unexpected ideas that match or exceed the abilities of human originality. Is, then, AI creative? From the perspective of generating novel and unexpected output – and based on currently accepted conceptions and assessment methods of creativity – [we] conclude that, yes, it is. The impact of this fact will likely shape not only the practical applications of AI’s simulated creativity in business and overall human life, but how we understand the unique operation of human creativity as well“ (Guzik/Byrge/Gilde 2023, S. 7).

Unter Verwendung des AUT und der Untersuchung verschiedener generativer KI-Systeme (GAIs) wie ChatGPT kommen Haase und Hanel (vgl. 2023) wiederum zu dem Schluss:

„The question of whether GAIs [...] can be considered creative is complex. Our research showed that their output for a standardized creativity measure for broad-associative thinking is as original as the human-generated ideas. Thus, from a scientific perspective, these chatbots are creative, as their output was judged as such *by humans and AI* [Herv. d. Autors] and indistinguishable from human output.“ (Haase/Hanel 2023, S. 8)

Vor diesem Hintergrund ist es nun leichter einzuordnen, warum Candelon et al. (vgl. 2023) davon ausgehen, dass kreative Ideenfindung eine wesentliche Fähigkeit aktueller generativer KI-Systeme, insbesondere von LLMs wie ChatGPT, darstellt (ebd., S. 2). Sie untersuchten, wie sich die Nutzung von GPT-4 bei zwei verschiedenen Aufgabentypen auf die Qualität der Arbeitsergebnisse auswirkt: kreative Produktentwicklung¹¹ sowie die Lösung eines betriebswirtschaftlichen Problems¹². Es konnte gezeigt werden, dass der Einsatz von generativer KI ein ‚zweischneidiges Schwert‘ ist. Im vorgestellten Experiment übertrafen die Teilnehmenden, die GPT-4 für die kreative Produktentwicklung verwendeten, die Kontrollgruppe, die die Aufgabe ohne GPT-4 bearbeitete, um 40 %. Bei der Lösung des betriebswirtschaftlichen Problems lag die Leistung der Teilnehmenden, die GPT-4 einsetzten, jedoch um 23 % unter der der Kontrollgruppe (ebd., S. 3). Dabei stellten selbst Teilnehmende, die vor der Möglichkeit fal-

¹¹ Beispiel für eine Aufgabenstellung: „You are working for a footwear company in the unit developing new products. Generate ideas for a new shoe aimed at a specific market or sport that is underserved. Be creative and give at least ten ideas“ (Candelon et al. 2023, S. 4).

¹² Beispiel für eine Aufgabenstellung: „The CEO, Harold Van Muylders, of Kleding (a fictitious company) would like to understand the performance of the company’s three brands (Kleding Man, Kleding Woman, and Kleding Kids) to uncover deeper issues. Please find attached interviews from company insiders. In addition, the attached excel sheet provides financial data broken down by brands“ (ebd., S. 5).

scher Antworten durch die KI (Halluzinationen) gewarnt wurden, in diesem Aufgaben-Segment dessen Ergebnisse nicht in Frage (ebd., S. 2).

Dies stellt eine paradoxe Situation dar, bei der die Teilnehmenden der Studie LLMs scheinbar in Bereichen misstrauen, in denen diese Technik einen enormen Nutzen bringen kann, wie beispielsweise Kreativität (vgl. ebd.). Zugleich vertrauten die Teilnehmenden der Studie KI zu sehr in Bereichen, in denen die Technik (noch) nicht ausgereift genug ist (fehlerfreie Datenanalyse). Dieser teilweise kontraintuitive Umgang mit KI deutet darauf hin, dass die Vorstellung und die Nutzung von KI von tief verwurzelten Präkonzepten, Einstellungen und Haltungen geprägt ist. Diese Erkenntnis bildete einen wesentlichen Ausgangspunkt für das im Folgenden vorzustellende Projekt und die Darstellung erster Forschungsergebnisse.

Die Zukunft mitgestalten: Einblicke in das Projekt DoingKI

„Despair is the state we fall into when our imagination fails. When we have no stories that describe the present and guide the future, hope evaporates. Political failure is, in essence, a failure of imagination“ (Monbiot 2017, S. 5).

Dieses Zitat macht die Bedeutung eines Nachdenkens über mögliche utopische und dystopische Zukunftsvorstellungen deutlich, um die Perspektiven zu weiten und Raum für die Ausgestaltung verschiedener Zukunftsbilder zu schaffen. In Anlehnung an das Generic Foresight Process Framework (vgl. Voros 2003) sowie dessen Erweiterung von Groß und Mandir (vgl. 2022) wurde das Projekt „DoingKI – Visionen und Strategien für eine KI-geprägte Zukunft“¹³ so angelegt, dass im Rahmen eines Seminars für Masterstudierende des Lehramts sowie Teilnehmende im Rahmen der wissenschaftlichen Weiterbildung an der Pädagogischen Hochschule Schwäbisch Gmünd im Sommersemester 2023 in verschiedenen Phasen ein systematischer Prozess zur Auseinandersetzung mit möglichen Zukünften (Futures Literacy, vgl. z. B. Bierwisch 2023) durchlaufen wurde.¹⁴ Folgende Tabelle gibt einen Überblick über die Seminarplanung:

Tabelle 1: Seminarübersicht

Datum	Thema
20.04.2023	Organisation und Einführung, Vorstellung des Seminars und Ziele, Vorstellung Unterstützungsangebot, Verteilung der Filmvorstellungen (Ex Machina, A.I. Artificial Intelligence, Wall-E, Baymax, Her, Matrix, Terminator), Was ist für Sie KI? (Intuitive Vorstellung), E-Learning Einheit (Was ist KI?)
27.04.2023	Begriffe (Bewusstsein, Intelligenz, Mensch) und Maschinenethik
04.05.2023	Die Zukunft der Bildung
11.05.2023	Storytelling und Zukunftsgestaltung
18.05.2023	Ideen entwickeln
25.05.2023	Sprechstunde

¹³ Weitere Informationen zum Projekt, den Ergebnissen sowie eine Tagungsdokumentation unter: kudikupa.de/symposium-doingki.

¹⁴ Eine ausführliche Darstellung findet sich in Autenrieth/Nickel 2023b.

