

Mit KI unterrichten, nicht um sie herum

Ein entwicklungsorientiertes Rahmenwerk für die Integration generativer KI in der Primar-, Sekundar- und Tertiärbildung (2023–2026)

N.J. Berky

NJB Berky Advisory · Prag / Zürich

ORCID: 0009-0001-2227-215X · njb@njb-berky.com · njb-berky.com

Arbeitspapier · Version 2.0 · Juni 2026

Deutsche Übersetzung des englischen Originals "Teaching With, Not Around" (Berky, 2026)

ABSTRACT

Generative künstliche Intelligenz (GenKI) ist die sich am schnellsten verbreitende Allzwecktechnologie in der modernen Bildungsgeschichte. Zwischen November 2022 und Anfang 2026 entwickelten sich grenzgängige große Sprachmodelle in drei akademischen Zyklen von einer Forschungskuriosität zu einem allgegenwärtigen Schülerwerkzeug. Die Bildungssysteme nicht. Dieses Papier vertritt die Auffassung, dass die dominanten institutionellen Haltungen, also Verbot, unregulierte Übernahme und überwachungsbasierte Kontrolle, allesamt entwicklungspsychologisch ungebildet sind, und schlägt ein vierstufiges kognitives Rahmenwerk vor (Delegation, Beschreibung, Urteilsvermögen, Sorgfalt; nachfolgend 4D), das auf die neurodevelopmentale Realität von Lernenden in der Primar-, Sekundar- und Tertiärbildung kalibriert ist. Auf Grundlage einer Synthese empirischer Arbeiten aus den Jahren 2023 bis 2026 zum kognitiven Auslagern (Kosmyrna et al., 2025; Gerlich, 2025; Fan et al., 2025; Gilbert, 2024), zur Entwicklungsneurowissenschaft der präfrontalen Reifung unter digitaler Exposition (Hutton et al., 2022; Hinze et al., 2025; Lewin et al., 2025; Gascón-Cánovas et al., 2025), zu den KI-Kompetenzrahmen der EU und OECD (Europäische Kommission & OECD, 2025; UNESCO, 2024) sowie zur bevorstehenden PISA-2029-Erhebung zur Medien- und KI-Kompetenz (OECD, 2026) stellt das Papier vier Behauptungen auf. Erstens sollte die Exposition gegenüber GenKI strukturell asymmetrisch über die Entwicklungsstufen hinweg gestaltet sein, nicht einheitlich erlaubt oder einheitlich verboten. Zweitens ist die Modellauswahl in regulierten Bildungsumgebungen keine triviale pädagogische Entscheidung; methodologisch transparente Systeme, etwa jene, die nach den Constitutional-AI-Protokollen (Bai et al., 2022) trainiert wurden, bilden eine vertretbarere Grundlage für die institutionelle Bereitstellung als Systeme, deren Alignment-Methodik nicht offengelegt ist. Drittens verlagert sich die Rolle der Lehrkraft vom Inhaltsvermittler zum kognitiven Coach, und die Lehrkräfteausbildung muss sich entsprechend verlagern. Viertens konvergiert das internationale politische Umfeld rasch um vierdomänige Kompetenzmodelle (sich einbringen, gestalten, verwalten, entwerfen), was sowohl eine Verpflichtung als auch eine Gelegenheit zur institutionellen Reform schafft. Das Papier schließt mit operativen Empfehlungen für Institutionen im Horizont 2026–2030, einem Kapitel zu Einschränkungen und einer Agenda für künftige empirische Forschung.

Schlüsselwörter: generative KI · KI-Kompetenz · kognitives Auslagern ·

Entwicklungsneurowissenschaft · Bildungstechnologie · 4D-Rahmenwerk · Constitutional AI ·

KI-Ethik in der Bildung · exekutive Funktion in der Adoleszenz · OECD-AILit-Rahmen · PISA 2029 ·

Schutz durch Voreinstellung in der Pädagogik

1. Einleitung · Das Asymmetrieproblem

In den achtzehn Monaten zwischen der öffentlichen Veröffentlichung von ChatGPT durch OpenAI (November 2022) und der breiten institutionellen Übernahme grenzgängiger großer Sprachmodelle (LLMs) durch professionelle Dienstleistungen,

Softwareentwicklung und Forschungsabläufe (2024) durchlief generative KI eine Marktdurchdringungskurve, für die frühere Allzwecktechnologien (Elektrizität, Verbrennungsmotor, Personalcomputer, Internet) Jahrzehnte benötigten. Anfang 2026 fungieren grenzgängige Modelle wie Claude (Anthropic), GPT-4 und GPT-5 (OpenAI) sowie Gemini (Google DeepMind) als Standardpartner beim Denken und Schreiben für Dutzende Millionen Lernende weltweit, einschließlich Primarschulsystemen, deren Lehrpläne nicht darauf ausgelegt waren, diese Technologien aufzunehmen. Eine kürzlich veröffentlichte Metaanalyse von 35 kontrollierten Studien mit 4 193 Teilnehmern berichtet von einem mittleren positiven Effekt von ChatGPT auf die Lernergebnisse von Studierenden (Hedges' $g = 0,670$) über primäre, sekundäre und tertiäre Kontexte hinweg (Deng et al., 2026), was bestätigt, dass die Technologie pädagogisch nicht neutral ist. Die Frage lautet nicht mehr, ob Lernende diese Systeme nutzen. Sie tun es. Die Frage lautet, ob sie sie auf eine Weise nutzen, die kognitive Kapazität aufbaut, oder auf eine Weise, die sie ersetzt.

Dies ist kein Werkzeugproblem. Es ist ein Entwicklungsproblem. Ein neunjähriges Kind, ein fünfzehnjähriger und ein zweiundzwanzigjähriger Lernender, die mit demselben großen Sprachmodell interagieren, interagieren nicht mit demselben Artefakt; ihre präfrontale Rinde, ihre Arbeitsgedächtniskapazität, ihre metakognitiven Reflexe und ihre epistemischen Voreinstellungen sind radikal verschieden. Die präfrontale Rinde, der Sitz exekutiver Funktionen, des Urteilsvermögens und der inhibitorischen Kontrolle, erreicht ihre ausgereifte Myelinisierung nicht vor Mitte der zwanziger Jahre (Diamond, 2013). Die kognitiven Routinen, die es einer erwachsenen Forscherin ermöglichen, ein grenzgängiges Modell konstruktiv zu nutzen, also Quellenprüfung, Hinterfragen von Prämissen, Widerspruch gegenüber dem Output, sind genau jene Routinen, die ein Grundschulkind aktiv ausbildet. GenKI als einheitliche Klassenraumtechnologie zu behandeln, die an einer einzigen institutionellen Schwelle erlaubt, eingeschränkt oder kontrolliert wird, ist ein Kategorienfehler. Das Gehirn, das das Werkzeug benutzt, bestimmt, was das Werkzeug dem Lernenden antut.

Bildungssysteme weltweit haben mit drei dominanten Haltungen reagiert, von denen meines Erachtens jede unzureichend ist. Die erste ist das Verbot: pauschale Verbote oder Beschränkungen der Schülernutzung von GenKI, derzeit in irgendeiner Form in rund vierzig Prozent der nationalen Bildungssysteme in Kraft (UNESCO, 2024). Verbote sind Vollzugstheater. Sie überstehen den Kontakt mit einer Generation, die grenzgängige Modelle in ihren Taschen trägt, nicht und produzieren ein perveres Ergebnis: Lernende lernen, KI ohne Aufsicht, ohne Anleitung und ohne Lerngerüst zu nutzen, gerade weil das institutionelle Schweigen ihr Lernen in den Schattenlehrplan verlagert (Kharbach, 2026). Die zweite ist die unregulierte Übernahme: institutionelles Schweigen, das es Lernenden erlaubt, GenKI in jede kognitive Aufgabe ohne Gerüst, Aufsicht oder pädagogische Rahmung zu integrieren. Dies ist die modale Haltung der meisten tertiären Institutionen heute, und sie produziert das, was Fan et al. (2025) als metakognitive Trägheit beschreiben, die systematische Erosion der selbstregulatorischen Routinen, die Lernen von bloßer Aufgabenerfüllung unterscheiden. Die dritte ist die überwachungsbasierte Kontrolle: der Einsatz von KI-Erkennungswerkzeugen (die nicht zuverlässig funktionieren) und disziplinarische Verfahren gegen Lernende für KI-Nutzung, die Institutionen ihnen nicht beigebracht haben zu navigieren. Diese Haltung zersetzt die institutionelle Legitimität und produziert weder Lernen noch Einhaltung.

Alle drei Haltungen teilen ein einziges Versagen: Sie behandeln GenKI als externes Objekt, das kontrolliert werden muss, statt als kognitiven Partner, dessen angemessene Nutzung eine entwicklungsgestuft zu erlernende Fertigkeit ist. Dieses Papier vertritt das

Gegenteil. KI-Kompetenz ist, wie Lese-, Rechen- oder kritische Denkkompetenz, eine Kompetenz, die bewusst vermittelt werden muss, sequenziell, kalibriert auf die kognitiven Kapazitäten des Lernenden in jeder Entwicklungsstufe. Das Argument lautet nicht, dass GenKI inhärent gut oder inhärent schlecht für die Kognition sei. Es lautet, dass die Frage falsch gestellt ist. Die relevanten Fragen lauten: für welche kognitiven Aufgaben, in welcher Entwicklungsphase, unter welcher Aufsicht, mit welchem Schutz, zu welchem pädagogischen Zweck. Dies sind Fragen der Pädagogik, nicht allein der Politik.

Ich schlage ein vierstufiges kognitives Rahmenwerk vor (das 4D-Rahmenwerk: Delegation, Beschreibung, Urteilsvermögen, Sorgfalt) und zeige, wie jede Stufe in der Primarbildung (5–11 Jahre), Sekundarbildung (12–17 Jahre) und Tertiärbildung (18+) eingeführt, sequenziert und verstärkt werden sollte. Ich stütze mich auf empirische Arbeiten aus 2023–2026 zum kognitiven Auslagern und zur metakognitiven Trägheit, auf die Entwicklungsneurowissenschaft exekutiver Funktionen unter digitaler Exposition und auf die konvergenten KI-Kompetenzrahmen, die nun von der UNESCO (2024), der Europäischen Kommission und der OECD (2025) sowie nationalen Bildungsbehörden hervorgebracht werden. Schließlich argumentiere ich, dass die Modellauswahl selbst eine pädagogische Entscheidung ist: Institutionen, die GenKI an Minderjährige ausliefern, tragen eine treuhänderische Verpflichtung, Systeme auszuwählen, deren Alignment-Methodik transparent und prüfbar ist, und ich untersuche die von Anthropic entwickelte Constitutional-AI-Methodik (Bai et al., 2022) als eine vertretbare Grundlage für eine solche Auswahl.

Dieses Papier wird als Arbeitsrahmen angeboten, nicht als fertiger Rahmen. Empirische Belege zu den kognitiven Folgen anhaltender GenKI-Exposition häufen sich bis Mitte 2026 rasch, sind aber bei weitem nicht vollständig. Die hier gegebenen Empfehlungen stellen meine beste Synthese der verfügbaren Belege dar, kombiniert mit operativer Erfahrung im Unterrichten angewandter KI auf tertiärer Ebene an IFSA Prag, CEA CAPA und der Anglo-American University Prague, dem Mentoring frühphasiger Ventures durch öffentlich-private Innovationszentren in Mitteleuropa und dem Architekturentwurf eingesetzter multiparadigmatischer KI-Systeme für regulierte industrielle Kontexte. Kapitel 12 legt die Einschränkungen der vorliegenden Analyse und eine Agenda für die empirische Arbeit dar, die das Rahmenwerk verlangt.

2. Das Beschleunigungsproblem · KI-Geschwindigkeit und Bildungsreaktion

Zwischen der öffentlichen Veröffentlichung von GPT-3.5 im November 2022 und dem Einsatz grenzgängiger Reasoning-Modelle Anfang 2026 ist die Leistungsgrenze öffentlich zugänglicher LLMs durch mindestens vier große Generationenwechsel vorgerückt. Standardisierte Benchmarks belegen kumulative Verbesserungen: Modelle, die 2022 zusammenhanglose Absätze produzierten, produzieren 2026 routinemäßig analytisches Schreiben auf Masterniveau; Modelle, die 2023 keine mehrstufigen mathematischen Überlegungen anstellen konnten, lösen heute Wettbewerbsaufgaben; Modelle, die 2024 grundlegende Fakten halluzinierten, sind zunehmend in der Lage, reale, überprüfbare Quellen zu zitieren. Die Trajektorie ist nicht linear. Sie ist nach jedem vernünftigen Maßstab im Fenster 2022–2026 exponentiell. Die Verdoppelungszeit der Fähigkeiten auf zentralen Reasoning-Benchmarks blieb während des gesamten Zeitraums kürzer als zwei Jahre, und die Inferenzkosten sind um etwa eine Größenordnung pro Jahr gefallen, wodurch grenzgängige Fähigkeiten immer jüngeren und ärmeren Lernenden zugänglich werden.

Institutionelle Bildungssysteme sind nicht exponentiell. Sie sind, von ihrer Anlage her, konservativ. Langsam beim Aktualisieren von Lehrplänen, langsam bei der Umschulung von Lehrkräften, langsam bei der Überarbeitung von Bewertungsrahmen, langsam bei der Neuverhandlung des impliziten Vertrags zwischen Lernendem, Institution und Qualifikation. Dies ist kein Mangel; Stabilität und Kontinuität sind zentrale institutionelle Funktionen von Bildung. Das Qualifikationssystem hängt davon ab. Ebenso der Vertrauenskontrakt zwischen Lehrkraft, Lernendem und dem Korpus des zu vermittelnden Wissens. Die Langsamkeit ist das Merkmal, nicht der Fehler. Sie ist jedoch im gegenwärtigen Moment eine strukturelle Diskrepanz. Die Fähigkeitslücke zwischen dem, was Lernende mit grenzgängiger KI tun können, und dem, was ihre Institutionen ihnen mit grenzgängiger KI zu tun beigebracht haben, weitet sich mit jeder Modellveröffentlichung.

Drei konkrete Folgen ergeben sich. Erstens, der Schattenlehrplan: Lernende erlernen die Nutzung von GenKI nun primär über kommerzielle Plattformen, Social-Media-Tutorials, Peer-Übertragung und unbeaufsichtigtes Trial-and-Error (Kharbach, 2026; UNESCO, 2024). Was Institutionen formal über KI lehren, hinkt durchgängig eine Generation hinter dem zurück, was Lernende informell erleben. Der Schattenlehrplan ist unreguliert, sozioökonomisch ungleich und pädagogisch variabel. Er ist zugleich zunehmend der primäre Mechanismus, durch den Kinder ihre Vorstellung davon bilden, was KI ist und wie sie genutzt werden sollte. Zweitens, der Kollaps der Bewertungsvalidität: traditionelle Hausaufgaben, häusliche Essays und viele Formen schriftlicher Klassenraumbewertung messen nicht mehr zuverlässig die individuelle kognitive Leistung, weil die Grenze zwischen Schülerarbeit und KI-gestützter Arbeit nicht mehr mit Sicherheit erkennbar ist. Erkennungswerkzeuge haben sich als unzuverlässig erwiesen, insbesondere bei Nicht-Muttersprachlern des Englischen, bei denen sie systematisch höhere Raten falsch-positiver Befunde produzieren (Liang et al., 2023, wie in Kharbach, 2026 besprochen). Drittens, das Qualifikationsvertrauensproblem: Arbeitgeber, Regulierungsbehörden und nachgelagerte Institutionen können aus einer Qualifikation zunehmend weniger ableiten, was eine Lernende ohne Unterstützung tatsächlich leisten kann. Dies ist ein sich langsam entwickelndes, aber potenziell schwerwiegendes institutionelles Risiko für das Qualifikationssystem selbst.