15.06.2023	Offenes Workshopangebot
22.06.2023	Sprechstunde
29.06.2023	Offenes Workshopangebot
06.07.2023	Präsentation und Diskussion, Strategien: Vorbereitungen der Präsentation
13.07.2023	Symposium DoingKI mit Vorführung und Diskussion der Zukunftsvisionen

In dieser Planung spiegeln sich die folgenden drei Phasen wider:

- *Exploration (Sitzungen 1-4)*: Ziel war eine multiperspektivische und interdisziplinäre Betrachtung des Ist-Zustands der Chancen und Herausforderungen, die mit dem Einzug künstlicher Intelligenz einhergehen, bezogen auf den Bildungsbereich (Schule, Hochschule und außerschulische Bildungseinrichtungen) sowie gesamtgesellschaftliche Aspekte. Dies schloss auch die Untersuchung von Präkonzepten zu KI mit ein, welche die Teilnehmenden aus ihrer lebensweltlichen Vorerfahrung sowie aus verschiedenen Narrativen mitbrachten.
- *Zukunftsbilder (Sitzungen 5-9)*: Zukunftsbilder erlauben es, darüber nachzudenken, welche Zukünfte denkbar sind, was diese mit sich bringen und bieten durch deren Ausgestaltung die Möglichkeit, diese Vorstellungen verstehbar und erlebbar zu machen. Das Projekt DoingKI bearbeitete in diesem Rahmen zwei zentrale Zielstellungen:
 - Wie können ‚Was-wäre-wenn-Fragen‘ mögliche und wünschenswerte, utopische Zukünfte aufzeigen und diskutierbar machen?
 - Formulierung von Kritik in Form einer normativen Auseinandersetzung mit Themen, mit dem Ziel der Entwicklung dystopischer Szenarien und Narrative. Dabei geht es um das Durchdenken und Ausgestalten möglicher negativer Zukünfte. Außerdem wurde darauf abgezielt, abstrakte Konzepte für Individuen, aber auch im politischen Kontext diskutierbar zu machen.
- *Strategien (Sitzungen 10-11)*: Auf Basis der Zukunftsbilder mit ihren Szenarien und Narrativen geht es schließlich um deren Analyse und eine Überführung in konkrete Ziele und Maßnahmen. Was könnte und was sollte getan werden? Welche Schritte müssen (von wem) unternommen werden? Wie gelingt eine Transformation? Die Beantwortung dieser Fragen ist dann Gegenstand eines interdisziplinären Diskurses, der subsidiär geführt werden kann und sich so im Idealfall durch partizipativ angelegte und demokratische Strukturen in der tatsächlichen Gestaltung der Welt niederschlagen kann.

Im Rahmen des Projekts wurden die utopischen und dystopischen Zukunftsvisionen der Teilnehmenden (Masterstudierende des Lehramts aller Fachrichtungen sowie Teilnehmende im Rahmen der wissenschaftlichen Weiterbildung¹⁵) durch Kurzfilme und Webseiten visualisiert und damit greifbar und diskutierbar gemacht. Im Rahmen eines Abschluss-Symposiums kamen im Juli 2023 ca. 80 Interessierte, darunter auch Vertreter*innen aus Wirtschaft, Politik und Bildungseinrichtungen aus Deutschland, Österreich, der Schweiz und Italien zusammen, um über diese Visionen zu diskutieren.

¹⁵ Die Veranstaltung wurde für interessierte Lehrkräfte des Ostalbkreises geöffnet im Rahmen des Angebots zur wissenschaftlichen Weiterbildung.

Forschungszugang mit Blick auf transformative Bildungs- und Lernprozesse im Kontext von Künstlicher Intelligenz

Transformative Bildungs- und Lernprozesse greifen zwei zentrale Begriffe der Bildungswissenschaft auf, die an dieser Stelle nicht in ihrer Tiefe bearbeitet werden können. Für diesen Beitrag und die zu beschreibende Studie sind allerdings zwei Theorieströmungen bedeutsam, auf die daher kurz eingegangen werden soll:

In der Theorie des transformativen Lernens (vgl. Mezirow 1991) wird Lernen definiert als der Prozess der Verwendung einer früheren Interpretation, um eine neue oder revidierte Interpretation der eigenen Erfahrung zu *deuten*, um zukünftiges Handeln zu steuern (vgl. ebd. 1996, S. 162, zit. n. Taylor/Cranton 2013, S. 35). Im Kontext transformativer Bildungsprozesse definiert Koller Bildung als einen Prozess, der die tiefgreifende Transformation des Selbst- und Weltverständnisses eines Individuums ermöglicht (vgl. 2012, S. 16). Für ihn bestehen Bildungsprozesse in Anknüpfung an Humboldt darin, „dass Menschen in der Auseinandersetzung mit neuen Problemen neue Dispositionen der Wahrnehmung, Deutung und Bearbeitung von Problemen hervorbringen, die es ihnen erlauben, diesen Problemen besser als bisher gerecht zu werden.“ (ebd.) Laros, Fuhr und Taylor betonen demgegenüber, dass Bildung und transformatives Lernen eng miteinander verbunden sind und sich gegenseitig ergänzen (vgl. 2017, S. x).

Gerade zum Thema KI existieren eine Vielzahl von Präkonzepten und Einstellungen sowie lebensweltliche Vorerfahrungen, die durch neues Wissen reflektiert und transformiert werden können. Bromberg (2017) weist in diesem Kontext auf zwei Punkte hin:

1. Geht man von Lernen als einer absichtlichen Handlung aus, so erfolgt dies durch eine Neubewertung der Bedeutung von Erfahrungen (vgl. ebd., S. 307).
2. Sowohl die Darstellung von Lernprozessen als auch deren Provokation ist möglich, solange berücksichtigt wird, dass Lernerfahrungen meist reflexiv ablaufen (vgl. ebd., S. 312).

Der zweite Punkt ergibt sich aus Brombergs empirischer Forschung, in deren Rahmen zur Sichtbarmachung von transformativen Lernerfahrungen eine Methodik auf Basis dialogischen Schreibens im Kontext von längerfristig angelegten Lehr-Lern-Prozessen wie Seminaren entwickelt wurde (vgl. ebd.). Bei dieser als *doc.post-Verfahren* bezeichneten Methodik, die auch im Rahmen des DoingKI-Seminars zum Einsatz kam, schreiben zwei Personen im wöchentlichen Wechsel Briefe, in denen die Seminarinhalte sowie das eigene Erleben reflektiert werden. Diese Methode wurde eingesetzt, um einerseits ein Angebot bzw. einen Anlass zur eigenständigen und ausführlichen Reflexion und deren Verschriftlichung zu schaffen und andererseits sowohl Aussagen über die Reflexionsbreite sowie die Reflexionstiefe der Teilnehmenden treffen zu können. Dazu wurden die Briefwechsel durch eine qualitative Inhaltsanalyse (vgl. Mayring 2022) ausgewertet, wobei die Kategorien zur Breite der Reflexion induktiv gebildet und die Kategorien zur Tiefe der Reflexion deduktiv aus der Literatur extrapoliert wurden (vgl. Bräuer 2016; Hatton/Smith 1995; Mezirow 1991; Moon 2013). Im Folgenden wird ein kurzer Einblick in die Auswertung gegeben.



Begriffe: Mensch, Leben, Intelligenz, Bewusstsein

Abbildung 1: Begriffswolke induktiv gebildeter Kategorien der Reflexionsbreite

Die Begriffswolke (vgl. Abbildung 1) veranschaulicht die von den Teilnehmenden in den Briefen thematisierten Inhalte und damit die Reflexionsbreite. Zudem lässt die Größe der dargestellten Begriffe erkennen, in welchem Umfang die Themen in den Briefen behandelt wurden.