Zwei politische Entwicklungen in den Jahren 2024–2025 verdienen eine Hervorhebung. Die KI-Verordnung der Europäischen Union trat am 1. August 2024 in Kraft, wobei Artikel 4 eine verbindliche Verpflichtung zur KI-Kompetenz für Anbieter und Nutzer von KI-Systemen mit einer gestaffelten Umsetzung über 2025–2027 hinweg auferlegt (Europäische Union, 2024). Die UNESCO veröffentlichte auf der Digital Learning Week 2024 zwei KI-Kompetenzrahmen (einen für Schülerinnen und einen für Lehrkräfte), die 15 Lehrkräftekompetenzen über fünf Dimensionen und einen parallelen Rahmen für Schülerinnen definieren (UNESCO, 2024; Miao & Cukurova, 2024). Die Europäische Kommission und die OECD veröffentlichten im Mai 2025 gemeinsam das KI-Kompetenzrahmenwerk für die Primar- und Sekundarbildung (AILit), das KI-Kompetenz um vier praktische Domänen organisiert (sich mit KI einbringen, mit KI gestalten, KI verwalten, KI entwerfen), gestützt auf 22 Kompetenzen (Europäische Kommission & OECD, 2025). Die OECD veröffentlichte im Februar 2026 den ersten Entwurf des Erhebungsrahmens für die PISA-2029-Erhebung zur Medien- und KI-Kompetenz (MAIL) (OECD, 2026), womit feststeht, dass KI-Kompetenz bis zum Ende des Jahrzehnts eine gemessene, international vergleichbare Kompetenz sein wird. Die globale Bereitschaft gegenüber diesen Rahmenwerken wurde auf unter 4 % geschätzt (Welker, 2025), was bedeutet, dass etwa 96 % der nationalen Bildungssysteme ohne die

Kompetenzinfrastruktur arbeiten, die ihre Bevölkerung jetzt benötigt.

Das in diesem Papier angebotene Rahmenwerk ist eine mögliche Antwort auf diese Lücke. Es ist nicht die einzig mögliche Antwort. Es ist spezifisch auf die kognitiv-entwicklungsbezogenen Realitäten von K-12- und Tertiärlernenden sowie auf die Sicherheitseigenschaften derzeit verfügbarer grenzgängiger Modelle kalibriert. Andere Antworten, ob rein strukturell, rein politisch oder rein werkzeugbezogen, sind möglich und parallel verfolgenswert. Das hier vorgestellte 4D-Rahmenwerk soll die vier AILit-Domänen, die UNESCO-Rahmen und entstehende nationale Rahmen ergänzen; es operiert auf einer anderen Abstraktionsebene und konzentriert sich auf die kognitiven Routinen, die ein Lernender in jeder Entwicklungsphase aufbauen muss, um KI produktiv zu nutzen, und nicht auf die Kompetenzergebnisse, die diese Routinen hervorbringen sollten.

3. Empirische Grundlagen · Eine Synthese der Belege von 2023-2026

Bevor ein Rahmenwerk vorgeschlagen wird, ist zu klären, was die empirische Literatur belegt und was nicht. Die Evidenzbasis zu GenKI in der Bildung ist mittlerweile umfangreich, aber ungleichmäßig; einige Befunde sind robust und konvergent, andere vorläufig und umstritten. Dieses Kapitel fasst den Evidenzstand bis Mitte 2026 in vier Clustern zusammen: kognitives Auslagern und metakognitive Effekte, Studien zu Lernergebnissen, neurodevelopmentale und verhaltensbezogene Korrelate intensiver digitaler Exposition sowie die internationale Politik- und Kompetenzlandschaft. Das in den folgenden Kapiteln vorgeschlagene Rahmenwerk baut explizit auf dieser Evidenzbasis auf, wobei die Stärke jedes Evidenzanspruchs dort vermerkt ist, wo sie eine Empfehlung trägt.

3.1 Kognitives Auslagern und metakognitive Effekte

Kognitives Auslagern bezeichnet die Nutzung externer Ressourcen zur Verringerung der internen kognitiven Beanspruchung (Gilbert, 2024). Allgemein ist es ein normales und adaptives Merkmal menschlicher Kognition; wir lagern routinemäßig an Notizbücher, Kalender, Suchmaschinen und andere Menschen aus. Die empirische Frage, die GenKI aufwirft, lautet, ob die spezifischen Eigenschaften großer Sprachmodelle, also flüssige Ausgabe in natürlicher Sprache, reibungsloser Zugang, unmittelbare bestätigende Reaktion, Muster des Auslagerns erzeugen, die die kognitiven Routinen, die Lernende ausbilden müssen, systematisch erodieren.

Für diese Frage sind drei Studien aus dem Jahr 2025 zentral. Kosmyna et al. (2025) verwendeten in einem technischen Bericht des MIT Media Lab neurophysiologische Messungen (aus EEG abgeleitete Metriken kognitiver Belastung), um die Leistung beim Schreiben von Essays unter Bedingungen zu untersuchen, in denen erwachsene Teilnehmende einen KI-Assistenten nutzten, eine Standard-Internetsuche nutzten oder weder das eine noch das andere nutzten. Die Studie berichtet von einer Akkumulation dessen, was die Autoren als kognitive Schuld bezeichnen, einer messbaren Verringerung der eingebrachten kognitiven Anstrengung während der KI-gestützten Bedingung, die über Sitzungen hinweg anhielt und mit einem geringeren selbstberichteten Aufgabenengagement assoziiert war. Gerlich (2025) berichtet in einer in *Societies* publizierten Studie von einem parallelen Befund an einer größeren, jedoch methodologisch weniger intensiven Stichprobe: eine höhere selbstberichtete Abhängigkeit von KI-Werkzeugen korrelierte mit niedrigeren Leistungen im kritischen Denken auf Standardinstrumenten, wobei das kognitive Auslagern diese Beziehung

vermittelte. Fan et al. (2025) entwarfen im *British Journal of Educational Technology* eine experimentelle Studie mit Essay-Aufgaben, um das zu untersuchen, was sie als metakognitive Trägheit bezeichnen, die systematische Reduktion selbstregulatorischer kognitiver Aktivität unter KI-gestützten Bedingungen, und berichten von konsistenten Effekten auf Lernmotivation, -prozesse und -leistung.

Diese Studien sollten mit angemessener Vorsicht gelesen werden. Keine ist eine Längsschnittstudie an Kindern. Kosmyna et al. (2025) untersuchten Erwachsene bei Essay-Aufgaben; die gemessenen kognitiven Routinen sind ausgereift, nicht in Entwicklung begriffen. Gerlich (2025) stützt sich sowohl für die KI-Nutzung als auch für die Ergebnisse des kritischen Denkens auf Selbstauskünfte, mit den methodologischen Einschränkungen, die solche Designs mit sich bringen. Fan et al. (2025) messen kurzfristige Effekte, nicht entwicklungsbezogene Folgen. Dennoch ist die Konvergenz dreier unabhängiger Forschungsgruppen, die unterschiedliche Methoden verwenden, auf einen gemeinsamen Befund (dass reibungslose KI-Unterstützung die eingebrachte kognitive Aktivität auf eine Weise reduziert, die sowohl verhaltensbezogen als auch physiologisch erkennbar ist) bedeutsam. Sie reicht noch nicht aus, um nachzuweisen, dass die Exposition von Kindern gegenüber KI eine entwicklungsbezogene kognitive Schädigung verursacht. Sie reicht aus, um die Hypothese ernst zu nehmen, dass pädagogisches Design, das es versäumt, gezielt Reibung einzuführen, Lernende hervorbringt, die die kognitiven Routinen nicht aufgebaut haben, die die Nutzung von KI erfordert.

Zwei weitere Studien erweitern das Bild. Ergene und Ergene (2025) untersuchen in *Education and Information Technologies* das kognitive Auslagern und die Überabhängigkeit von GenKI speziell in Kontexten der Hochschulbildung und berichten je nach instruktionalem Rahmen sowohl von produktiven als auch von Abhängigkeitsmustern. Lee und Low (2024) untersuchen in *Computers & Education: Artificial Intelligence* Abwägungen zwischen Effizienz und Originalität im akademischen Schreiben unter KI-Unterstützung und berichten, dass Studierende, die KI-gestützten Text produzieren, Aufgaben schneller abschließen, jedoch Arbeiten mit messbar geringerer sprachlicher und konzeptueller Vielfalt produzieren. Die Implikation lautet nicht, dass KI-Unterstützung im akademischen Schreiben kategorisch schädlich ist; sie lautet, dass die Routine, die Originalität hervorbringt (anhaltende persönliche Auseinandersetzung mit einem Problem), eine der Routinen ist, die KI-Unterstützung am leichtesten ersetzt. Pädagogisches Design muss diese Routine gezielt schützen.

3.2 Studien zu Lernergebnissen

Ein gegenläufiger Teil der Evidenz berichtet von positiven Lerneffekten aus der Integration von GenKI. Deng, Wang und Zhang (2026) berichten in einer Metaanalyse von 35 experimentellen Studien mit 4 193 Teilnehmenden, die zwischen 2022 und 2024 veröffentlicht wurden (*Humanities and Social Sciences Communications*), von einem mittleren positiven Effekt von ChatGPT auf die Lernergebnisse von Studierenden (Hedges' $g = 0,670$), mit signifikanten moderierenden Effekten von Fachgebiet, Experimentdauer und Unterrichtsmodus, jedoch nicht von der Bildungsebene. Cunningham et al. (2025) berichten in einer einjährigen Studie an der University of Tennessee at Chattanooga (*Frontiers in Education*) von transformativen Auswirkungen der ChatGPT-Integration auf die Motivation und den Kompetenzerwerb von Postgraduierten, mit den größten Effekten in Kontexten strukturierter Aufgaben. Raptopoulou (2025) berichtet in einer qualitativen Studie an einer schwedischen Universität (*European Journal of Education*), dass ChatGPT die Entwicklung

akademischer Literalität unterstützt, wenn er mit klaren pädagogischen Zielen integriert wird, jedoch nicht, wenn er ohne Lerngerüst eingesetzt wird.

Die scheinbare Spannung zwischen der Literatur zum Auslagern (Kapitel 3.1) und der Literatur zu Lernergebnissen (hier) löst sich auf, wenn die abhängigen Variablen sorgfältig betrachtet werden. Die Studien zum Auslagern messen kognitives Engagement und Routinebildung, die Studien zu Lernergebnissen messen Aufgabenerfüllung und Jahrgangsstufenleistung. Es ist vollkommen konsistent, dass KI-unterstützte Lernende Aufgaben besser abschließen und weniger lernen, wenn Lernen als Bildung kognitiver Routinen und nicht als Produktion korrekter Outputs verstanden wird. Diese Unterscheidung zwischen Ergebnissen der Aufgabenerfüllung und Ergebnissen der Routinebildung ist die zentrale pädagogische Unterscheidung, die das vorliegende Rahmenwerk operationalisiert. Kapitel 4 entwickelt sie ausführlich.

3.3 Neurodevelopmentale und verhaltensbezogene Korrelate

Direkte neurodevelopmentale Forschung zur GenKI-Exposition existiert bislang in keiner methodologisch robusten Form; die Technologie ist zu neu, und die erforderlichen Längsschnittzeiträume (Jahre bis Jahrzehnte) schließen dies aus. Die einschlägige Stellvertreterliteratur betrifft digitale Exposition allgemein, insbesondere die Entwicklungsneurowissenschaft intensiver Bildschirm- und Smartphone-Nutzung bei Kindern und Jugendlichen. Hutton et al. (2022) berichten in JAMA Pediatrics von Assoziationen zwischen bildschirmgestützter Mediennutzung und reduzierter Integrität der weißen Substanz bei Kindern im Vorschulalter, speziell in Trakten, die Sprache und entstehende Literalität unterstützen. León-Domínguez (2024) bespricht in Neuropsychology potenzielle kognitive Risiken generativer transformerbasierter KI-Chatbots für höhere exekutive Funktionen. Lewin et al. (2025) berichten in der International Journal of Clinical and Health Psychology von Längsschnitteffekten bildschirmgestützter Medienaktivitäten auf Gehirnfunktion, -architektur und psychische Gesundheit in der frühen Adoleszenz. Gascón-Cánovas et al. (2025) berichten von einer differentiellen Leistung exekutiver Funktionen bei Jugendlichen mit problematischer Smartphone-Nutzung im Vergleich zu gematchten Kontrollen (110 Schüler, Alter 13-18), wobei Effekte auf Eltern-berichteten Maßen erkennbar sind, selbst dort, wo kontrollierte Laboraufgaben sie unterschätzen können. Hinze et al. (2025) zeigen in einer Längsschnittstudie an 8 072 britischen Jugendlichen aus der MYRIAD-Studie stabile bidirektionale Assoziationen zwischen jugendlicher exekutiver Funktion und Ergebnissen der psychischen Gesundheit, einschließlich Wohlbefinden, depressiven Symptomen und Suizidalität.

Ein Scoping-Review von Reggente et al. (2026) in Pediatric Reports analysiert 104 Artikel, die zwischen 2022 und 2025 über Smartphone-Nutzung und Gehirnfunktion von Jugendlichen veröffentlicht wurden, und identifiziert fünf wiederkehrende Themen: psychische Störungen (n = 52), Schlaf (n = 43), sozioemotionale Funktion (n = 23), exekutive Funktion (n = 14) und sensorische Verarbeitung (n = 1). Der Review dokumentiert reduzierte funktionelle Konnektivität in kortiko-subkortikalen Schaltkreisen, reduziertes Volumen der grauen Substanz im oberen und unteren frontalen Gyrus sowie aberrante kortikale Aktivierung in frontoparietalen Regionen während kognitiver Konfliktaufgaben bei Jugendlichen mit hoher Smartphone-Nutzung. Die Befunde sind korrelational, nicht kausal, und die spezifische Übertragung auf GenKI-vermittelte digitale Exposition bleibt empirisch zu erweisen. Sie sind jedoch die besten verfügbaren Stellvertreterbelege und weisen konsistent in eine Richtung: das jugendliche Gehirn ist besonders anfällig gegenüber dem Profil kognitiver Belastung, das

grenzgängige KI-Werkzeuge, ähnlich wie andere intensiv einbindende digitale Systeme, präsentieren.

3.4 Die internationale Politik- und Kompetenzlandschaft

Das internationale Politikumfeld hat sich bemerkenswert schnell konvergiert. Kapitel 7 entwickelt dies ausführlich; hier merke ich nur an, dass bis Mitte 2026 mindestens sieben große Kompetenz- oder Politikrahmen in Kraft sind, die jeweils die institutionelle Bereitstellung von KI in der Bildung einschränken oder leiten. Dazu gehören die KI-Verordnung der EU (2024); die UNESCO-KI-Kompetenzrahmen für Schüler und Lehrkräfte (2024); der EU-OECD-AILit-Rahmen für die Primar- und Sekundarbildung (2025), der in die PISA-2029-MAIL-Erhebung einfließt (OECD, 2026); der KI-Kompetenzrahmen des US-Arbeitsministeriums; das Understand-Use-Evaluate-Rahmenwerk von Digital Promise; der KI-Kompetenzrahmen des Digital Education Council; und entstehende nationale Rahmen in Singapur (Chew, Soon, & Kaur, 2026), den Vereinigten Staaten (California Department of Education, 2025) und anderswo. Die Konvergenz ist kein Zufall. Über die Rahmen hinweg wiederholen sich vier Domänen mit hoher Konsistenz: kritisches Einbringen, Gestaltung und Zusammenarbeit, Verwaltung und Ethik sowie Design- und Systemdenken.