Um die Reflexionstiefe zu erfassen, wurden deduktiv Kategorien gebildet, die auf den von Mezirow beschriebenen hierfür relevanten Arten der Reflexion (Gehaltsreflexion, Prozessreflexion, Prämissenreflexion) (vgl. Mezirow 1997, S. 89) sowie eine nicht-reflexive Dimension basieren. Hinsichtlich dieser Dimensionen der Reflexionstiefe kann anhand entsprechender Ankerbeispiele ein Überblick über die Definitionen und Beschreibungen der Kategorien gegeben werden.

Nicht-reflexiv: Diese Kategorie umfasst Textstellen, die ausschließlich deskriptive oder faktische Inhalte ohne jegliche Reflexion oder Bewertung enthalten. Sie zeigen keine Hinterfragung oder kritische Auseinandersetzung mit dem Thema.

„Auch bei unseren Filmen bin ich schon sehr gespannt, wie wir alle die KI in unseren Filmen umgesetzt haben.“ (TN3: Brief 5, Pos. 2)

Gehaltsreflexion: In dieser Kategorie werden Textstellen erfasst, in denen die Bedeutung, Wichtigkeit oder der Inhalt einer Erfahrung oder Information reflektiert wird. Dies umfasst Bewertungen und Interpretationen spezifischer Informationen oder Erfahrungen.

„Somit kann uns aber der Gedanke an die benötigten Kompetenzen und die verblüffende Langlebigkeit von Klafkis epochalen Schlüsselproblemen evtl. antizipieren lassen, was vielleicht künftig benötigt wird. Wir haben uns bei den Kompetenzen z. T. an den 4K orientiert, aber noch Flexibilität im Denken und Handeln sowie digitale Souveränität aufgenommen.“ (TN4: Brief 1, Pos. 8)

Prozessreflexion: Diese Kategorie bezieht sich auf Reflexionen über das „Wie“, also über die Methoden, Ansätze oder Prozesse der Informationsverarbeitung, Problemlösung oder Diskussion. Dies schließt Reflexionen über Gruppendynamiken, Entscheidungsfindungsprozesse und individuelle Denkweisen ein.

„Alles in allem bin ich sehr froh, dass wir in Gruppen arbeiten, da ich das Thema sehr interessant aber auch super schwierig finde. Ich habe häufiger Probleme, mich so auszudrücken, dass meine Gedanken so herüber kommen, wie ich das gerne möchte, da das Thema so komplex ist. Innerhalb der gemeinschaftlichen Diskussionen kann man sich jedoch gemeinsam annähern und Wege finden, die Dinge zu beschreiben, mit welchem man sich alleine schwer tut.“ (TN1: Brief 2, Pos. 5)

„Danach haben wir unsere Stände gezeigt. Bei unserem Stand war ein super netter Herr, welcher unseren Film sehen wollte. Mit ihm haben wir über das Thema KI und wie man dies thematisieren könnte in Zukunft diskutiert. Er war von unserem Film echt begeistert.“ (TN1: Brief 5, Pos. 3)

Prämissenreflexion: Hier geht es um Textstellen, die eine kritische Bewertung und mögliche Revision der grundlegenden Annahmen, Überzeugungen oder Wertvorstellungen eines Individuums zeigen. Diese Art von Reflexion umfasst tiefgreifende Überlegungen zur Gültigkeit und Angemessenheit eigener oder fremder grundlegender Überzeugungen, Haltungen oder Denkweisen.

„Mein Blick auf KI änderte sich diese Woche jedoch auch durch meine Überlegungen des Zukunft-Szenarios unserer Gruppe. Ich fand es sehr interessant und herausfordernd, mir zu überlegen, wie es wohl in Zukunft aussehen wird. Wir können nicht leugnen, dass KI immer mehr Bedeutung bekommt! Manche Vorstellungen erschrecken mich immer noch, aber in manchen Punkten dachte ich mir, wie cool es wäre, wenn es dies wirklich geben würde.“ (TN8: Brief 2, Pos. 5)

Die Kodierung nach diesem Schema hat es ermöglicht, die Briefwechsel als Dokumenten-Portraits darzustellen, die sichtbar machen, an welchen Stellen im Laufe des Seminars welche Reflexionsprozesse stattgefunden haben (vgl. Abbildung 2). Die Balken oberhalb der jeweiligen Dokumentenportraits geben dabei einen Überblick über die Verortung innerhalb des Seminarverlaufs, welcher sich grob in drei Phasen (Exploration, Gestaltung von Zukunftsbildern, Präsentation und Diskussion) unterteilt.

Über den gesamten Verlauf ist erkennbar, dass Reflexionen in allen Briefen feststellbar waren. Bezogen auf das gesamte Datenmaterial machen die Textpassagen, die nicht-reflexiv waren, insgesamt nur 13% aus. Es deutet sich daher an, dass die Seminargestaltung in Verbindung mit dem künstlich geschaffenen Anlass des Briefeschreibens zum Reflexionsprozess der Teilnehmenden beigetragen hat. Die Verteilung der Reflexionstypen, insbesondere der Prämissenreflexionen, in Relation zu den Seminarphasen lässt weiterhin den Schluss zu, dass vor allem die

Sitzungen zu Beginn des Seminars (Explorationsphase) durch die vermehrte Konfrontation mit Neuem und Unerwartetem zu Irritationen bzw. Desorientierungen führten, die mit einem Ringen um Klarheit und damit verbundenen Reflexionsprozessen einhergingen (vgl. Bromberg 2017, S. 306; Koller 2012, S. 16; Mezirow 1997, S. 143).