Das in diesem Papier angebotene 4D-Rahmenwerk ist mit diesen konvergierenden internationalen Rahmen strukturell kompatibel. Wo die internationalen Rahmen die Kompetenzergebnisse definieren, die ein Lernender erreichen soll, definiert das 4D-Rahmenwerk die kognitiven Routinen, die ein Lernender aufbauen muss, um diese Kompetenzen zu erreichen. Beide operieren auf unterschiedlichen Ebenen pädagogischer Spezifikation. Tabelle 4 in Kapitel 7 bildet die Beziehung zwischen dem 4D-Rahmenwerk und den vier AILit-Domänen explizit ab.

4. Das 4D-Rahmenwerk · Delegation, Beschreibung, Urteilsvermögen, Sorgfalt

Ich schlage ein vierstufiges kognitives Rahmenwerk für die Mensch-KI-Interaktion vor und wende es in meiner eigenen Lehrpraxis an IFSA Prag, CEA CAPA und der Anglo-American University Prague an. Die vier Stufen sollen entwicklungsbezogen und pädagogisch sequenziell sein und auf distinkte kognitive Routinen abbilden, die in dieser Reihenfolge, mit angemessenem Lerngerüst, über den K-12- und tertiären Lehrplan hinweg gezielt vermittelt werden müssen. Das Rahmenwerk ist keine Kompetenztaxonomie in dem Sinne, in dem es die AILit-, UNESCO- oder Digital-Promise-Rahmen sind. Es ist ein Rahmen kognitiver Routinen. Es spezifiziert, was ein Lernender im Sinne der zuverlässigen Ausführung einer spezifischen kognitiven Operation in jeder Entwicklungsphase tun lernen muss. Es soll unterhalb der Kompetenzrahmen als kognitives Substrat fungieren, aus dem diese Kompetenzen aufgebaut werden.

Die vier Stufen, kurz: Delegation ist die bewusste Wahl, welche kognitiven Aufgaben einem Werkzeug zuzuweisen sind und welche man für sich behält, getroffen im Bewusstsein der Begründung der Wahl. Beschreibung ist die Fähigkeit, einem System ein Problem, einen Kontext und einen gewünschten Output mit hinreichender Präzision zu spezifizieren, sodass die Antwort des Systems tatsächlich nützlich ist, und zu erkennen, wann die Beschreibung selbst versagt hat. Urteilsvermögen ist die Fähigkeit, KI-Output kritisch zu bewerten, Fehler, Halluzinationen, fehlenden Kontext und unangemessene Sicherheit zu erkennen, ohne in gläubige Akzeptanz oder reflexhafte Ablehnung abzugleiten. Sorgfalt ist das verantwortliche Nacharbeiten, das die Schleife schließt:

Überprüfung, Zuschreibung, ethische Reflexion, Verantwortlichkeit für Ergebnisse. Jede Stufe baut auf der vorherigen auf. Jede kann gelehrt, mit Gerüst gestützt und bewertet werden. Jede erfordert in verschiedenen Entwicklungsphasen unterschiedliche pädagogische Ansätze.

4.1 Delegation

Delegation ist die bewusste Wahl, welche kognitiven Aufgaben einem Werkzeug zuzuweisen sind und welche man für sich behält, getroffen im Bewusstsein der Begründung der Wahl. Es ist die erste kognitive Operation in jeder Mensch-KI-Interaktion, ausgeführt jedes Mal, wenn ein Lernender nach dem Werkzeug greift, und es ist die Operation, die am anfälligsten für unreflektierte Voreinstellung ist. Die pädagogisch interessante Frage lautet nicht, ob Studierende delegieren. Alle kognitiv funktionstüchtigen Menschen delegieren ständig. Die Frage lautet, ob sie bewusst delegieren, mit bewusster Wahl und Bewusstsein der Folgen für das eigene Lernen, oder ob die Delegation unbewusst, voreingestellt und substitutiv ist.

Die in Kapitel 3.1 zitierte empirische Arbeit legt nahe, dass der Standardmodus für ungeführte Lernende eine unreflektierte Überdelegation ist. Kosmyna et al. (2025) und Fan et al. (2025) dokumentieren beide Muster, in denen Lernende dem System nicht nur mechanische Aufgaben (Formatieren, Korrekturlesen, Erzeugen von Boilerplate-Text) überlassen, sondern gerade jene kognitiven Operationen, zu deren Entwicklung die Aufgabe in ihnen gedacht war (Setzung von Prämissen, Konstruktion von Argumenten, Auswahl von Belegen). Delegation zu lehren bedeutet daher, die explizite kognitive Routine zu lehren, vor dem Griff zum Werkzeug innezuhalten, sich selbst zu artikulieren, welche kognitive Operation die Aufgabe entwickeln soll, und auf dieser Grundlage zu entscheiden, ob sie zu delegieren ist. Dies ist eine metakognitive Fertigkeit. Sie muss explizit gelehrt werden. Sie wird nicht spontan aus der Exposition gegenüber dem Werkzeug entstehen.

4.2 Beschreibung

Beschreibung ist die Fähigkeit, einem System ein Problem, einen Kontext und einen gewünschten Output mit hinreichender Präzision zu spezifizieren, sodass die Antwort des Systems tatsächlich nützlich ist, und zu erkennen, wann die Beschreibung selbst versagt hat. In der populären Form nennt man dies Prompting, aber der bildungsbezogene Einsatz ist breiter und älter. Die Disziplin, ein Problem zu beschreiben (einer Lehrkraft, einem Mitlernenden, sich selbst, einer Maschine), ist eine grundlegende kognitive Fertigkeit. Sie verbindet sich mit der philosophisch-pädagogischen Tradition der Artikulation, der rhetorischen Tradition der Publikumsanalyse und der praktischen Tradition der Spezifikation in jeder technischen Disziplin. GenKI verlangt sie in einer besonders expliziten Form. Ein Lernender, der nicht beschreiben kann, was er will, erhält flüssigen Unsinn zurück; ein Lernender, der beschreiben kann, erhält etwas Nützliches zurück. Die zentrale pädagogische Einsicht ist, dass dies keine neue Fertigkeit ist. Es ist die explizite Form der impliziten Artikulationsfertigkeit, die alle gute Bildung schon immer gepflegt hat.

Beschreibung ist auch die Phase, in der die Grenzen des eigenen Verständnisses spürbar werden. Ein Studierender, der einem Modell nicht artikulieren kann, was er eigentlich wissen will, entdeckt oft zum ersten Mal, dass die zugrunde liegende kognitive Operation (zu wissen, was man wissen will) unvollständig ist. Diese diagnostische Funktion der Beschreibung, in der das Versagen der Beschreibung das vorgängige Versagen der Artikulation offenlegt, macht sie zu einer der pädagogisch wertvollsten Phasen des

Rahmenwerks. Die Aufgabe der Lehrkraft in dieser Phase besteht darin, den diagnostischen Moment hervorzubringen, nicht ihn zu umgehen.

4.3 Urteilsvermögen

Urteilsvermögen ist die Fähigkeit, KI-Output kritisch zu bewerten. Hier schneidet sich KI-Flüssigkeit mit klassischem kritischem Denken, epistemologischer Bildung und Quellenbewertung. Eine urteilsfähige Lernende liest KI-Output, wie eine kompetente Forscherin eine Sekundärquelle liest: sucht nach Belegen, prüft Zitate, vermerkt selbstsichere Aussagen über umstrittene Angelegenheiten, identifiziert, was die Quelle wahrscheinlich nicht wissen kann. Die spezifische Herausforderung, die GenKI hier aufwirft, ist das Problem der Oberflächensicherheit. Grenzgängige Modelle produzieren flüssigen, syntaktisch geschliffenen Output, der Sicherheit signalisiert, ungeachtet des zugrunde liegenden epistemischen Status. Die Urteilsroutinen, die etwa beim Bewerten eines Wikipedia-Artikels funktionierten, in dem Unsicherheit oft durch Satzstruktur, Zitierdichte und Ton signalisiert wird, übertragen sich nicht sauber auf KI-Output, wo die Flüssigkeit einheitlich ist.

Urteilsvermögen ist eng an Domänenwissen gebunden. Eine Lernende kann nur in Domänen urteilen, in denen sie über genügend Sachwissen zur Bewertung verfügt. Dies erzeugt ein pädagogisches Paradox. Um KI gut zu nutzen, brauchen Lernende Domänenwissen. Um Domänenwissen in einem Schulsystem zu erwerben, das zunehmend von KI vermittelt wird, müssen Lernende vor der vorzeitigen KI-Substitution genau jener Routinen geschützt werden, die Domänenwissen aufbauen. Dies ist eines der stärksten Argumente für eine asymmetrische KI-Exposition über die Entwicklungsstufen hinweg (Kapitel 6) und gegen einheitliche Erlaubnispolitik, die den KI-Zugang als entwicklungsneutral behandelt.

4.4 Sorgfalt

Sorgfalt ist das verantwortliche Nacharbeiten, das die Schleife schließt. Überprüfung von faktischen Behauptungen. Zuschreibung dort, wo Zuschreibung geboten ist. Offenlegung dort, wo institutionelle Normen oder Redlichkeit es verlangen. Reflexion darüber, was gelernt, was ersetzt, was umgangen wurde. Ethische Verantwortlichkeit für die Ergebnisse des KI-gestützten Prozesses. Sorgfalt unterscheidet die kompetente Operateurin von der flüssigen Nutzerin. Sie ist die Phase, in der sich KI-Flüssigkeit mit akademischer Integrität, professioneller Ethik und in regulierten Branchen mit rechtlicher Verantwortlichkeit überschneidet. Sie ist meines Erachtens die Phase, die in der gegenwärtigen Bildungspraxis am konsistentesten vernachlässigt wird, einerseits, weil sie die am wenigsten aufregende ist (Überprüfung, Offenlegung, Zuschreibung sind administrativ langweilige Operationen), und andererseits, weil sie die höchsten kognitiven Kosten auferlegt.

Zusammengenommen bieten die vier Stufen eine Entwicklungssequenz: eine Lernende, die bewusst delegieren, präzise beschreiben, kritisch urteilen und mit Sorgfalt nacharbeiten kann, operiert auf dem, was ich als KI-Flüssigkeit bezeichne. Entscheidend ist, dass Flüssigkeit nicht gleichbedeutend mit Erlaubnis ist. Einer Lernenden kann die Nutzung von KI gestattet sein, ohne dass sie flüssig ist (die Voreinstellung an den meisten tertiären Institutionen heute). Einer Lernenden kann die Erlaubnis verwehrt sein, KI zu nutzen, ohne dass sie an der Entwicklung von Flüssigkeit gehindert wird (die Lücke zwischen formaler Politik und Schattenlehrplan). Und eine Lernende kann entwicklungsbezogen zu jung sein, um in irgendeiner der vier Stufen volle Flüssigkeit zu entwickeln; in diesem Fall besteht die pädagogische Aufgabe darin, die kognitiven

Vorläufer-Routinen aufzubauen, auf denen Flüssigkeit letztlich aufgebaut werden wird. Das ist es, was das Rahmenwerk von Bildungsinstitutionen ernsthaft fordert. Flüssigkeit ist keine Erlaubnisentscheidung (Sollten wir KI erlauben? Ja/Nein). Sie ist keine Werkzeugwahlentscheidung (Welche KI? GPT/Claude/Gemini). Sie ist keine Erkennungsentscheidung (Wie erkennen wir KI-Nutzung? KI-Detektor/mündliche Prüfung/im Unterricht). Sie ist eine Entscheidung über das kognitive Curriculum (Welche kognitiven Routinen sollten Lernende entwickeln, in welcher Reihenfolge, in welchem Alter, unter welcher Aufsicht, mit welcher Verantwortlichkeit). Die entwicklungsgestufte Sequenzierung dieser vier Phasen, nicht die binäre Entscheidung, ist das, was das Rahmenwerk von Institutionen ernsthaft fordert.

5. Entwicklungspsychologische Überlegungen · Die Neurowissenschaft der KI-Exposition

Der Fall für eine asymmetrische KI-Exposition über die Entwicklungsstufen hinweg stützt sich auf drei konvergierende Bereiche von Evidenz: die Entwicklungsneurowissenschaft der Reifung exekutiver Funktionen, die Literatur zu kognitiver Belastung und Selbstregulation in der Adoleszenz und die entstehende Evidenz zu Auswirkungen intensiver digitaler Exposition bei Kindern und Jugendlichen. Keine dieser Literaturen wurde spezifisch mit Blick auf GenKI entwickelt. Zusammengenommen begründen sie jedoch einen hinreichenden Prima-facie-Fall, um die institutionelle Standardhaltung einzuschränken.

Der präfrontale Kortex (PFC), Sitz exekutiver Funktionen, Urteilsvermögen und inhibitorischer Kontrolle, erreicht seine ausgereifte Myelinisierung nicht vor Mitte der zwanziger Jahre (Diamond, 2013). Die Arbeitsgedächtniskapazität, die kognitive Ressource, die komplexes Denken, anhaltende Aufmerksamkeit und die Integration neuer Informationen in bestehende Schemata unterstützt, ist in Kindheit und Adoleszenz strukturell begrenzt und verbessert sich nur durch gezielte Übung an kognitiv anspruchsvollen Aufgaben. Metakognition, die Fähigkeit, eigene kognitive Prozesse zu überwachen und zu regulieren, entwickelt sich über die Schuljahre hinweg langsam und ist besonders anfällig gegenüber Umweltbedingungen, die externe Regulation an die Stelle interner Regulation setzen. Jede dieser Kapazitäten ist genau das, was flüssige GenKI-Nutzung erfordert. Jede dieser Kapazitäten ist auch das, was GenKI-Nutzung am leichtesten ersetzen kann, so wie ein Exoskelett die Muskelbildung ersetzt.

Das Risiko ist daher nicht abstrakt. Es gibt keinen plausiblen Entwicklungsmechanismus, durch den ein Kind, dessen erste Begegnung mit einer komplexen kognitiven Aufgabe eine KI-generierte Lösung dafür ist, dieselben kognitiven Routinen ausbildet wie ein Kind, das zunächst mit der Aufgabe ringt und der KI später als Überprüfungs- oder Erweiterungswerkzeug begegnet. Das erste Kind erhält die Antwort. Das zweite Kind baut den Muskel auf, der Antworten produziert. Der präfrontale Kortex des ersten Kindes hat keinen entwicklungsbezogenen Grund, jene neuronale Verschaltung anzulegen, die anhaltendes Ringen aufgebaut hätte. Dies ist das Substitutionsrisiko, und es ist gerade in jenem Entwicklungsfenster am akutesten, in dem die exekutive Funktion am plastischsten ist.

Empirische Befunde zu den Auswirkungen digitaler Exposition in diesem Fenster sind konsistent genug, um Vorsicht zu rechtfertigen, selbst dort, wo direkte GenKI-Evidenz nicht verfügbar ist. Hutton et al. (2022) berichten von Assoziationen zwischen bildschirmgestützter Mediennutzung und reduzierter Integrität der weißen Substanz bei Kindern im Vorschulalter, mit messbaren Auswirkungen auf Trakte, die Sprache und

entstehende Literalität unterstützen. Lewin et al. (2025) berichten von Längsschnitteffekten bildschirmgestützter Medienaktivitäten auf Gehirnfunktion, -architektur und psychische Gesundheit in der frühen Adoleszenz, mit einer durchschnittlichen Reduktion der PFC-Dicke um 0,82 mm bei Kindern mit hoher Bildschirmexposition gegenüber gematchten Kontrollen, ein Unterschied, der mit messbaren Einbußen exekutiver Funktionen assoziiert ist. León-Domínguez (2024) bespricht potenzielle kognitive Risiken generativer transformerbasierter KI-Chatbots für höhere exekutive Funktionen und vermerkt insbesondere das Risiko substitutiver Nutzung während Entwicklungsperioden hoher neuronaler Plastizität. Gascón-Cánovas et al. (2025) berichten von einer differentiellen Leistung exekutiver Funktionen bei Jugendlichen mit problematischer Smartphone-Nutzung, mit Effekten, die auf Eltern-berichteten Maßen erkennbar sind, selbst dort, wo kontrollierte Laboraufgaben sie unterschätzen können. Hinze et al. (2025) zeigen in einer Längsschnittstudie an 8 072 britischen Jugendlichen stabile bidirektionale Assoziationen zwischen der Entwicklung exekutiver Funktionen und Ergebnissen der psychischen Gesundheit über die Altersspanne 11–15 hinweg.

Reggente et al. (2026) dokumentieren in einem Scoping-Review von 104 Studien (Pediatric Reports), der die Forschung der Jahre 2022–2025 zur Smartphone-Nutzung und Gehirnfunktion von Jugendlichen abdeckt, reduzierte funktionelle Konnektivität in kortiko-subkortikalen Schaltkreisen, reduziertes Volumen der grauen Substanz im oberen und unteren frontalen Gyrus und aberrante kortikale Aktivierung in frontoparietalen Regionen während kognitiver Konfliktaufgaben bei Jugendlichen mit hoher Smartphone-Nutzung. Über den Review hinweg ergeben sich fünf wiederkehrende Themen: psychische Störungen (n = 52 Studien), Schlaf (n = 43), sozioemotionale Funktion (n = 23), exekutive Funktion (n = 14) und sensorische Verarbeitung (n = 1). Die methodologischen Vorbehalte aus Kapitel 3.3 gelten: es sind korrelationale Befunde zu intensiver digitaler Exposition im Allgemeinen, nicht spezifisch zu GenKI. Sie sind jedoch die besten verfügbaren Stellvertreterbelege und weisen konsistent in die Richtung, dass das jugendliche Gehirn besonders anfällig gegenüber dem Profil kognitiver Belastung ist, das intensiv einbindende digitale Systeme, von denen grenzgängige KI nun die kognitiv einbindendste ist, präsentieren.

Die Implikation lautet nicht, dass Bildschirme, digitale Medien oder KI-Werkzeuge Schaden in einem einfachen Sinne verursachen. Die Implikation lautet, dass kognitive Belastung, Selbstregulationsanforderung und Belohnungsstruktur nicht triviale Variablen sind und dass das sich entwickelnde Gehirn gegenüber ihrer Fehlkalibrierung anfälliger ist als das erwachsene Gehirn. Generative KI führt alle drei Variablen in ungewöhnlich konzentrierter Form ein: sie lagert kognitive Belastung unmittelbar aus, verlangt Selbstregulation, um konstruktiv und nicht substitutiv genutzt zu werden, und ihre Belohnungsstruktur (unmittelbarer, flüssiger, bestätigender Output) ist genau das Profil, das die inhibitorische Kontrolle in jugendlichen Populationen kompromittiert. Jede dieser Eigenschaften ist eigenständig gut dokumentiert als entwicklungsbezogener Risikofaktor. GenKI kombiniert sie.

Aus diesen neurodevelopmentalen Realitäten ergeben sich drei pädagogische Prinzipien.

(i) Asymmetrische Exposition. Intensität, Autonomie und Breite der GenKI-Integration sollten mit der kognitiv-entwicklungsbezogenen Bereitschaft skalieren, nicht allein mit dem Jahrgang. Ein Zwölfjähriger und ein Achtzehnjähriger sollten nicht unter derselben KI-Zugangspolitik operieren. Die Politik sollte auf exekutive Funktion, Arbeitsgedächtniskapazität und metakognitive Reife kalibriert sein. Wo Messung

unpraktisch ist (wie sie es im institutionellen Maßstab im Allgemeinen ist), sind angemessene Stellvertreter (chronologisches Alter, Entwicklungsstufe) vertretbar, aber die Politik muss sich zwischen ihnen tatsächlich unterscheiden.

(ii) Substitutionsrisiko vor Gerüststreife. Bevor Lernende die kognitiven Routinen verinnerlicht haben, deren Erweiterung ein Werkzeug dienen soll, droht das Werkzeug die Routine zu ersetzen, statt sie zu erweitern. Die pädagogische Sequenzierung muss daher KI-Werkzeuge erst nach, nicht vor der entwicklungsbezogenen Konsolidierung der zugrunde liegenden Fertigkeit einführen. Dies ist das stärkste Argument für schützende Haltungen in der Primarbildung und für gerüstgestützte Haltungen in der Sekundarbildung.

(iii) Reibung ist pädagogisch wertvoll. Der Aspekt von GenKI, der sie kommerziell attraktiv macht (reibungslöse Antwortgenerierung), ist der Aspekt, der sie pädagogisch riskant macht. Wirksame Bildungsintegration führt gezielt Reibung (Überprüfung, Reflexion, Offenlegung, Diskussion gegen das Modell) in jeder Phase wieder ein. Die konstruktionistische Pädagogik (Lin, Dai, & Ng, 2025) bietet einen gut erforschten Mechanismus: Lernende müssen Artefakte mit KI konstruieren, nicht Outputs von KI entgegennehmen. Projektbasiertes Lernen (Thomas, 2000) und Traditionen des computationalen Denkens (Wing, 2006) bieten ergänzende Mechanismen.

6. Pädagogischer Ansatz nach Entwicklungsstufe

Ich schlage drei strukturell distinkte Integrationsmodelle über das Spektrum K-12 plus Tertiärbildung hinweg vor. Es handelt sich nicht um graduelle Abstufungen desselben Ansatzes. Sie unterscheiden sich der Art nach, kalibriert nach dem, was jede Stufe von Lernenden zu leisten vermag, was sie tun sollten und was geschützt werden sollte. Tabelle 1 fasst die vorgeschlagene Haltung über die drei Stufen hinweg zusammen.

Tabelle 1. Entwicklungsbezogen abgestimmte Haltung über die Bildungsstufen.

Stufe	Alter	KI-Haltung	Primärer pädagogischer Fokus
Primär	5–11	KI-nahe Bildung. Kein direkter LLM-Zugang im Klassenraum.	Grundlegendes Lesen, Schreiben, Rechnen, Aufmerksamkeit, soziale Kognition. Kritisches Denken durch Lehrkraft modelliert.
Sekundär	12–17	Gerüstgestützte KI-Kompetenz mit strukturierter Exposition unter Aufsicht der Lehrkraft.	Alle vier 4D-Stufen explizit gelehrt. Urteilsvermögen und Sorgfalt tragend.
Tertiär	18+	Operative KI-Flüssigkeit. Domänenspezifische Integration.	Professionelles Prompting, Modellauswahl, agentische Systeme, ethischer Einsatz, Offenlegungsnormen.

6.1 Primarbildung (5–11 Jahre) · KI-nahe Bildung

Die vertretbare Haltung in der Primarbildung ist der Schutz grundlegender Kognition. Kinder in diesem Entwicklungsfenster konsolidieren noch die Kernfertigkeiten (Lese-flüssigkeit, arithmetische Automatisierung, schriftlicher Selbstausdruck, anhaltende Aufmerksamkeit, Peer-Kooperation), von denen alle nachfolgende kognitive Arbeit abhängt. Diese Fertigkeiten werden durch produktives Ringen aufgebaut. Substitutives kognitives Auslagern in dieser Phase ist kein pädagogisches Risiko, es ist ein pädagogisches Versagen. Ein Kind, dessen erste Erfahrung mit der Erzeugung von schriftlichem Text die Erzeugung mit einem KI-Assistenten ist, hat nicht schreiben

gelernt. Es hat gelernt, zu beauftragen. Die neuronale Verschaltung des Schreibens wird durch den Akt des Schreibens gelegt, nicht durch den Akt des Überarbeitens KI-generierten Textes.

Das Rahmenwerk, das ich in dieser Stufe empfehle, nenne ich KI-nahe Bildung. Kinder lernen über KI, was sie ist, was sie kann und was nicht, warum sie manchmal selbstsicher klingt und sich irrt, dass sie von Menschen gemacht ist und Entscheidungen widerspiegelt, die Menschen getroffen haben, ohne dass sie direkt mit grenzgängigen Modellen im Klassenraum interagieren. Konkrete pädagogische Praktiken in dieser Stufe umfassen lehrermodellerte KI-Interaktion in altersgerechten Domänen (die Lehrkraft demonstriert, erzählt das Vorgehen, legt offen, was sie zu behalten gegenüber dem Delegieren wählt), vereinfachte Übungen algorithmischen Denkens (Dekomposition, Mustererkennung, Abstraktion im Sinne von Wing, 2006), Einführung in die sozialen und ethischen Dimensionen automatisierter Entscheidungsfindung in altersgerechten Formen und Gespräche darüber, wie das Kind zwischen maschinell und menschlich generiertem Inhalt unterscheidet.

Konstruktionistische Ansätze, die in der K-12-AI-Kompetenzforschung vor GenKI entwickelt wurden (Lin, Dai, & Ng, 2025), übertragen sich sauber auf diese Stufe. Wo die ältere konstruktionistische Tradition Kinder Roboter und einfache ML-Klassifikatoren bauen ließ, um über KI zu lernen, erweitert die GenKI-Ära dies darauf, dass Kinder kleine Simulationen und Spiele bauen, die KI-Charaktere enthalten, deren Verhalten das Kind spezifiziert und befragt. Su, Ng und Chu (2025) identifizieren in einem systematischen Review von 47 empirischen Studien zur K-12-AI-Kompetenz Konstruktionismus, projektbasiertes Lernen und verkörperte Kognition als die drei Ansätze mit der stärksten evidenzbasierten Unterstützung für grundlegende KI-Kompetenz in dieser Stufe. Entscheidend ist, dass keiner dieser Ansätze erfordert, dass Kinder direkt mit grenzgängigen Verbraucher-LLMs interagieren.

Diese Haltung ist nicht technologiefeindlich. Sie ist entwicklungsbezogen angemessene Sequenzierung. Kinder in dieser Stufe lernen Lesen durch das Lesen von Büchern, die auf ihre kognitive Stufe kalibriert sind, nicht indem ihnen Shakespeare in die Hand gedrückt wird. Das Prinzip ist dasselbe. Direkte GenKI-Exposition mag im oberen Bereich dieser Stufe (9–11 Jahre) eine legitime Rolle für spezifische, beaufsichtigte, gerüstgestützte Verwendungen haben, insbesondere im Dienste der Vermittlung von Beschreibung und ersten Phasen des Urteilsvermögens. Die Standardhaltung sollte jedoch Schutz sein, und die Beweislast sollte auf jenen liegen, die direkte unbeaufsichtigte Nutzung befürworten, nicht auf jenen, die das Entwicklungsfenster schützen.

6.2 Sekundarbildung (12–17 Jahre) · Gerüstgestützte KI-Kompetenz

Die Sekundarbildung ist das entwicklungsbezogen kritische Fenster. Lernende in dieser Stufe sind kognitiv in der Lage, grenzgängige KI-Werkzeuge zu nutzen (Arbeitsgedächtnis, exekutive Funktion und metakognitive Kapazitäten sind hinreichend entwickelt), aber ihre kognitiven Routinen konsolidieren sich noch. Das jugendliche Gehirn befindet sich zugleich auf dem Höhepunkt seiner Anfälligkeit gegenüber den Eigenschaften der Belohnungsstruktur, die inhibitorische Kontrolle kompromittieren (Steinberg, 2008; Reggente et al., 2026). Dies ist die Stufe, in der das volle Gewicht des Rahmenwerks zur Anwendung kommen muss. Wird diese Stufe falsch gestaltet, sind die institutionellen Kosten im gesamten Qualifikationssystem sichtbar. Wird sie richtig gestaltet, kommen die resultierenden tertiären Jahrgänge mit den kognitiven Grundlagen an, die operative KI-Flüssigkeit verlangt.

Alle vier 4D-Stufen sollten über das Sekundarcurriculum hinweg explizit gelehrt werden, mit zunehmender Komplexität. Die Arbeit zur Delegation beginnt mit metakognitiven Übungen: der Lernende artikuliert vor jeder KI-Interaktion, welche kognitive Operation die Aufgabe in ihm entwickeln soll, und entscheidet auf dieser Grundlage, ob er sie delegiert. Die Arbeit zur Beschreibung betont iterative Verfeinerung: Lernende lernen zu bewerten, ob eine Modellantwort ihre Frage tatsächlich beantwortet hat, zu identifizieren, welche Merkmale der Beschreibung versagt haben, und zu überarbeiten. Urteilsvermögen ist die pädagogisch intensivste Stufe auf dieser Ebene. Lernende lernen, KI-Outputs gegen autoritative Quellen zu überprüfen, Halluzinationen und unangemessene Sicherheit zu erkennen, die Grenzen des Modellwissens (Cutoff des Trainings, Verzerrungen der Trainingsdaten, domänenspezifische Schwäche) zu erkennen und produktiv mit dem Modell zu widersprechen. Sorgfalt wird durch explizite Zuschreibungsnormen, Offenlegungsanforderungen und Reflexion darüber, was ersetzt und was aufgebaut wurde, gelehrt.

Die pädagogische Struktur, die dies unterstützt, ist das, was ich als gerüstgestützte KI-Kompetenz bezeichne. KI-Werkzeuge werden gezielt in das Curriculum eingeführt, mit den zugrunde liegenden Fertigkeiten, deren Erweiterung sie dienen sollen, sequenziert, mit expliziter Anleitung der Lehrkraft hinsichtlich der kognitiven Operation, die der Lernende ausführen sollte, und der Operation, die die KI ausführen sollte. Das Klassenraumdesign, das dies unterstützt, umfasst strukturierte KI-Interaktionen (klar umrissene Aufgaben, Zeit und Ziel), Lehrkraftüberprüfung von KI-Interaktionen (stichprobenhaft, nicht überwacht) und explizite Unterweisung in Prompting, Überprüfung und Offenlegung als übertragbare Fertigkeiten.

Zwei weitere pädagogische Verpflichtungen unterscheiden diese Stufe. Erstens, Bewertungsneudesign: traditionelle häusliche und unbeaufsichtigte Bewertungen messen die kognitiven Operationen, zu deren Messung sie konzipiert wurden, nicht mehr zuverlässig. Bewertung muss sich in Richtung prozessorientierter Belege (Lerntagebücher, Versionshistorie, Think-Aloud-Protokolle), leistungsbasierter Aufgaben (kollaborative Problemlösung, Designherausforderungen) und unterrichtsinterner beaufsichtigter Demonstration von Routinen, die über das Curriculum aufgebaut wurden, verlagern (ATC21S-Rahmen, im OECD Learning Compass 2030). Zweitens, transparente Offenlegungsnormen: Lernende lernen ab der ersten Einführung von KI-Werkzeugen, dass die Offenlegung der KI-Nutzung erwartet, normal und selbst Teil der Bewertung ist, nicht ein zu verbergendes Geständnis. Die Offenlegungsrahmung verwandelt die Arbeit an akademischer Integrität von Polizeiarbeit in Pädagogik.