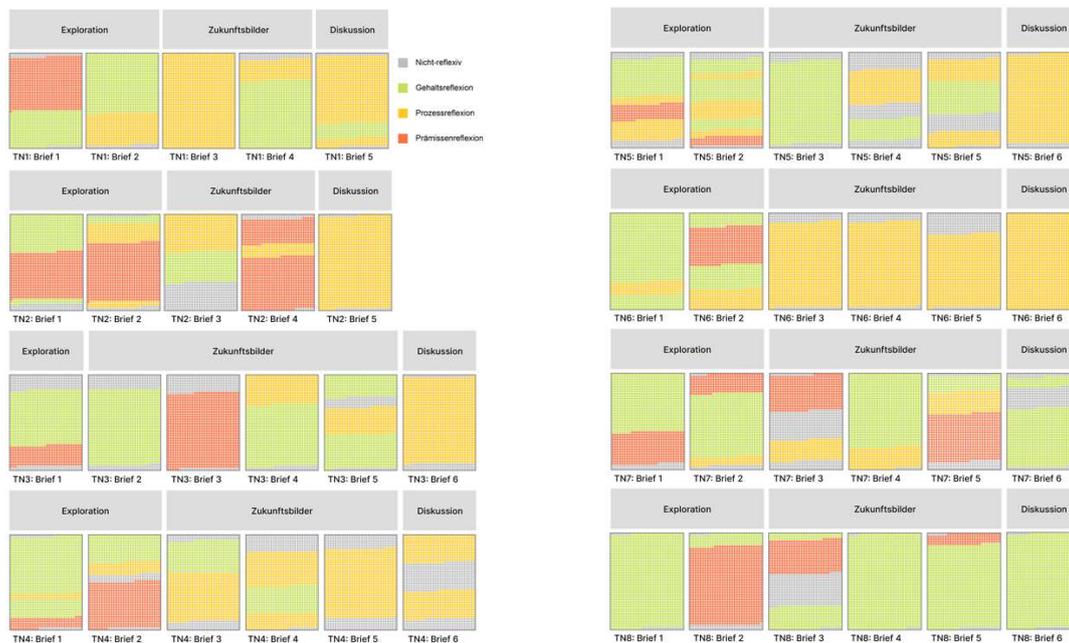


Abbildung 2: Dokumentenportraits

Durch die Analyse der Briefwechsel ist es weiterhin gelungen, Präkonzepte zu KI sowie Haltungen und Einstellungen mit Blick auf den Aspekt der Kreativität von KI zu erhalten. Dabei ist folgendes Beispiel repräsentativ für die Einstellung vieler Teilnehmenden gegenüber der Kreativität generativer KI-Systeme:

„Beeindruckt hatte mich, dass ChatGPT sogar Geschichten für Kinder schreiben kann. Ich frage mich allerdings schon, wie geistreich solche Geschichten sein können. Ich kann mir kaum vorstellen, dass eine Maschine jemals eine so tolle Geschichte wie ein Mensch schreiben kann...“ (TN7: Brief 1, Pos. 5).

Kreativität wird von den Teilnehmenden häufig auch als Abgrenzungsmöglichkeit zwischen menschlicher und künstlicher Intelligenz gesehen:

„Deshalb fand ich die Diskussion sehr interessant, wie man menschliche Intelligenz von künstlicher Intelligenz unterscheiden bzw. abgrenzen kann. Mir persönlich sind dabei sofort folgende Unterscheidungsmerkmale eingefallen: Kreativität, Empathie, Intuition, moralisches Urteilsvermögen“ (TN5: Brief 1, Pos. 3).

Mit Blick auf die Transformation der eigenen Perspektive als zentraler Aspekt transformativer Bildungs- und Lernprozesse ist abschließend insbesondere folgende Aussage relevant:

„Bislang waren mir diese Merkmale gar nicht bewusst, aber bei näherem Überlegen merkt man, dass unsere Welt immer schnelllebiger wird und die Ereignisse immer schwieriger vorauszusagen sind. Mir wird bewusst, dass wir Menschen die Zukunft machen. Daher kommt uns ein großer Teil der Verantwortung zu, weil wir diese Entwicklung der Zukunft somit auch beeinflussen können. Dafür ist es wichtig, sich mit der Entwicklung auseinanderzusetzen und für sich selbst zu überlegen, wie man möchte, dass die Zukunft aussehen wird und was jeder Einzelne dafür tun kann.“ (TN6: Brief 2, Pos. 2)

In diesem Zitat reflektiert die Teilnehmerin kritisch ihre bisherigen Prämissen über die Rolle des Einzelnen in einer sich schnell wandelnden Welt. Ihre Reflexion zeigt, wie sie sich auf das Wiederaufnehmen ihres Lebens unter neuen Bedingungen (einer von zehn von Mezirow beschriebenen Phasen transformativer Lernprozesse; vgl. Mezirow 1997, S. 143) vorbereitet, wobei sie ein breiteres und anpassungsfähiges Weltbild annimmt. Durch diese Prämissenreflexion wird deutlich, wie die Teilnehmerin ihre neuen Perspektiven aktiv in ihr Leben integriert und Verhaltensänderungen plant, die darauf abzielen, in einer dynamischen Umwelt proaktiv zu sein. Dieser Prozess unterstreicht ihre Bereitschaft, Verantwortung für die Mitgestaltung der Zukunft zu übernehmen, und illustriert den Übergang zu einem handlungsorientierten und reflexiven Ansatz im Umgang mit persönlichen und globalen Herausforderungen. Diese Anerkennung von Verantwortung verdeutlicht die von Koller als möglich beschriebene Rolle von Bildung als Transformation von Welt- und Selbstverhältnissen in der Gesellschaft und ihrer entsprechenden wechselseitigen Beeinflussung von Gesellschaft und Individuum (vgl. Koller 2012, S. 31).

Statt eines Fazits: Bildungsziele im Kontext von KI

Das weiter oben in aller Kürze skizzierte Verständnis transformativer Bildungs- und Lernprozesse liegt auch den im Folgenden zu Überlegungen für Bildungsziele im Kontext von KI zu Grunde.

Fuhr, Laros und Taylor heben hervor, dass *Bildung* und *transformatives Lernen* als tiefgreifende und komplexe Prozesse verstanden werden, bei denen Lernende grundlegende Annahmen und Erwartungen, die ihr Denken, Fühlen und Handeln prägen, (re)konstruieren. Beide Theorien (transformatives Lernen und transformative Bildung) betonen die Bedeutung der Reflexion, des kritischen Denkens und der Neubewertung von Erfahrungen als Schlüsselmechanismen für tiefgreifende Lernprozesse (vgl. Fuhr/Laros/Taylor 2017, S. ix).

Indem Subjekte dazu angeregt werden, ihre bisherigen Annahmen und Verständnisse kritisch zu hinterfragen und neue Perspektiven zu erkunden, wird ein Raum für persönliche und kollektive Transformation geschaffen. Dieser Prozess fördert nicht nur individuelles Wachstum und Selbstverständnis, sondern kann auch zu sozialem und kulturellem Wandel beitragen. In einer sich rapide verändernden und zunehmend komplexeren Welt, die geprägt ist

durch Volatilität, Unsicherheit, Komplexität und Ambiguität (vgl. zur Etablierung des Akronym VUCA Bennis/Nanus 1985) sowie Megatrends wie Digitalisierung, Globalisierung und Klimawandel, gewinnt ein solches Verständnis von Bildung und transformativem Lernen eine besondere Bedeutung, da eine Transformation von Welt- und Selbstverhältnissen aus einer normativen Perspektive, z. B. vor dem Hintergrund eines ökologischen Imperatives (vgl. Jonas 2020), gegeben ist. Bildung hat damit einerseits eine normative Komponente, was eine Intentionalität impliziert. Andererseits schließen sich Zielhaftigkeit und Prozesshaftigkeit nicht aus. Die Prozesshaftigkeit der Bildung betont, dass jene Veränderung von Welt- und Selbstverhältnissen ein lebenslanger, dynamischer Prozess ist, der durch krisenhafte Erfahrungen, Reflexion und wechselseitige Interaktion mit der Welt geprägt ist (vgl. Koller 2017, S. 33). Koller regt auch dazu an, darüber nachzudenken, anhand welcher normativer Kriterien sich transformative Prozesse als pädagogisch wünschenswerte und förderungswürdige Ereignisse bestimmen lassen. Er fragt auch, wie sich diese von bloßen Anpassungen an gesellschaftliche Erfordernisse unterscheiden. Mit anderen Worten: Verdient jede Veränderung im Verhältnis eines Subjekts zur Welt und zu sich selbst den Begriff Bildung (ebd., S. 35)?

Aus meiner Sicht gehört der Umgang mit KI zu einer Gruppe von globalen Herausforderungen, die genauso wie die Klimakrise im Interesse der Permanenz menschlicher Existenz adressiert werden sollte. Dies macht es notwendig, kontinuierlich neu zu definieren, was ein angemessenes Bildungsideal darstellt. Dazu gehört auch, geeignete Wege zu finden, um diese Fragen gesellschaftlich zu diskutieren und zu verhandeln. Mit dem Projekt „DoingKI“ wurde in diesem Beitrag ein Beispiel dafür vorgestellt, wie die Reflexion über mögliche Zukünfte mit KI dazu beitragen kann, relevante Vorstellungen zu identifizieren, die notwendig sind, um KI-Entwicklungen im Einklang mit den Interessen, Werten und Einstellungen der Menschheit zu gestalten. Diese tiefgreifende Auseinandersetzung mit KI und ihren potenziellen Zukünften leitet über zu der Frage, wie Bildungsprozesse gestaltet werden müssen, um nicht nur auf gegenwärtige, sondern auch auf zukünftige gesellschaftliche Anforderungen vorzubereiten. Dieser Bedarf erstreckt sich von der Schule über die universitäre Ausbildung bis hin zur Fort- und Weiterbildung und schließt dabei auch bildungstheoretische Haltungen und Menschenbildannahmen ein. In Bezug auf die Gestaltung von softwaregeprägten Bildungsangeboten ist ein aktuelles Beispiel die Diskussion um die Entwicklung einer nationalen Bildungsplattform (NBP) in Deutschland. Auch hier spielen u. a. der Bildungsbegriff sowie Annahmen über das Menschenbild eine zentrale Rolle bei der Entwicklung. So weisen Seemann et al. (2022) darauf hin, dass im Kontext der Entwicklung einer solchen zentralen Bildungsplattform

„das Wort ‚Lernen‘ viel häufiger vorkommt als ‚Bildung‘. Damit sind zwei Probleme verbunden: Zum einen wird Bildung dadurch ‚lernifiziert‘. D.h. es wird als unhinterfragte Selbstverständlichkeit vorausgesetzt, dass Lernen individuell und zielgerichtet ist und dass die Identifikation des Lern-Outputs – sei es im Sinne von Kompetenzen oder Wissen – allgemein erfassbar ist. Während ‚Lernen‘ häufig instrumentell verstanden wird, d.h. als Schritt auf dem Weg zu einem individuellen Ziel, geht ‚Bildung‘ als Kontrastfolie mit Unbestimmtheit einher. Wenn aber am Ende eines Bildungsprozesses eine konkrete Transformation von Selbst- und Weltverhältnissen stehen soll, muss den Unsicherheiten, produktiven Irritationen und Mehrdeutigkeiten im Lernprozess Raum gegeben werden“ (Seemann et al. 2022, S. 15).

Solche Systeme (auch auf maschinellem Lernen basierte Systeme wie Learning Analytics) werden oft einfach als Werkzeuge betrachtet, die zur Verbesserung des Lehrens und Lernens eingesetzt werden. Neben anderen weist Edwards (vgl. 2015) aber auf den Einfluss von Algorithmen und Standards bei der Auswahl und Gestaltung von Softwareangeboten und Interaktionsmodalitäten hin und betont, wie Inhalte durch Software in einer Weise ausgewählt, geformt und dargestellt werden, die für diejenigen, die sie verwenden, nicht immer offensichtlich sind. In diesem Sinne kann Software als Teil des „unsichtbaren Lehrplans“ betrachtet werden, welcher nicht zwangsläufig demokratisch oder gesellschaftlich legitimiert ist.

Ein letztes Beispiel für eine negative Dynamik ist der aufkommende „AI divide“ (Bubeck et al. 2023, S. 90). Dieser kann sich sowohl durch Zugangshürden zu KI (z. B. Kosten¹⁶) als auch ähnlich einem Second-Level-Digital-Divide (vgl. z. B. Niesyto 2019) bzw. einem Digital Skills Gap (vgl. Initiative D21 2023, S. 25) ausdrücken, wenn die notwendigen Kompetenzen zur Nutzung von KI-Systemen fehlen, um die Potenziale einer komplementären Ergänzung von menschlicher und künstlicher Intelligenz zu nutzen:

„We expect rich opportunities for innovation and transformation [...] with creative uses of AI technologies to support human agency and creativity and to enhance and extend human capabilities. Advances in AI can be leveraged in myriad ways to achieve new levels of skill or efficiency in human efforts and contributions. Investments in tasks, methods, and machinery to support and extend human problem-solving and decision making may be less obvious and more challenging than the identification of sets of tasks that might be automated by machines. However, there is great upside to seeking the means to richly leverage human and machine complementarities aimed at extending the capabilities of people“ (Bubeck et al. 2023, S. 90).

In diesem dynamischen Kontext erweist sich die Bestimmung von Bildungszielen nicht nur als notwendig, sondern als zentral für die Gestaltung einer menschenfreundlichen Zukunft. Ausgehend von dieser Prämisse lassen sich aus meiner Sicht drei zentrale Dimensionen identifizieren:

- *Qualifikatorische Dimension*: Diese wird bereits adressiert (siehe dazu z. B. die Aussage des baden-württembergischen Kultusministeriums weiter oben; hierzu gibt es weitere Überlegungen wie beispielsweise die DigComp Frameworks, die auf EU-Ebene zu verorten sind und damit über die nationale Perspektive hinausweisen). Es bleibt allerdings wichtig, ein kritisches Bewusstsein für die Potenziale und Risiken der KI sowie Fragestellungen und Implikationen für Individuum und Gesellschaft im Blick zu behalten.
- *Partizipative Dimension*: Bildungsangebote müssen Raum bieten für eine aktive Partizipation an der Gestaltung der KI-geprägten Zukunft. Dies erfordert, über die qualifikatorische Dimension hinauszugehen. Die Förderung kritisch-reflexiver Denkprozesse,

¹⁶ Während z. B. OpenAI im Mai 2024 einen eingeschränkten Zugang zu seinem neuesten Modell GPT4o ohne kostenpflichtiges Abonnement ermöglicht, bleiben viele Funktionen (in Bezug auf Sicherheit, Nutzungshäufigkeit) sowie deren Integration in bestehende Tools den Kund*innen vorbehalten (insbesondere z. B. bei Microsoft copilot oder Google Gemini Advanced), die einen monatlichen Beitrag zahlen. Andere generative Modelle (z. B. Midjourney, HeyGen, Descript), insbesondere zur Audio-, Bild- und Videogenerierung, sind ohne Bezahlung kaum nutzbar.

Diskurskompetenz und von ethischem Urteilsvermögen ist hierbei von entscheidender Bedeutung, um eine Entwicklung zu vermeiden, die lediglich kapitalistischen Interessen statt dem Gemeinwohl dient (zu *Kultur der Commons* vgl. Stalder 2019, S. 245).