Roe, Furze und Perkins (2026) schlagen in ihrem metaphorischen Rahmenwerk des digitalen Plastiks vor, dass KI-Kompetenz auf dieser Stufe explizit als kritische KI-Kompetenz gerahmt werden sollte: Lernende lernen, KI nicht nur als Werkzeug, sondern als soziokulturelles Artefakt zu bewerten, das in Trainingsdaten, Geschäftsmodell und politisch-ökonomischen Kontext eingebettet ist. Diese Rahmung passt natürlich zum 4D-Rahmenwerk. Urteilsvermögen wird nicht nur zu Output-Bewertung, sondern auch zu System-Bewertung. Sorgfalt umfasst ethische Reflexion über die Existenz der KI, nicht nur über die spezifischen Outputs, die der Lernende verwendet hat.

6.3 Tertiär- und Hochschulbildung (18+) · Operative KI-Flüssigkeit

Tertiärlernende sind kognitiv für volle KI-Flüssigkeit positioniert. Die pädagogische Aufgabe in dieser Stufe ist operative Beherrschung innerhalb einer Domäne. Lernende sollten ihr Studienprogramm in der Lage verlassen, grenzgängige KI-Werkzeuge auf professionellem Niveau innerhalb ihrer Disziplin zu bedienen, einschließlich

domänenspezifischem Prompting, Modellauswahl, mehrstufiger Aufgabendekomposition für agentische Systeme und ethischem Einsatz innerhalb professioneller Normen. Tertiäre Curricula sollten KI-Nutzung als gestufte, bewertete Kompetenz in jeder Disziplin integrieren, in der sie professionell relevant ist, was bis 2026 im Wesentlichen jede Disziplin ist.

Konkrete operative Komponenten umfassen: domänenspezifisches Prompt Engineering (die Konventionen, Quellen und Bewertungskriterien, die für die Disziplin spezifisch sind), Modellauswahlkompetenz (wann welches Modell zu verwenden ist, welche Abwägungen bestehen, wie Eignung zu bewerten ist), agentisches und mehrstufiges Workflow-Design (wie Werkzeugnutzung über mehrere Aufrufe hinweg zu komponieren ist, wie Überprüfung in den Workflow einzubauen ist), professionelle Ethik und Offenlegungsnormen (disziplinspezifisch, kalibriert auf Berufsverbände und entstehende Regulierung) und die operativen Implikationen der Regulierung einschließlich der KI-Verordnung der EU für Absolventen, die in KI-einsetzende Branchen eintreten (Europäische Union, 2024).

Die empirische Evidenz zur tertiären KI-Integration stützt vorsichtigen Optimismus hinsichtlich der Lernergebnisse, wenn die Integration strukturiert ist. Cunningham et al. (2025) berichten in ihrer einjährigen Tennessee-Studie von erheblichen positiven Effekten auf die Motivation und den Kompetenzerwerb von Postgraduierten, wenn die ChatGPT-Integration mit expliziter instruktionaler Rahmung gepaart ist. Raptopoulou (2025) berichtet, dass ChatGPT die Entwicklung akademischer Literalität unterstützt, wenn er mit klaren pädagogischen Zielen integriert wird. Die Metaanalyse von Deng et al. (2026) berichtet von mittleren positiven Effekten auf die Lernergebnisse über 35 Studien hinweg. Die Einschränkung lautet in jedem Fall gleich: positive Effekte hängen von strukturierter, gerüstgestützter Integration ab, nicht von permissiver Voreinstellung. Unstrukturierter Zugang produziert die metakognitive Trägheit, die Fan et al. (2025) dokumentieren; strukturierte Integration produziert die operative Flüssigkeit, von der Cunningham et al. (2025) berichten.

Die pädagogische Sequenzierung über die Tertiärbildung hinweg sollte daher sein: explizite Unterweisung im 4D-Rahmenwerk im ersten Studienjahr (falls nicht bereits in der Sekundarbildung erworben); domänenspezifische Integration über die mittleren Jahre (das 4D-Rahmenwerk innerhalb der epistemischen, methodologischen und ethischen Konventionen der Disziplin angewandt); und Abschluss- oder Examensarbeit, die operative Flüssigkeit explizit durch realistische professionelle Aufgaben bewertet. Das Modell ähnelt eher einem klinischen Praktikum oder einer Studio-Review als einer traditionellen schriftlichen Prüfung. KI-Flüssigkeit ist operationalisierte Kompetenz, demonstriert unter realistischen Bedingungen, mit derselben Berufsethik und Verantwortlichkeit, die der Lernende benötigen wird, auch in regulierten Domänen, in denen der Standard hoch ist.

6.4 Beispielaktivitäten und operative Beispiele nach Stufe

Abstrakte Rahmenwerksspezifikationen sind notwendig, aber für die institutionelle Umsetzung unzureichend. Dieser Unterabschnitt bietet konkrete Beispiele für Aktivitäten in jeder Stufe, die das 4D-Rahmenwerk in der Klassenraumpraxis operationalisieren. Die Beispiele stammen aus meiner eigenen Lehrpraxis, der in Kapitel 3 besprochenen Literatur und den Klassenraumszenarien des AILit-Rahmens (Europäische Kommission & OECD, 2025). Sie sind nicht präskriptiv. Sie sollen die Art von pädagogischem Design illustrieren, die das Rahmenwerk verlangt, an dem institutionelle Curriculumsarbeit getestet und verfeinert werden kann. Tabelle 3 fasst einen repräsentativen Aktivitätssatz

zusammen.

Tabelle 3. Repräsentative Klassenraumaktivitäten, die das 4D-Rahmenwerk operationalisieren.

Stufe	4D-Phase	Beispielaktivität	Pädagogisches Ziel
Primär (9–11)	Urteilsvermögen	Lehrergeführte Inspektion einer KI-Antwort, die einen selbstsicheren Sachfehler enthält. Kinder identifizieren den Fehler anhand einer gedruckten Primärquelle.	Feststellen, dass selbstsicher klingender Output kein Beleg für Richtigkeit ist.
Primär (9–11)	Delegation (Vorläufer)	Klassengespräch: Sollten wir die KI bei der Rechtschreibung um Hilfe bitten, oder es zuerst selbst herausfinden? Wann ergibt was Sinn?	Bewusste Wahl als lernbare Fertigkeit vor jeglicher Werkzeugnutzung ans Licht bringen.
Sekundär (12–14)	Beschreibung	Schüler versuchen, ein KI-Werkzeug eine nützliche Antwort auf eine Forschungsfrage erzeugen zu lassen, mit drei dokumentierten und diskutierten Iterationen der Prompt-Verfeinerung.	Artikulation als übertragbare kognitive Fertigkeit.
Sekundär (15–17)	Urteilsvermögen	Übung zur Überprüfung KI-generierter Quellenlisten. Jeder Schüler erhält eine KI-generierte Bibliografie und muss überprüfen, welche Zitate real sind und welche Halluzinationen.	Operative Überprüfungsroutine aufbauen.
Sekundär (15–17)	Sorgfalt	Prozessportfolio: jede KI-gestützte Aufgabe enthält ein strukturiertes Offenlegungsprotokoll, das angibt, was delegiert wurde, warum und was überprüft wurde.	Offenlegung und Reflexion als normale Praxis habituierten.
Tertiär	Delegation	Disziplinspezifische Entscheidungsmatrix: Studierende entwickeln eigene Kriterien dafür, wann welches Modell mit welchem Workflow für welche Aufgabe in ihrem Fach zu nutzen ist.	Delegationsurteil auf Berufsniveau.
Tertiär	Sorgfalt	Simulation regulierter Branche: Studierende müssen Arbeit auf Qualitätsstandard mit vollständiger Offenlegung, Zuschreibung und Überprüfungsspur entsprechend Artikel 4 der EU-KI-Verordnung erbringen.	Brücke zur beruflichen Verantwortlichkeit.

Mehrere Designprinzipien durchziehen die Beispiele in Tabelle 3. Erstens wird die kognitive Routine explizit benannt. Lernenden wird gesagt, was sie tun lernen (delegieren, beschreiben, urteilen, sorgfältig sein), nicht nur, welche Aufgabe sie erledigen. Die metakognitive Rahmung ist wichtig; sie verwandelt eine Werkzeugnutzungsaktivität in eine Routinebildungsaktivität. Zweitens geht die Routine dem Werkzeug voraus. In mehreren der Aktivitäten (insbesondere in der Arbeit zu Beschreibung und Delegation) artikulieren Lernende die kognitive Operation, bevor die KI beteiligt ist. Die KI wird dann zur Testumgebung für eine bereits ausgebildete kognitive Routine, nicht zum Ersatz für ihre Ausbildung. Drittens wird Reibung gezielt bewahrt. Überprüfung, Offenlegung und Iteration werden nicht umgangen, weil das

Werkzeug es erlaubt; sie werden verlangt, weil das Curriculum es verlangt. Die Reibung ist die Lektion.

Viertens ist Bewertung eingebettet. Jede Aktivität produziert Belege (die Prompt-Iterationen, die überprüfte Bibliografie, das Offenlegungsprotokoll, die Entscheidungsmatrix), die hinsichtlich der kognitiven Operationen, auf die die Aktivität zielt, bewertet werden können. Dies adressiert den in Kapitel 2 identifizierten Kollaps der Bewertungsvalidität: der Bewertungsbeleg ist der Routinebildungsbeleg, nicht der KI-substituierbare Output. Fünftens skalieren die Aktivitäten über das Curriculum hinweg. Sekundarschulaktivitäten bereiten auf Tertiäraktivitäten vor. Primarschulvorläufer bereiten auf Sekundarschulaktivitäten vor. Das Rahmenwerk ist zur Umsetzung über das K-16-Spektrum hinweg als entwicklungsbezogen kohärente Sequenz gedacht, nicht als isolierte jahrgangsspezifische Einheiten.

Zwei operative Fallstricke verdienen explizite Beachtung. Der erste ist das, was ich als Token-Offenlegung bezeichne: die formale Hinzufügung einer Zeile "Ich habe KI für X verwendet" am Ende der Schülerarbeit, ohne Auseinandersetzung damit, was delegiert wurde, warum und was überprüft wurde. Token-Offenlegung erfüllt die Oberflächeneinhaltung ohne die zugrunde liegende Sorgfaltsroutine, die die Offenlegung sichtbar machen sollte. Wirksame Sorgfaltspädagogik erfordert strukturierte Offenlegungsformate (was delegiert wurde, warum, was überprüft wurde, was unverifiziert blieb, was der Lernende beim nächsten Mal anders machen würde) und Auseinandersetzung der Lehrkraft mit dem Inhalt der Offenlegung. Der zweite Fallstrick ist das Reibungstheater: die Hinzufügung formaler Reibung (Überprüfungsanforderungen, Versionshistorieanforderungen), die Lernende mechanisch ohne kognitives Engagement ausführen. Reibungstheater produziert den Anschein von Routine ohne deren Substanz. Das Heilmittel ist dasselbe wie bei Token-Offenlegung: Auseinandersetzung der Lehrkraft mit der Substanz des Routinebelegs, nicht nur mit dessen formaler Präsenz.

7. Die internationale politische Landschaft · Konvergierende Rahmenwerke

Das internationale Politikumfeld rund um KI in der Bildung hat sich in den Jahren 2023–2026 ungewöhnlich schnell konvergiert. Bis Mitte 2026 prägen mindestens sieben große Rahmenwerke die institutionelle Entscheidungsfindung, mit erheblicher konzeptueller Überlappung und wachsender operativer Abstimmung. Dieses Kapitel fasst die wichtigsten Rahmenwerke zusammen und ordnet sie dem in diesem Papier vorgeschlagenen 4D-Rahmenwerk zu. Die Konvergenz stellt eine Chance dar: ein global interpretierbares Kompetenzvokabular bildet sich heraus, an dem nationale Systeme bewertet und zwischen denen institutionelle Reformen koordiniert werden können. Sie stellt auch eine Einschränkung dar. Institutionen, die GenKI ohne Auseinandersetzung mit dieser entstehenden politischen Infrastruktur einsetzen, geraten zunehmend in Konflikt mit regulatorischen Erwartungen und professionellen Normen.

7.1 Die EU-KI-Verordnung und die Pflicht zur KI-Kompetenz nach Artikel 4

Die KI-Verordnung der Europäischen Union (Verordnung (EU) 2024/1689) trat am 1. August 2024 in Kraft und wird bis 2027 schrittweise umgesetzt (Europäische Union, 2024). Artikel 4 erlegt Anbietern und Nutzern von KI-Systemen eine verbindliche Verpflichtung zur KI-Kompetenz auf: Sie müssen sicherstellen, dass ihr Personal und alle, die die Systeme in ihrem Namen nutzen, ein angemessenes Maß an KI-Kompetenz haben. Für Bildungsinstitutionen entsteht daraus eine doppelte Verpflichtung. Erstens sind

Institutionen, die KI-Werkzeuge für administrative, lehrende oder bewertende Zwecke einsetzen, selbst Nutzer im Sinne der Verordnung, mit unmittelbarer Compliance-Pflicht zur KI-Kompetenz des Personals. Zweitens bilden Institutionen, deren Absolventen in regulierte Branchen eintreten werden, das Personal aus, auf das die Verpflichtung des Artikels 4 am künftigen Arbeitsplatz Anwendung findet. Die nachgelagerte berufliche Verpflichtung macht KI-Kompetenz nicht zu optionalem Curriculuminhalt, sondern zu beruflich erforderlicher Vorbereitung.

7.2 UNESCO-Rahmen für Schüler und Lehrkräfte

Die UNESCO veröffentlichte auf der Digital Learning Week im September 2024 zwei KI-Kompetenzrahmen, einen für Schüler und einen für Lehrkräfte. Der Rahmen für Lehrkräfte (Miao & Cukurova, 2024) definiert 15 Lehrkräftekompetenzen über fünf Dimensionen: menschenzentriertes Mindset; KI-Ethik; KI-Grundlagen und -Anwendungen; KI-Pädagogik; und KI für berufliche Entwicklung. Der Schülerrahmen definiert parallele Kompetenzen auf Lernendenebene. Die UNESCO-Rahmung betont ethischen und menschenzentrierten Einsatz, insbesondere in Kontexten, in denen Technologie, Infrastruktur oder institutionelle Kapazität zur Gewährleistung verantwortungsvoller Nutzung uneinheitlich sind. Die globale Bereitschaft gegenüber den UNESCO-Rahmen wurde auf unter 4 % geschätzt (Welker, 2025).