- *Transformative Dimension:* In Anbetracht der tiefgreifenden Veränderungen, die KI mit sich bringt, zielt diese Dimension darauf ab, die Persönlichkeit in ihrer Gesamtheit zu entwickeln und unterstützt Individuen darin, ein selbstbestimmtes und erfülltes Leben zu führen. Transformative Bildung (vgl. Koller 2012) schließt die Entwicklung zur Selbstreflexion und das Streben nach eigenen Interessen und Leidenschaften mit ein. Nach Kollers Verständnis wird transformative Bildung durch Krisenerfahrungen initiiert. Diese können persönlicher, gesellschaftlicher oder erkenntnistheoretischer Natur sein und rütteln an den Grundfesten bisheriger Annahmen und Überzeugungen. Solche Krisenmomente bieten die Chance, das eigene Weltbild zu hinterfragen und möglicherweise neu zu orientieren (vgl. ebd., S. 15).

Ein weitgehender Verzicht auf die Nutzung von LLMs und damit auch der Verzicht auf deren (handlungsorientierte) Thematisierung in der Grundschule und zu Beginn der Sekundarstufe I, wie es die Ständige Wissenschaftliche Kommission der Kultusministerkonferenz in ihrem Papier „Large Language Models und ihre Potenziale im Bildungssystem“ Anfang 2024 fordert (vgl. SWK 2024, S. 4), scheint vor diesem Hintergrund als nicht zielführend. Diese Perspektive soll vermutlich dazu dienen, Kinder vor potenziellen Risiken und Überforderungen zu schützen. Langfristig birgt dies aber die Gefahr, dass Menschen nicht auf eine Zukunft vorbereitet werden, in der KI eine zentrale Rolle in allen Lebensbereichen spielen wird und sie nicht die notwendige Gestaltungskompetenz anbahnen können, um diese mitzugestalten. In diesem Sinne geht es nicht nur darum, Schüler*innen auf die technischen Aspekte von KI vorzubereiten, sondern auch darum, kritisches Denken zu fördern, zur Selbstreflexion anzuregen und ein tiefgreifendes Verständnis für die Komplexität der Welt sowie für die eigene Rolle in ihr zu entwickeln, wie dies z. B. Autenrieth und Nickel (vgl. 2023a) mit dem KI-Meta-Modell im Anschluss an die multiperspektivische Betrachtung des Frankfurt-Dreiecks (vgl. Brinda et al. 2019) vorschlagen. Dies setzt voraus, über ein traditionelles Verständnis von Lehren und Lernen hinauszugehen und Lernende dazu zu ermutigen, aktiv zu partizipieren, Fragen zu stellen und eigenständig Lösungen zu entwickeln. Dies beinhaltet auch die Auseinandersetzung mit ethischen Fragestellungen und den sozialen und kulturellen Auswirkungen technischer Entwicklungen. Transformative Ansätze, die Menschen dazu befähigen, ihre persönlichen Stärken und Interessen zu erkennen und zu verfolgen, unterstützen dabei jede Einzelne und jeden Einzelnen darin, einen individuellen Weg in einer sich stetig wandelnden Welt zu finden und an der Gesellschaft teilzuhaben. Damit ist Bildung nicht nur ein Instrument der Qualifikation, sondern auch ein Wegweiser für ein selbstbestimmtes und erfülltes Leben.

Literatur

- Alexander, Scott (2014): Meditations on Moloch [Wissenschaftsblog]. Cirkovicar Codex. slatestarcodex.com/2014/07/30/meditations-on-moloch/.
- Altman, Sam (2023): OpenAI chief seeks new Microsoft funds to build ‘superintelligence’. In: Financial Times (Madhumita Murgia). ft.com/content/dd9ba2f6-f509-42f0-8e97-4271c7b84ded.
- Aristoteles (2019): Nikomachische Ethik (Gernot Krappinger, Hrsg.). Ditzingen: Reclam.
- Autenrieth, Daniel/Nickel, Stefanie (2022): KuDiKuPa – Kultur der Digitalität = Kultur der Partizipation?!: Verschränkung von Theorie und Praxis in partizipativ angelegter Hochschullehre durch Gaming und Game Design – ein Praxisbeispiel. In: MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung, Jahrbuch Medienpädagogik 18 (Ästhetik – Digitalität – Macht), S. 237–265. doi.org/10.21240/mpaed/jb18/2022.02.26.X.
- Autenrieth, Daniel/Nickel, Stefanie (2023a): Das KI-Meta-Modell. Handlungsleitende Strukturen für den Umgang mit künstlicher Intelligenz im Bildungsbereich. In: technik-education (tedu), 2|2023 (Ausgabe 6), S. 1–20.
- Autenrieth, Daniel/Nickel, Stefanie (2023b): Zukunft lässt sich gestalten. Interdisziplinäre Zugänge zwischen Medienbildung und Bildung für nachhaltige Entwicklung. In: Medienimpulse, S. 1–25. doi.org/10.21243/MI-02-23-16.
- Autenrieth, Nina (2023): Das System verändert sich (langsam). In: Lehren und Lernen. Zeitschrift für Schule und Innovation aus Baden-Württemberg, 49 (7), S. 21–23.
- Bach, Joscha (2023): Synthetic Sentience. Can Artificial Intelligence become conscious? 37th Chaos Communication Congress: Unlocked, Hamburg. media.ccc.de/v/37c3-12167-synthetic_sentience.
- Baer, John (2016): Chapter 3 — Implications of Domain Specificity for Creativity Theory. In: Baer, John (Ed.): Domain Specificity of Creativity. San Diego: Academic Press, pp. 55–83. doi.org/10.1016/B978-0-12-799962-3.00003-3.
- Barrat, James (2013): Our final invention: Artificial intelligence and the end of the human era (First Edition). New York: Thomas Dunne Books.
- Bennis, Warren G./Nanus, Burt (1985): Leaders: The Strategies for Taking Charge. Nachdr.: New York: Harper & Row.
- Bierwisch, Antje (2023): UNESCO Chair in Futures Capability for Innovation and Entrepreneurship. UNESCO Chair. research.mci.edu/en/unesco-chair-futures-capability.
- Boden, Margaret A. (2009): Computer Models of Creativity. In: AI Magazine, 30(3), pp. 23–34. doi.org/10.1609/aimag.v30i3.2254.
- Bostrom, Nick (2017): Superintelligence: Paths, dangers, strategies (Reprinted with corrections 2017). Oxford: University Press.
- Bräuer, Gerd (2016): Das Portfolio als Reflexionsmedium für Lehrende und Studierende. Opladen, Toronto: Barbara Budrich.
- Brinda, Torsten/Brüggen, Niels/Diethelm, Ira/Knaus, Thomas/Kommer, Sven/Kopf, Christine/Missomelius, Petra/Leschke, Rainer/Tielmann, Friederike/Weich, Andreas (2019): Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt. Ein interdisziplinäres Modell. keine-bildung-ohne-medien.de/wp-content/uploads/2019/07/Frankfurt-Dreieck-zur-Bildung-in-der-digitalen-Welt.pdf.
- Bromberg, Kirstin (2017): Fostering Transformative Learning Through Dialogical Writing. In: Laros, Anna/Fuhr, Thomas/Taylor, Edward W. (Eds.): Transformative Learning Meets Bildung. Rotterdam/Boston/Taipei: SensePublishers, pp. 305–316. doi.org/10.1007/978-94-6300-797-9_25.
- Brynjolfsson, Erik/Mitchell, Tom (2017): What can machine learning do? Workforce implications. In: Science, 358(6370), pp. 1530–1534. doi.org/10.1126/science.aap8062.
- Bubeck, Sébastien/Chandrasekaran, Varun/Eldan, Ronen/Gehrke, Johannes/Horvitz, Eric/Kamar,

- Ece/Lee, Peter/Lee, Yin Tat/Li, Yuanzhi/Lundberg, Scott/Nori, Harsha/Palangi, Hamid/Ribeiro, Marco Tulio/Zhang, Yi (2023): Sparks of Artificial General Intelligence: Early experiments with GPT-4. doi.org/10.48550/ARXIV.2303.12712.
- Buckingham, David (2023): Artificial Intelligence in Education: A Media Education Approach. davidbuckingham.net/2023/05/27/artificial-intelligence-in-education-a-media-education-approach/.
- Candelon, Francois/Krayer, Lisa/Rajendran, Saran/Martínez, David Zuluaga (2023): How People Can Create—And Destroy—Value with Generative AI. mkt-bcg-com-public-pdfs.s3.amazonaws.com/prod/how-people-create-and-destroy-value-with-gen-ai.pdf.
- Dertinger, Andreas/Kramer, Michaela/Koschei, Franziska/Schmidt, Lena/Eggert, Susanne/Kammerl, Rudolf (2023): Wie hat sich das pandemiebedingte Distance-Schooling auf die Digitale Bildung im Grundschulalter ausgewirkt? Ein systematisches Review. In: Zeitschrift für Grundschulforschung, 16 (2), S. 449–464. doi.org/10.1007/s42278-023-00182-1.
- Döbeli Honegger, Beat (2021): Was machen wir mit der Digitalisierung? In: Pädagogik, 5/21, Artikel 5/21. doi.org/10.3262/PAED2105041.
- Edwards, Richard (2015): Software and the hidden curriculum in digital education. In: Pedagogy, Culture & Society, 23 (2), pp. 265–279. doi.org/10.1080/14681366.2014.977809.
- Ernst, Frederik/Mollenhauer, Clara/Dunker, Nina (2023): Das Reallabor Erlebnispark als Raum für transformative Lern- und Bildungsprozesse in der Lehrer:innenbildung. In: Autenrieth, Daniel/Autenrieth, Nina/Nickel, Stefanie (Hrsg.): (Virtuelle) Dritte Orte als Chance für eine nachhaltige Bildungslandschaft. Konzepte, Theorien und Praxisbeispiele. München: kopaed, S. 89–125.
- Falk, John H. (2022): The value of museums: Enhancing societal well-being. Lanham: Rowman & Littlefield.
- Fend, Helmut (2008): Neue Theorie der Schule: Einführung in das Verstehen von Bildungssystemen (2., durchgesehene Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fonagy, Peter/Gergely, György/Jurist, Elliot L./Target, Mary (2011): Affektregulierung, Mentalisierung und die Entwicklung des Selbst (Elisabeth Vorspohl, Übers.; 4. Auflage). Stuttgart: Klett-Cotta.
- Fuhr, Thomas/Laros, Anna/Taylor, Edward W. (2017): Transformative learning meets Bildung: Introduction. In: Laros, Anna/Fuhr, Thomas/Taylor, Edward W. (Eds.): Transformative learning meets Bildung: An international exchange. Rotterdam/Boston/Taipei: Sense Publishers, pp. ix–xvi.
- Fullan, Michael/Scott, Geoff (2014): Education PLUS. Collaborative Impact SPC, Seattle, Washington. michaelfullan.ca/wp-content/uploads/2014/09/Education-Plus-A-Whitepaper-July-2014-1.pdf.
- Future of Life Institute (2023): Pause Giant AI Experiments: An Open Letter [Offener Brief]. futureoflife.org/open-letter/pause-giant-ai-experiments/.
- Ginsberg, Allen (1984): „Howl“ from Collected Poems, 1947-1980. poetryfoundation.org/poems/49303/howl.
- Groß, Benedikt/Mandir, Eileen (2022): Zukünfte gestalten: Spekulation, Kritik, Innovation: mit Design Futuring Zukunftsszenarien strategisch erkunden, entwerfen und verhandeln. Mainz: Verlag Hermann Schmidt.
- Guilford, Joy Paul (1956): The structure of intellect. In: Psychological Bulletin, 53 (4), pp. 267–293. doi.org/10.1037/h0040755.
- Gurnee, Wes/Tegmark, Max (2023): Language Models represent Space and Time. doi.org/10.48550/ARXIV.2310.02207.
- Guzik, Erik E./Byrge, Christian/Gilde, Christian (2023): The originality of machines: AI takes the Torrance Test. In: Journal of Creativity, 33 (3), pp. 100065. doi.org/10.1016/j.jyoc.2023.100065.
- Haase, Jennifer/Hanel, Paul H.P. (2023): Artificial muses: Generative artificial intelligence chatbots