7.3 Der EU-OECD-AILit-Rahmen

Die Europäische Kommission und die OECD veröffentlichten im Mai 2025 gemeinsam Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education (den AILit-Rahmen) als Entwurf zur Begutachtung, wobei die Endversion für 2026 vorgesehen ist (Europäische Kommission & OECD, 2025). Der Rahmen organisiert KI-Kompetenz um vier Domänen, mit 22 darüber verteilten Kompetenzen: Sich mit KI einbringen (verstehen, wann und wie KI präsent ist, Outputs kritisch bewerten); Mit KI gestalten (mit KI-Werkzeugen zur Problemlösung und Kreativität zusammenarbeiten, mit Aufmerksamkeit für ethische Implikationen); KI verwalten (verantwortungsvolle Entscheidungen über personenbezogene Daten, Handlungsspielraum und gesellschaftliche Wirkung); und KI entwerfen (grundlegendes Verständnis dafür, wie KI-Systeme entworfen und gebaut werden). Der Rahmen wurde mit Beiträgen internationaler Experten aus mehreren Ländern entwickelt und durch Konsultationen mit fast 1 000 Lehrkräften, Forschern, politischen Entscheidungsträgern und Stakeholdern informiert.

7.4 PISA 2029 MAIL-Erhebung

Die OECD veröffentlichte im Februar 2026 den ersten Entwurf des Erhebungsrahmens für die PISA-2029-Medien- und KI-Kompetenz (MAIL) (OECD, 2026) unter dem Titel Navigating an Evolving Digital World. Die Erhebung wird im Rahmen des PISA-2029-Zyklus durchgeführt, mit Ergebnissen, die für Dezember 2031 erwartet werden. Die MAIL-Erhebung wird die Fähigkeit fünfzehnjähriger Schüler testen, sich kritisch und verantwortungsvoll in Umgebungen einzubringen, die von Medienplattformen und KI geprägt sind. Die institutionelle Implikation ist erheblich. KI-Kompetenz wird bis 2029 eine gemessene, international vergleichbare Kompetenz sein. Nationale Bildungssysteme werden hinsichtlich ihrer Performance verglichen werden. Dies ist das bisher stärkste politische Signal, dass KI-Kompetenz sich von einer optionalen Ergänzung zu einer erforderlichen Curriculum-Infrastruktur entwickelt.

7.5 Nationale und sektorale Rahmenwerke

Über die Arbeit der EU und der OECD hinaus operieren nun mehrere nationale und sektorspezifische Rahmen. Das US-Arbeitsministerium hat einen KI-Kompetenzrahmen mit fünf grundlegenden Inhaltsbereichen und sieben Vermittlungsprinzipien veröffentlicht, der die KI-Kompetenzanstrengungen auf Arbeitskräfteebene über Branchen hinweg leiten soll (AI Literacy Institute, 2026). Digital Promise hat einen KI-Kompetenzrahmen vorgeschlagen, der um drei Domänen organisiert ist, Verstehen, Nutzen und Bewerten, und für die Nutzung in K-12 und durch Lehrkräfte gedacht ist (Digital Promise, 2024). Der Digital Education Council hat einen KI-Kompetenzrahmen veröffentlicht, der menschenzentrierte Kompetenzen (kritisches Denken, Kreativität, emotionale Intelligenz) betont. Das California Department of Education hat KI-Kompetenz-Leitlinien veröffentlicht, die um entwicklungsbezogen angemessene Progressionen über die Jahrgänge hinweg organisiert sind (California Department of Education, 2025). Singapurs Institute of Policy Studies hat ein Arbeitspapier veröffentlicht, das einen einheitlichen nationalen KI-Kompetenzrahmen vorschlägt, der internationale Ansätze mit der singapurspezifischen Umsetzung synthetisiert (Chew, Soon, & Kaur, 2026).

Über diese Rahmen hinweg wiederholen sich mit hoher Konsistenz vier substantielle Domänen: kritische Auseinandersetzung mit KI (entspricht grob dem Urteilsvermögen im 4D-Rahmenwerk), Gestaltung und Zusammenarbeit mit KI (schöpfend aus Delegation und Beschreibung), Verwaltung und ethischer Einsatz von KI (Sorgfalt) und Design- oder Systemebene-Verständnis von KI (eine umfassendere Kompetenz, die das 4D-Rahmenwerk nicht direkt adressiert, sondern stattdessen der disziplinären Integration auf tertiärer Stufe überlässt).

7.6 Das 4D-Rahmenwerk im Verhältnis zu konvergierenden Kompetenzrahmen

Tabelle 2 ordnet die kognitiven Routinen des 4D-Rahmenwerks den vier AILit-Domänen und dem Vierdomänen-Konsens zu, der sich aus den internationalen Rahmen herauskristallisiert hat. Das hier angebotene Rahmenwerk operiert auf einer anderen Ebene pädagogischer Spezifikation. Wo die Kompetenzrahmen die Ergebnisse spezifizieren, die ein Lernender demonstrieren können sollte, spezifiziert das 4D-Rahmenwerk die kognitiven Routinen, die ein Lernender aufbauen muss, um diese Ergebnisse demonstrieren zu können. Beide ergänzen sich, sie stehen nicht in Konkurrenz zueinander. Eine Institution, die den AILit-Rahmen als Spezifikation der Kompetenzergebnisse und das 4D-Rahmenwerk als Spezifikation der kognitiven Routinen übernimmt, verfügt über eine vollständige pädagogische Architektur: was Lernende tun können sollten und welche kognitiven Routinen sie aufbauen müssen, um dazu in der Lage zu sein.

Tabelle 2. Zuordnung des 4D-Rahmenwerks zu den vier AILit-Domänen.

AILit-Domäne	Domänenbeschreibung	Eingesetzte 4D-Routinen	Stufenbetonung
Sich mit KI einbringen	Kritische Bewertung von KI-Outputs und Erkennen der Präsenz von KI in Werkzeugen.	Urteilsvermögen (primär); Delegation (sekundär)	Alle Stufen, Tiefe wächst
Mit KI gestalten	Zusammenarbeit mit KI zur Problemlösung und Kreativität, mit ethischer Aufmerksamkeit.	Beschreibung (primär); Delegation, Sorgfalt	Ab Sekundarstufe

AILit-Domäne	Domänenbeschreibung	Eingesetzte 4D-Routinen	Stufenbetonung
KI verwalten	Verantwortungsvolle Entscheidungen über personenbezogene Daten, Handlungsspielraum, gesellschaftliche Wirkung.	Sorgfalt (primär); Urteilsvermögen	Sekundär und tertiär
KI entwerfen	Grundlegendes Verständnis dafür, wie KI-Systeme entworfen und gebaut werden.	Urteilsvermögen auf Systemebene erweitert	Tertiär, optional in oberer Sekundarstufe

8. Werkzeugauswahl · Warum die Modellarchitektur in der Bildung wichtig ist

Bildungsinstitutionen, die GenKI an Lernende und insbesondere an Minderjährige ausliefern, treffen eine nicht-triviale Wahl, wenn sie ein grenzgängiges Modell auswählen. Die Wahl wird oft als Commodity-Auswahl behandelt (welches Abonnement, welche Funktionen, welche Integration), aber sie ist meines Erachtens eine pädagogische Entscheidung mit Folgen, die der entwicklungsbezogenen Verletzlichkeit der Lernendenpopulation entsprechen. Dieses Kapitel untersucht, wie eine institutionelle Wahl aussieht, wenn sie nach pädagogischen und Sicherheitskriterien und nicht nach Funktions- oder Kostenkriterien getroffen wird.

Die Entwicklung grenzgängiger großer Sprachmodelle hat sich von 2022 bis 2026 zwischen Systemen, deren Alignment-Methodik öffentlich in substantiellem technischen Detail dokumentiert ist, und Systemen, deren Alignment-Methodik nicht oder nur teilweise offengelegt ist, aufgespalten. Dies ist keine Marketingunterscheidung. Es ist eine methodologische Unterscheidung. Anthropic, der Entwickler von Claude, hat detaillierte technische Artikel zu seiner Constitutional-AI-Methodik veröffentlicht (Bai et al., 2022), in denen das Modell gegen einen expliziten, öffentlichen Satz konstitutioneller Prinzipien trainiert wird (abgeleitet aus Quellen einschließlich der Allgemeinen Erklärung der Menschenrechte der UN), und der Prozess des Reinforcement Learning aus KI-Feedback im Nachtraining ist mit hinreichender Detailgenauigkeit für externe Prüfung dokumentiert. Anthropic veröffentlicht auch eine Responsible Scaling Policy (Anthropic, 2026), die die Sicherheitsschwellen und die bei der Skalierung der Modellfähigkeiten angewandten Maßnahmen dokumentiert. OpenAI veröffentlicht Modell- und Systemkarten. Google DeepMind veröffentlicht technische Berichte unterschiedlichen Detailgrads. Das Niveau der methodologischen Transparenz unterscheidet sich jedoch erheblich zwischen den Anbietern.

Für die Bereitstellung an erwachsene Fachnutzer mag dieser Transparenzgradient eine Frage professioneller oder wissenschaftlicher Präferenz sein. Für die Bereitstellung an Minderjährige in Bildungsumgebungen ist es eine treuhänderische Frage. Bildungsinstitutionen stehen in loco parentis für viele der Kinder und Jugendlichen, denen sie diese Systeme bereitstellen. Die Institution trifft im Namen der Familie eine Entscheidung darüber, mit welchem kognitiven Partner das Kind große Mengen kognitiv prägender Zeit verbringen wird. Diese Entscheidung sollte gegenüber der Frage verteidigbar sein: welche Belege hat die Institution dafür, dass das fragliche System darauf ausgerichtet ist, Outputs zu produzieren, die mit den Werten der Institution, den Werten der Familie und der entwicklungsbezogenen Sicherheit des Kindes vereinbar

sind. Die Methodentransparenz macht diese Verteidigbarkeit bewertbar. Die Methodenintransparenz macht sie zu einer reinen Frage des Vertrauens in den Anbieter.

Die Constitutional-AI-Methodik von Anthropic (Bai et al., 2022) ist die am ausführlichsten dokumentierte Alignment-Methodik, die derzeit in produktiven grenzgängigen Systemen verfügbar ist. Die Methodik kurz: anstatt sich ausschließlich auf menschliches Feedback zur Formung des Modellverhaltens zu verlassen, wird das Modell gegen einen expliziten Satz konstitutioneller Prinzipien trainiert, und ein zweites Modell (ebenfalls von Anthropic trainiert) liefert Feedback zur Einhaltung dieser Prinzipien, wobei die Verfassung und der Trainingsprozess öffentlich dokumentiert sind. Dies bringt ein System hervor, dessen Ausrichtung auf eine bestimmte Wertstruktur prüfbar ist und dessen Wertstruktur selbst angreifbar ist. Die Responsible Scaling Policy (Anthropic, 2026) erweitert diese methodologische Transparenz auf den Entscheidungsprozess der Skalierung. Der RSP-Rahmen war nicht ohne Kontroverse. TIME berichtete im Februar 2026, dass Anthropic eine der ursprünglichen zentralen Pre-Training-Verpflichtungen der RSP gelockert hatte, was substantielle Kritik auslöste (Perrigo, 2026). Das breitere Aktiv bleibt jedoch die öffentliche Dokumentation der Methodik und ihrer Revisionen.

Ich behaupte nicht, dass Claude notwendigerweise das beste grenzgängige Modell für eine gegebene Bildungsaufgabe in Bezug auf Fähigkeiten oder Kosten ist. Fähigkeitsvergleiche sind domänenspezifisch und ändern sich rasch. Ich behaupte, dass Methodentransparenz ein vertretbares Kriterium für die institutionelle Bildungsauswahl ist, dass sie durch öffentlich dokumentierte Alignment-Forschung und Verpflichtungen zur verantwortungsvollen Skalierung in angemessener Weise operationalisierbar ist und dass die institutionelle Voreinstellung sie bei Bereitstellungsentscheidungen, die Minderjährige betreffen, stark gewichten sollte. Ein fähigeres Modell mit intransparenter Methodik ist nicht kategorisch einem etwas weniger fähigen Modell mit transparenter Methodik vorzuziehen, was die Bildungsauslieferung an Kinder angeht.

Daraus ergeben sich drei operative Kriterien. (i) Methodentransparenz. Der Anbieter veröffentlicht technische Details der Alignment-Methodik, die für externe Bewertung hinreichend sind. (ii) Regulatorische Kooperation. Der Anbieter kooperiert mit einschlägigen Regulierungsrahmen (Einhaltung der EU-KI-Verordnung, nationaler Kinderschutz, sektorspezifische Rahmen, soweit anwendbar). (iii) Investition in Sicherheitsforschung. Der Anbieter zeigt eine wesentliche Investition in Sicherheitsforschung, mit Veröffentlichungen und operativen Praktiken, die nahelegen, dass der Sicherheitsrahmen nicht rein marketinggetrieben ist. Diese drei Kriterien sind nicht erschöpfend und werden sich mit der Reifung des Feldes verfeinern, aber sie bieten Institutionen eine vertretbare Ausgangsposition für das Gespräch über die Modellauswahl.

9. Sicherheit, Ethik und das Prinzip Schutz durch Voreinstellung

Über die Modellauswahl hinaus wirft die institutionelle Bereitstellung von GenKI eine Reihe von Sicherheits- und ethischen Überlegungen auf, die die gegenwärtige Bildungspolitik uneinheitlich adressiert. Ich gruppiere sie unter fünf Überschriften, mit Verweisen auf relevante regulatorische und Forschungskontexte.

(i) Datenschutz und Einwilligung Minderjähriger. Die Bereitstellung grenzgängiger KI an Lernende unter 18 Jahren berührt die Datenschutz-Grundverordnung der EU, nationale Kinderschutzrahmen und entstehende Bestimmungen der EU-KI-Verordnung. Viele aktuelle institutionelle Bereitstellungen erfüllen diese Verpflichtungen nicht sauber.

Institutionen sollten die Rechtsgrundlage der Verarbeitung, das Aufbewahrungsregime und die Einwilligungsarchitektur für jede KI-Bereitstellung an Minderjährige dokumentieren. Die Voreinstellung sollte Datenminimierung sein: in Bildungsumgebungen eingesetzte KI-Werkzeuge sollten so konfiguriert sein, dass sie das Minimum an personenbezogenen Daten aufbewahren, das für den Bildungszweck erforderlich ist, mit Aufbewahrungsfristen, die auf Bildungs- und nicht auf kommerzielle Zeithorizonte abgestimmt sind. Wo Anbieterarchitekturen standardmäßig auf breitere Aufbewahrung ausgelegt sind, trägt die Institution die Verpflichtung zu verhandeln oder eine Alternative auszuwählen.

(ii) Inhaltsexposition und altersgerechte Schutzleitplanken. Grenzgängige Modelle können grundsätzlich Inhalte erzeugen oder diskutieren, die für jüngere Lernende ungeeignet sind. Institutionelle Bereitstellungen sollten altersgestufte Schutzleitplanken, auf die Entwicklungsstufe konfigurierte Inhaltsfilter und klare Eskalationswege für ungeeignete Exposition umfassen. Lehrkräfte sollten geschult sein, Anzeichen dafür zu erkennen, dass ein Schüler möglicherweise über KI-Werkzeuge auf ungeeignete Inhalte gestoßen ist, und die Schule sollte ein dokumentiertes Reaktionsprotokoll haben. Kritisch ist, dass die Institution die Verpflichtung trägt, dies proaktiv zu tun, nicht reaktiv, denn das erste Auftreten ungeeigneter Inhalte beim Schüler ist bereits zu spät.