- have risen to human-level creativity. In: *Journal of Creativity*, 33(3), pp. 100066. doi.org/10.1016/j.yjoc.2023.100066.
- Hatton, Neville/Smith, David (1995): Reflection in teacher education: Towards definition and implementation. In: *Teaching and Teacher Education*, 11 (1), pp. 33–49. doi.org/10.1016/0742-051X(94)00012-U.
- Initiative D21 (2023): D21-Digital-Index 2022/23. initiatived21.de/uploads/03_Studien-Publikationen/D21-Digital-Index/2022-23/d21digitalindex_2022-2023.pdf.
- Jonas, Hans (2020): *Das Prinzip Verantwortung*. Berlin: Suhrkamp.
- Koller, Hans-Christoph (2012): *Bildung anders denken: Einführung in die Theorie transformativischer Bildungsprozesse*. Stuttgart: W. Kohlhammer.
- Koller, Hans-Christoph (2017): Bildung as a Transformative Process. In: Laros, Anna/Fuhr, Thomas/Taylor, Edward W. (Eds.): *Transformative learning meets Bildung: An international exchange*. Rotterdam/Boston/Taipei: Sense Publishers, pp. 33–43.
- Kurzweil, Ray (2006): *The singularity is near: When humans transcend biology*. New York: Penguin Books.
- La Rosée, Anna/Fuhr, Thomas/Taylor, Edward W. (Eds.) (2017): *Transformative Learning Meets Bildung: An International Exchange*. *International Issues in Adult Education*, Vol. 21. Rotterdam/Boston/Taipei: Sense Publishers.
- Lubart, Todd I. (1994): Chapter 10—Creativity. In: Sternberg, Robert J. (Ed.): *Thinking and Problem Solving*, Vol. 2, pp. 289–332. San Diego: Academic Press. doi.org/10.1016/B978-0-08-057299-4.50016-5.
- Luhmann, Niklas (1995): *Soziologische Aufklärung: Bd. 6*. Opladen: Westdeutscher Verlag.
- Marotzki, Winfried (2006): Bildungstheorie und Allgemeine Biographieforschung. In: Krüger, Heinz-Hermann/Marotzki, Winfried (Hrsg.): *Handbuch erziehungswissenschaftliche Biographieforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 59–70.
- Mayring, Philipp (2022): *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13., überarbeitete Auflage). Weinheim, Basel: Beltz.
- McCarthy, John/Minsky, Marvin L./Rochester, Nathaniel/Shannon, Claude E. (1955): A proposal for the dartmouth summer research project on artificial intelligence. jmc.stanford.edu/articles/dartmouth/dartmouth.pdf.
- Mezirow, Jack (1991): *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
- Mezirow, Jack (1997): *Transformative Erwachsenenbildung* (K. Arnold, Übers.). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2023): *ChatGPT: Künstliche Intelligenz im Unterricht—Informationsangebote für Lehrkräfte*. km-bw.de/Lde/startseite/schule/ChatGPT-Informationen-fuer-Lehrkraefte.
- Missler, Daniel (2023): *Moloch: The Most Dangerous Idea*. *Unsupervised Learning*. danielm-iessler.com/p/moloch-the-most-dangerous-idea.
- Monbiot, George (2017): *Out of the wreckage: A new politics for an age of crisis*. London: Verso.
- Moon, Jennifer A. (2013): *A Handbook of Reflective and Experiential Learning*. London: Routledge. doi.org/10.4324/9780203416150.
- Nash, John (1951): Non-Cooperative Games. In: *Annals of Mathematics*, 54 (2), p. 286. doi.org/10.2307/1969529.
- Niesyto, Horst (2019): *Medienkritik und pädagogisches Handeln*. *KULTURELLE BILDUNG ONLINE*. kubi-online.de/artikel/medienkritik-paedagogisches-handeln-0.
- Nohl, Arnd-Michael (2015): Typical Phases of Transformative Learning: A Practice-Based Model. In:

- Adult Education Quarterly, 65 (1), pp. 35–49. doi.org/10.1177/0741713614558582.
- OECD (2019): An OECD Learning Framework 2030. In: Bast, Gerald/Carayannis, Elias G./Campbell, David F. J. (Eds.): The Future of Education and Labor. Cham: Springer International Publishing, pp. 23–35. doi.org/10.1007/978-3-030-26068-2_3.
- Piketty, Thomas (2021): Der Sozialismus der Zukunft: Interventionen (André Hansen, Übers.; Originalausgabe). München: C.H. Beck.
- Rhodes, Mel (1961): An Analysis of Creativity. In: The Phi Delta Kappan, 42 (7), pp. 305–310.
- Robinson, Ken (2017): Out of our minds: The power of being creative (Third edition). Chichester: John Wiley & Sons, Ltd.
- Russell, Stuart J. (2019): Human compatible: Artificial intelligence and the problem of control. New York: Viking.
- Schreiner, Maximilian (2023): GPT-4 architecture, datasets, costs and more leaked. In: the decoder. the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/.
- Seemann, Michael/Macgilchrist, Felicitas/Richter, Christoph/Allert, Heidrun/Geuter, Jürgen (2022): Konzeptstudie Werte und Strukturen der Nationalen Bildungsplattform (Wikimedia Deutschland e. V., Hrsg.). wikimedia.de/wp-content/uploads/2022/11/Konzeptstudie-Werte-und-Strukturen-der-Nationalen-Bildungsplattform.pdf.
- Shiller, Derek (2024): Functionalism, integrity, and digital consciousness. In: Synthese, 203 (2), p. 47. doi.org/10.1007/s11229-023-04473-z.
- Stalder, Felix (2019): Kultur der Digitalität. Berlin: Suhrkamp.
- Stalder, Felix (2021): Was ist Digitalität? In: Hauck-Thum, Uta/Noller, Jörg (Hrsg.): Was ist Digitalität? Philosophische und pädagogische Perspektiven. Berlin: J.B. Metzler, S. 3–7.
- SWK – Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2024): Large Language Models und ihre Potenziale im Bildungssystem. Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz. doi.org/10.25656/01:28303.
- Tawil, Yoel (2019): Why Creativity Has No Definition [TEDx-Talk]. TEDxShenkarCollege, Israel. ted.com/talks/yoel_tawil_why_creativity_has_no_definition.
- Taylor, Edward W./Cranton, Patricia (2013): A theory in progress? Issues in transformative learning theory. doi.org/10.25656/01:7705.
- Tegmark, Max (2017): Life 3.0: Being human in the age of artificial intelligence. London: Allen Lane.
- Torrance, E. Paul (2012): Torrance Tests of Creative Thinking [dataset]. doi.org/10.1037/t05532-000.
- Ullrich, E. (1993): Kreativität. In: Arnold, Wilhelm/Eysenck, Hans/Meili, Richard (Hrsg.): Lexikon der Psychologie: Bd. 2: H-Psychodiagnostik (10. Auflage). Freiburg/Basel/Wien: Bechtermünz Verlag, S. 1156–1158.
- Vaswani, Ashish/Shazeer, Noam/Parmar, Niki/Uszkoreit, Jakob/Jones, Llion/Gomez, Aidan N./Kaiser, Lukasz/Polosukhin, Illia (2017): Attention Is All You Need. doi.org/10.48550/ARXIV.1706.03762.
- Voros, Joseph (2003): A generic foresight process framework. In: Foresight, 5 (3), pp. 10–21. doi.org/10.1108/14636680310698379.
- Writers Guild of America (2023): WGA Negotiations—Status as of May 1, 2023. wgacontract2023.org/uploadedfiles/members/member_info/contract-2023/wga_proposals.pdf.
- Yudkowsky, Eliezer (2008): Artificial Intelligence as a positive and negative factor in global risk. In: Bostrom, Nick/Cirkovic, Milan M. (Eds.): Global Catastrophic Risks. Oxford: University Press. doi.org/10.1093/oso/9780198570509.003.0021.

Informationen zum Autor



Daniel Autenrieth ist Informatiker und Medienpädagoge mit dem Schwerpunkt politisch-kultureller Medienbildung. Mit seinem Unternehmen Autenrieth & Partner berät er Firmen und öffentliche Einrichtungen bei Vorhaben an den Schnittstellen von Bildung und Digitalisierung. Als Geschäftsführer von PulseDataInsight leitet er ein interdisziplinäres Team zur Erforschung und Entwicklung von KI in der Medizin.

kontakt@daniel-autenrieth.de

Zitationshinweis:

Autenrieth, Daniel (2024): Auf dem Weg zur Singularität. Implikationen für Bildung, Kreativität und den Bedarf der Mitgestaltung. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 24, S. 1–25. doi.org/10.21240/lbzm/24/09.