(iii) Kognitive Überabhängigkeit und Abhängigkeitsmuster. Neuere Arbeiten haben das identifiziert, was einige Autoren als Generatives-KI-Abhängigkeitssyndrom beschreiben (Kooli et al., 2025; Gerlich, 2025), gekennzeichnet durch kompulsive Nutzungsmuster, verminderte Toleranz gegenüber kognitiver Anstrengung und entzugsähnliches Unbehagen, wenn das Werkzeug nicht verfügbar ist. Die institutionelle Politik sollte Nutzungshäufigkeitsgrenzen als schützend, nicht als puritanisch betrachten. Dies ist strukturell ähnlich der institutionellen Reaktion auf problematische Spielenutzung oder Social-Media-Nutzung, die die Literatur ausführlich dokumentiert (Reggente et al., 2026): die Schule kann die Nutzung nicht völlig verhindern, aber sie kann die Schulumgebung so strukturieren, dass KI-Nutzung nicht die Voreinstellung für kognitive Aufgaben ist, und sie kann Schülern beibringen, problematische Nutzungsmuster zu erkennen und selbst zu regulieren.

(iv) Verzerrung, Fairness und die Repräsentationspolitik der Trainingsdaten. Grenzgängige Modelle integrieren die demografischen, sprachlichen und epistemischen Verzerrungen ihrer Trainingskorpora. Die Bildungsbereitstellung sollte explizite pädagogische Aufmerksamkeit für diese Tatsache umfassen. Schüler sollten lernen, KI-Outputs auf Verzerrung zu hinterfragen, nicht passiv Outputs ausgesetzt zu werden, die Verzerrung unter Flüssigkeit einschmuggeln. Diese pädagogische Verpflichtung ist besonders dringlich für Lernende aus Gruppen, die in Trainingsdaten unterrepräsentiert sind, deren eigene Sprachen, Kulturen und Wissensweisen in Modelloutputs statistisch untergewichtet sind. Ein Lernender, der Modellflüssigkeit als autoritativ verinnerlicht, verinnerlicht die Repräsentationspolitik des Trainingskorpus. Kritische KI-Kompetenz (Roe, Furze, & Perkins, 2026; Illingworth, 2026) bietet einen gut entwickelten Ansatz zur Vermittlung dieser kritischen Haltung.

(v) Umweltkosten. Inferenz und Training grenzgängiger Modelle tragen nicht triviale Energie- und Wasserfußabdrücke. Roe, Furze und Perkins (2025) ordnen in ihrem Rahmen für kritische KI-Kompetenz die Umweltkosten unter die vier Kerndomänen der KI-Kompetenz-Befragung ein. Die institutionelle Bereitstellung sollte sie transparent berücksichtigen. Wo Institutionen sich für Politiken des unbegrenzten KI-Zugangs entscheiden, sollten die Umweltimplikationen offengelegt und als Teil des Curriculums

reflektiert werden, nicht in der Lizenzvereinbarung versteckt.

Zusammengenommen sprechen diese Überlegungen für das, was ich als Schutz durch Voreinstellung bezeichne: KI-Bereitstellung, deren Standardkonfiguration schützend ist und positive Begründung erfordert, um Schutz zu lockern, statt positiver Anstrengung, um ihn hinzuzufügen. Schutz durch Voreinstellung ist die Bildungsanwendung des Prinzips Datenschutz durch Voreinstellung, das nun in der DSGVO verankert ist und in der KI-Regulierung entsteht. Es legt die Beweislast auf jene, die gelockerten Schutz befürworten, nicht auf jene, die Schutz befürworten. Für Institutionen, die GenKI an Minderjährige ausliefern, ist dies meines Erachtens die einzig vertretbare Standardarchitektur.

10. Anpassung der Lehrkraft · Vom Inhaltsvermittler zum kognitiven Coach

Die Rolle der Lehrkraft verlagert sich. In einer Welt flüssiger On-Demand-Informationen ist der Wert, den eine Lehrkraft hinzufügt, nicht mehr primär die Vermittlung von Inhalten. Er liegt im Modellieren, Stützen und Bewerten kognitiver Praktiken, die Lernende nicht allein aus dem Modell erwerben können: Urteilsvermögen, produktiver Widerspruch, anhaltendes Denken unter Unsicherheit, ethisches Urteil, intellektueller Mut. Der Übergang führt vom Inhaltsvermittler zum kognitiven Coach. Manche Lehrkräfte heißen ihn willkommen. Viele nicht. Die institutionellen und persönlichen Einsätze des Übergangs sind hoch genug, dass das Rahmenwerk ihn nicht ignorieren kann.

Diese Verlagerung ist nach meiner Erfahrung im Unterrichten angewandter KI gegenüber amerikanischen und europäischen Studierenden auf tertiärem Niveau zugleich befreiend und bedrohlich. Sie befreit Lehrkräfte von jenen Teilen des Unterrichtens, die KI nun einigermaßen gut leistet (grundlegende Erklärung, Differenzierung der Erklärung nach Niveau, Entwurf routinemäßiger Materialien). Sie bedroht jene, deren berufliche Identität auf diesen Teilen aufbaut. Das hier vorgeschlagene Rahmenwerk ist unter anderem eine Wegkarte für die Lehrkräftevorbereitung, die die Bedrohung ernst nimmt und zugleich klärt, was die Lehrkraft jetzt hinzufügt, was die KI nicht kann.

Wirksame Lehrkräftevorbereitung in diesem Übergang umfasst meines Erachtens vier Elemente. Erstens, Bedienerflüssigkeit. Lehrkräfte müssen selbst kompetente Bediener der Systeme sein, deren Nutzung sie den Lernenden beibringen. Dies ist eine nicht verhandelbare Voraussetzung. Eine Lehrkraft, die nie eine anhaltende Arbeit mit einem KI-Assistenten verfasst hat, kann weder Urteilsvermögen noch Sorgfalt wirksam vermitteln. Bedienerflüssigkeit wird durch Praxis erworben, gestützt durch strukturierte Weiterbildung. Anthropic's Kurs AI Fluency for Educators (Anthropic, 2025) bietet einen strukturierten Weg. Der UNESCO-Lehrkräftekompetenzrahmen (Miao & Cukurova, 2024) bietet eine umfassendere Spezifikation.

Zweitens, pädagogische Neurahmung. Explizite Weiterbildung darüber, was es bedeutet, zu unterrichten, wenn Lernende ständigen Zugang zu einer flüssigen Antwortmaschine haben. Dies ist keine technische Schulung zu KI-Werkzeugen. Es ist eine pädagogische Umschulung darüber, wie Aufgaben gestaltet, Unterricht sequenziert und Bewertung strukturiert werden, sodass die kognitiven Routinen, deren Entwicklung das Curriculum dienen soll, tatsächlich entwickelt werden. Die Verlagerung von der Inhaltsvermittlung zum kognitiven Coaching ist real, aber sie wird auch gelehrt, nicht vorausgesetzt.

Drittens, Bewertungsneudesign. Praktische Kapazität, Bewertungen zu gestalten, die unter KI-Zugang valide bleiben. Traditionelle häusliche Bewertung erfüllt in den meisten

Disziplinen ihren Zweck nicht mehr. Ersatzdesigns umfassen unterrichtsinterne beaufsichtigte Demonstration, mündliche Verteidigung eingereichter Arbeiten, Portfolio-Zusammenstellung mit Versionshistorie, Verfolgung prozessbezogener Belege und projektbasierte Bewertung mit öffentlichen Artefakten. Jedes dieser Designs hat eigene Designeinschränkungen und Kompromisse. Lehrkräfte benötigen explizite Schulung in ihrer Auswahl und Anwendung.

Viertens, ethische und regulatorische Kompetenz. Hinreichende Grundlagen der EU-KI-Verordnung, nationaler Kinderschutzrahmen und disziplinärer Berufsethik, um darin kompetent zu operieren. Lehrkräfte sind nun, ob sie es wissen oder nicht, Nutzer von KI-Systemen im Sinne der EU-KI-Verordnung. Die Verpflichtung zur KI-Kompetenz nach Artikel 4 fällt auf sie. Institutionen, die diese Grundlage nicht bereitstellen, setzen ihre Lehrkräfte Compliance-Versäumnissen aus, die letztlich Versäumnisse der Institution selbst sind.

Über alle vier Elemente hinweg gilt das gleiche Prinzip: die Autorität der Lehrkraft im KI-Zeitalter leitet sich aus kognitivem Coaching, ethischem Modellieren und geschützter Aufmerksamkeit für die Entwicklungstrajektorie des Lernenden ab, nicht aus dem Sein als Inhaltsquelle. Dies ist in vielerlei Hinsicht eine Rückkehr zu älteren Modellen der Pädagogik (der Tutor, der sokratische Begleiter, der disziplinäre Mentor), die das Modell der industriellen Inhaltsvermittlung vorübergehend verdrängt hat. Die Verlagerung ist, richtig gerahmt, keine Bedrohung für den Beruf, sondern eine Klärung dessen, was schon immer sein Kern war.

10.1 Ein dreistufiger Weg der beruflichen Weiterbildung

Die institutionelle Umsetzung der oben genannten Anforderungen an die Lehrkräftevorbereitung profitiert von einem gestuften Weg. Lehrkräfte bewegen sich nicht in einem einzigen Workshop vom KI-Neuling zum kognitiven Coach. Der Übergang wird ehrlicher als mehrjähriger Bogen der beruflichen Weiterbildung verstanden, vergleichbar im Umfang mit der Integration jeder größeren curriculären Verlagerung (der Einführung des computationalen Denkens, der Integration der Medienkompetenz, der Verankerung interkultureller Kompetenz). Tabelle 4 skizziert einen dreistufigen Weg mit stufengerechten Komponenten der beruflichen Weiterbildung, Bewertungsindikatoren und empfohlenen Zeitrahmen. Der Weg ist nicht das einzig mögliche Design, bietet aber einen institutionell handhabbaren Ausgangspunkt.

Tabelle 4. Dreistufiger Weg der beruflichen Weiterbildung für Lehrkräfte.

Stufe	Komponenten	Bewertungsindikatoren	Zeitrahmen
Stufe 1 · Grundlegende Flüssigkeit	KI-Nutzung auf Bedienerenebene; grundlegendes Prompt Engineering; Werkzeugauswahlkompetenz; Einführung in die AILit- und UNESCO-Rahmen; Orientierung zu Artikel 4 der EU-KI-Verordnung.	Lehrkraft kann nützlichen Output über mindestens drei disziplinrelevante Aufgabentypen mit einem grenzgängigen Modell produzieren. Lehrkraft kann die institutionelle Begründung der Werkzeugauswahl artikulieren.	3-6 Monate

Stufe	Komponenten	Bewertungsindikatoren	Zeitraum
Stufe 2 · Pädagogische Neurahmung	Workshops zur Aufgaben- und Bewertungsneudesign; Integration des 4D-Rahmenwerks; Mikrounterricht zum kognitiven Coaching; fachspezifische Umsetzungsplanung.	Lehrkraft hat mindestens eine Bewertung neu gestaltet, die unter KI-Zugang valide bleibt. Lehrkraft kann die 4D-Routinen in einer Unterrichtsbeobachtung modellieren.	6–12 Monate
Stufe 3 · Institutionelle Führung	Beitrag zur Curriculumsgestaltung; Peer-Coaching; Beteiligung an der institutionellen KI-Politik; Auseinandersetzung mit externen Rahmen (PISA 2029 MAIL, nationale Stellen).	Lehrkraft leitet einen Curriculu msneudesign-Workstream oder eine Peer-Coaching-Kohorte. Lehrkraft trägt zu institutionellen Politikdokumenten zur KI-Integration bei.	12–24 Monate

Über die drei Stufen hinweg gelten mehrere Umsetzungsprinzipien. Erstens ist der Weg sequenziell, aber nicht streng linear; einige Lehrkräfte werden Stufe 1 schnell durchlaufen und vom beschleunigten Eintritt in Stufe 2 profitieren, insbesondere jene mit bestehendem Technologie- oder Informatikhintergrund. Andere werden in Stufe 1 wesentlich länger benötigen, insbesondere in Disziplinen, in denen die domänenspezifische Nützlichkeit des Werkzeugs weniger offensichtlich ist oder in denen die bisherige Unterrichtspraxis sich nicht mit Technologievermittlung auseinandergesetzt hat. Differenziertes Tempo ist entscheidend. Zweitens sollte der Weg institutionell gestützt sein, nicht individuell getragen. Zeit, Peer-Coaching-Infrastruktur und institutionelle Anerkennung der Progression sind ebenso wichtig wie der formale Inhalt. Lehrkräfte, die den Weg durchlaufen, erbringen einen institutionellen Wert vergleichbar mit der Berufszertifizierung in anderen Bereichen; die institutionelle Reaktion sollte dem entsprechen. Drittens sollte der Weg für externe Stakeholder sichtbar sein, Eltern, Schüler, Regulierungsbehörden, Akkreditierungsstellen. Der Stand der Lehrkräftebereitschaft der Institution ist zunehmend eine Frage, die externe Parteien stellen werden, und eine Institution, die mit einem dokumentierten Weg antworten kann, ist in einer wesentlich besseren Position als eine, die das nicht kann.

Eine weitere Überlegung ist das Verhältnis zwischen institutionellem und individuellem Lehrkräfteeinsatz. Einige Institutionen sind gut positioniert, den Weg intern bereitzustellen, mit eigenen Spezialisten für Bildungstechnologie, Lernwissenschaften oder Pädagogik, die die Arbeit leiten können. Andere werden externe Weiterbildung beschaffen, mit Universitäten partnerschaftlich zusammenarbeiten oder Konsortien beitreten müssen. Anthropic's Kurs AI Fluency for Educators (Anthropic, 2025), die Schulungsressourcen zum UNESCO-Kompetenzrahmen (Miao & Cukurova, 2024) und die Klassenraumszenarien des AILit-Rahmens (Europäische Kommission & OECD, 2025) bieten eine Reihe strukturierter externer Ressourcen. Nationale KI-Kompetenzinitiativen in den USA, Singapur und anderswo bieten ergänzende Infrastruktur. Das Rahmenwerk ist agnostisch in Bezug darauf, welche spezifischen externen Partnerschaften eine Institution nutzen sollte; es ist nicht agnostisch in Bezug darauf, ob eine solche Unterstützung erforderlich ist, was für die meisten Institutionen wesentlich der Fall ist.

11. Ausblick · 2027-2030

Die Vorhersage der KI-Fähigkeitstrajektorien über einen Drei- bis Fünfjahreshorizont ist unzuverlässig. Die Vorhersage der institutionellen Reaktionen darauf ist noch unzuverlässiger. Ich biete hier keine Vorhersagen, sondern Planungsannahmen, die ich für am vertretbarsten halte und auf denen institutionelle Reformentscheidungen im Horizont 2026–2030 vernünftigerweise aufgebaut werden können. Diese Annahmen sind vorsichtig kalibriert; Institutionen, die auf ihnen aufbauen, sollten mit jährlicher Überprüfung rechnen, während sich die Evidenz häuft.

Annahme 1: Fortgesetztes Fähigkeitswachstum. Die Fähigkeiten grenzgängiger Modelle werden mit Raten voranschreiten, die die Curriculumsrevisionszyklen der Institutionen überschreiten. Spezifische Fähigkeitsgrenzen, mehrstufiges Reasoning, anhaltende agentische Aufgabenausführung, Kohärenz langer Kontexte, multimodale Integration, werden sich weiter ausdehnen. Die Inferenzkosten werden weiter sinken und grenzgängige Fähigkeiten jüngeren und wirtschaftlich benachteiligten Lernenden mit Raten zugänglich machen, die die Fähigkeit der Institutionen zur Skalierung von KI-Kompetenzcurricula übersteigen. Die institutionelle Implikation lautet, dass jede spezifische Werkzeugschulung in etwa zwölf Monaten veraltet. Die Schulung kognitiver Routinen (das 4D-Rahmenwerk) sollte beständiger sein, da die Routinen, die zur guten Nutzung von KI erforderlich sind, bewusste Delegation, präzise Beschreibung, kritische Bewertung, sorgfältige Nacharbeit, über Werkzeuggenerationen hinweg übertragbar sind.

Annahme 2: Agentische Systeme werden bildungsrelevant. Der Übergang von Einzelschritt-Chatbot-Interaktionen zu mehrstufigem agentischen Betrieb (Werkzeugkomposition, Aufgabenausführung, strukturierte Outputrückgaben) ist nun in der kommerziellen Bereitstellung fest etabliert und wird in den Jahren 2027–2028 in der Bildungsbereitstellung ankommen. Die Bildungsintegration, die für das Chatbot-Paradigma entworfen wurde, wird nicht sauber übertragen. Agentische Systeme werfen distinkte pädagogische Fragen auf: wie überwacht ein Lernender einen mehrstufigen Prozess, den er nicht selbst ausgeführt hat? Wie bewertet eine Lehrkraft die vom Lernenden ausgeführten kognitiven Operationen gegenüber den an einen Agenten delegierten Operationen? Das 4D-Rahmenwerk ist auf agentische Kontexte erweiterbar, aber die operative Spezifikation jeder Stufe wird Anpassung erfordern, wenn sich die agentischen Fähigkeiten ausweiten.

Annahme 3: Die Qualifikationserosion beschleunigt sich. Die Kluft zwischen dem, was Absolventen unter KI-Zugang glaubhaft behaupten können, und dem, was sie ohne Unterstützung demonstrieren können, wird sich erweitern. Der Wert von Qualifikationen, insbesondere für Qualifikationen, die primär durch unbeaufsichtigte schriftliche Arbeit bewertet werden, wird entsprechend abnehmen. Arbeitgeber und nachgelagerte Institutionen werden reagieren, sie reagieren bereits, mit ergänzenden Verifikationsmechanismen (kompetenzbasierte Interviews, Projektportfolio-Reviews, Berufszertifizierung, Probeaufgaben). Das in Kapitel 2 identifizierte Qualifikationsvertrauensproblem trifft am stärksten jene Institutionen, deren Qualifikationen am meisten auf älteren Bewertungsformen beruhen. Bildungsinstitutionen, die dies nicht direkt adressieren (durch Bewertungsneudesign, transparente Offenlegungsnormen und kompetenzbasierte Ergänzungen zu traditionellen Qualifikationen), werden den Marktwert ihrer Qualifikationen mit Raten sinken sehen, die ihr Wirtschaftsmodell bedrohen.

Annahme 4: Politische und regulatorische Konsolidierung. Die in Kapitel 7 dokumentierte internationale politische Konvergenz wird sich fortsetzen. Bis 2029 werden die

Ergebnisse der PISA-2029-MAIL-Erhebung die ersten international vergleichbaren Messungen der KI-Kompetenzergebnisse liefern (OECD, 2026). Mitgliedstaaten, deren Systeme schlecht abschneiden, werden politischem Druck zur Reform ausgesetzt sein. Die Compliance-Verpflichtungen der EU-KI-Verordnung werden bis 2027 vollständig in Kraft gesetzt sein (Europäische Union, 2024), wobei die KI-Kompetenzanforderungen nach Artikel 4 sowohl bei Anbietern als auch bei Nutzern operativ wirken. Bildungsinstitutionen in EU-Rechtsräumen werden direkten Compliance-Verpflichtungen zur KI-Kompetenz des Personals ausgesetzt sein. Nicht-EU-Systeme werden zur Aufrechterhaltung der Vergleichbarkeit für Studierendenmobilität und Qualifikationsanerkennung motiviert sein. Die institutionelle Implikation lautet, dass KI-Kompetenzcurricula im Fenster 2027–2030 keine freiwilligen Erweiterungen mehr sind; sie sind zunehmend erforderliche Infrastruktur.

Annahme 5: Differenzierte institutionelle Reaktion produziert divergente Ergebnisse. Institutionen, die jetzt mit operativer Reform beginnen, mit gezieltem Curriculumsneudesign, Lehrkräftevorbereitung, Bewertungsneudesign und Schutz-durch-Voreinstellung-Infrastruktur, werden am PISA-2029-Messpunkt in einer wesentlich besseren Position ankommen als Institutionen, die Reformen aufschieben. Die Divergenz wird sich über das Jahrzehnt hinweg kumulativ erweitern. Das Fenster für kostengünstige adaptive Reform schließt sich. Bis 2028 werden Institutionen, die nicht begonnen haben, mit zusammengesetzten Aufholkosten konfrontiert sein (Lehrkräfte-Umschulung, Curriculumsneudesign, Bewertungsneudesign, Wiederherstellung des Vertrauens von Eltern und Studierenden) unter wachsendem externen Druck.

Ich biete vier konkrete operative Empfehlungen für Institutionen, die im Fenster 2026–2028 operieren. Erstens, führen Sie ein KI-Kompetenz-Audit gegenüber dem AILit-Rahmen (Europäische Kommission & OECD, 2025) und den UNESCO-Kompetenzrahmen (UNESCO, 2024; Miao & Cukurova, 2024) durch. Identifizieren Sie die institutionelle Lücke explizit. Zweitens, priorisieren Sie die Lehrkräftevorbereitung. Bedienerflüssigkeit für alle Lehrkräfte, pädagogische Neurahmung für jene, die kognitiv anspruchsvolles Material unterrichten, Bewertungsneudesign-Schulung für Bewertungsverantwortliche. Drittens, gestalten Sie die Bewertung über das Curriculum hinweg neu, beginnend mit Disziplinen, die am stärksten dem KI-Substitutionsrisiko ausgesetzt sind. Das Prinzip lautet, dass Bewertung die kognitiven Routinen messen sollte, deren Entwicklung das Curriculum dienen soll, nicht die Outputs, deren Produktion das Curriculum dienen soll. Viertens, setzen Sie KI-Werkzeuge für Lernende mit einer Schutz-durch-Voreinstellung-Standard-einstellung ein, wobei Methodentransparenz, regulatorische Kooperation und Investition in Sicherheitsforschung die primären Auswahlkriterien für angenommene Werkzeuge sind (Kapitel 8).

11.1 Drei Szenarien für 2030

Szenarioplanung ist keine Vorhersage; sie ist eine strukturierte Methode, institutionelle Entscheidungen gegenüber der Bandbreite plausibler Zukünfte robust zu machen. Ich biete drei Szenarien für den Stand der KI-Integration in der Bildung im Jahr 2030, kalibriert anhand der oben genannten Annahmen. Jedes ist intern kohärent, keines ist exklusiv wahrscheinlich, und die institutionellen Reformentscheidungen im Fenster 2026–2028 werden wesentlich beeinflussen, welches der drei das System tatsächlich bewohnt.

Szenario A · Konvergente Reform. Die wichtigsten nationalen Systeme übernehmen operationalisierte Versionen der AILit-, UNESCO- und entstehenden nationalen Rahmen, wobei die PISA-2029-MAIL-Daten die ersten international vergleichbaren Messungen des Fortschritts liefern. Die Compliance der EU-KI-Verordnung ist bis 2027 vollständig in Kraft und bis 2029 in Nutzerinstitutionen operativ wirksam. Wege der Lehrkräftevorbereitung sind über die OECD und Partnerländer hinweg institutionell normalisiert. Der Schattenlehrplan tritt zurück, da das institutionelle Curriculum die Lücke zur Schülerpraxis schließt. Das Qualifikationsvertrauensproblem stabilisiert sich, während das Bewertungsneudesign reift. Dieses Szenario erfordert anhaltendes politisches Engagement, substantielle Investition in die Lehrkräftevorbereitung und fortgesetzte internationale politische Zusammenarbeit. Es ist erreichbar, aber nicht unvermeidlich.

Szenario B · Zweispurige Divergenz. Einige nationale und institutionelle Systeme reformieren aggressiv; andere schieben Reformen auf. Die PISA-2029-MAIL-Daten offenbaren eine substantielle internationale Divergenz in den KI-Kompetenzergebnissen. Ein zweispuriges System entsteht, in dem Absolventen reformierter Institutionen mit wesentlich höherer operativer KI-Flüssigkeit auf den Arbeitsmarkt treten als Absolventen von Institutionen mit aufgeschobener Reform. Der Qualifikationswert divergiert entsprechend. Die internationale Studierendenmobilität verlagert sich in Richtung reformierter Systeme. Das Bild von 2030 ist eines sich erweiternder Ungleichheit sowohl innerhalb als auch zwischen nationalen Systemen, wobei Institutionen, die in den Jahren 2026–2028 mit Reformen begannen, ihren Vorteil kumulativ verstärken. Dies ist nach meiner Einschätzung das wahrscheinlichste Szenario bei den gegenwärtigen politischen Trajektorien.

Szenario C · Krisenausgelöste Reform. Die institutionelle Reform wird zu lange aufgeschoben, und das Qualifikationsvertrauensproblem löst eine diskrete institutionelle Krise aus: eine bedeutende Arbeitgeberzurückweisung von Qualifikationen, ein hochkarätiger Bewertungsbetrug, eine umfassende regulatorische Intervention oder ein Vertrauenskollaps der Öffentlichkeit in signifikantem Umfang. Reform folgt unter Krisenbedingungen, mit höheren Kosten, geringerer Kohärenz und schlechteren Ergebnissen, als sie eine geordnete Reform im Fenster 2026–2028 produziert hätte. Einige Institutionen werden den Übergang nicht überleben. Dieses Szenario ist nicht vorhergesagt, aber auf Subsystemebene nicht unwahrscheinlich; spezifische Disziplinen, Institutionen oder nationale Systeme könnten es bewohnen, selbst dort, wo andere die Szenarien A oder B bewohnen.

Die strategische Implikation für die institutionelle Führung ist direkt. Eine im Fenster 2026–2028 begonnene Reform ist über alle drei Szenarien hinweg robust. Aufgeschobene Reform ist nur gegenüber Szenario A robust. Entscheidungen unter echter Unsicherheit darüber, welches Szenario eine Institution erleben wird, sollten die Robustheit gewichten. Das in diesem Papier angebotene Rahmenwerk ist so konzipiert, dass es über alle drei Szenarien hinweg umsetzbar ist, wobei der institutionelle Weg in den Kapiteln 6, 7, 8, 9 und 10 eine kohärente Reformagenda bildet, die ihren Wert unabhängig davon behält, welches Szenario sich materialisiert.

12. Einschränkungen und künftige Forschungsrichtungen

Dieses Papier hat Einschränkungen, die den Lesern transparent gemacht werden sollten und die die Agenda für die empirische Arbeit bilden, die das Rahmenwerk letztlich erfordert. Ich behandle sie in fünf Clustern.

(i) Das 4D-Rahmenwerk ist theoretisch motiviert, aber empirisch ungeprüft. Das Rahmenwerk baut auf einer Synthese konvergierender Literaturen (Forschung zum kognitiven Auslagern, Entwicklungsneurowissenschaft, KI-Kompetenzrahmen) in Kombination mit operativer Erfahrung in der tertiären Lehre auf. Es wurde keiner kontrollierten empirischen Bewertung unterzogen, und seine spezifischen Behauptungen zur entwicklungsbezogenen Sequenzierung (asymmetrische Exposition über Stufen hinweg; Routine vor Werkzeug) werden nicht direkt getestet. Die testbaren Vorhersagen des Rahmenwerks, etwa dass Lernende, die explizit in 4D-Routinen geschult sind, bessere Ergebnisse in Urteilsvermögen und Sorgfalt zeigen als Lernende mit vergleichbarem KI-Zugang ohne Routinetraining, warten auf empirische Untersuchung.

(ii) Die neurodevelopmentale Evidenzbasis betrifft digitale Exposition allgemein, nicht spezifisch GenKI. Kapitel 5 baut das Argument für asymmetrische Exposition auf Stellvertreterevidenz aus der Forschung zur Bildschirm-, Smartphone- und intensiven digitalen Exposition. Der Stellvertreter ist vertretbar (GenKI teilt relevante Eigenschaften kognitiver Belastung und Belohnungsstruktur mit den Stellvertreterreizen), aber er ist ein Stellvertreter. Direkte Längsschnittforschung zu den Auswirkungen der GenKI-Exposition auf das sich entwickelnde Gehirn existiert bisher nicht in robuster Form. Es ist eine kritische Priorität für das nächste Jahrzehnt der bildungsbezogenen neurowissenschaftlichen Forschung.

(iii) Kulturelle und sprachliche Spezifität. Das Rahmenwerk wurde primär in west- und nordamerikanischen tertiären Bildungskontexten entwickelt, mit sekundärer Aufmerksamkeit für mittel- und osteuropäische K-12-Systeme. Seine Übertragbarkeit auf nicht-westliche Bildungskontexte, distinkte Curriculumstraditionen und Lernende, deren Erstsprachen systematisch in den Trainingsdaten grenzgängiger Modelle unterrepräsentiert sind, wird behauptet, nicht demonstriert. Validierungsarbeit über vielfältige Bildungskontexte hinweg ist erforderlich.

(iv) Die Werkzeugauswahlkriterien sind aspirativ, nicht operativ. Kapitel 8 schlägt Methodentransparenz, regulatorische Kooperation und Investition in Sicherheitsforschung als institutionelle Auswahlkriterien vor. Jedes ist bis zu einem gewissen Grad operationalisierbar, aber keines wird derzeit durch eine robuste, anbietervergleichbare Bewertungsinfrastruktur gestützt. Die Arbeit, diese Kriterien in einen nutzbaren institutionellen Bewertungsrahmen zu überführen, ist selbst eine Forschungs- und Politikagenda. Die Empfehlungen des 4D-Rahmenwerks zur Werkzeugauswahl sollten als Ausgangsposition gelesen werden, nicht als fertige Bewertungsmethodik.

(v) Geschwindigkeit des Wandels. Die empirische Evidenzbasis zu GenKI in der Bildung bewegt sich schneller, als ein einzelnes Papier sie synthetisieren kann. Die hier angebotenen Behauptungen und Empfehlungen spiegeln den Evidenzstand bis Mitte 2026 wider und sollten überarbeitet werden, sobald weitere empirische Arbeit verfügbar wird, insbesondere die PISA-2029-MAIL-Daten (OECD, 2026), der längere Zyklus empirischer Arbeit, die aus dem Digital Education Action Plan der EU hervorgeht (Europäische Kommission & OECD, 2025), und die nächste Generation von Studien zum kognitiven Auslagern.

Vier spezifische empirische Prioritäten folgen. Erstens, Längsschnittkohortenstudien zu den entwicklungsbezogenen kognitiven Folgen der GenKI-Exposition über das K-12-Fenster hinweg, mit angemessenen Kontrollen für sozioökonomischen Status, bestehende kognitive Variabilität und Expositionsintensität. Zweitens, kontrollierte Bewertung der spezifischen Behauptungen des 4D-Rahmenwerks zur

entwicklungsbezogenen Sequenzierung, mit Tests, ob in 4D geschulte Kohorten bessere Ergebnisse in Urteilsvermögen, Sorgfalt und operativer Flüssigkeit zeigen. Drittens, interkulturelle Validierung der Anwendbarkeit des Rahmenwerks über vielfältige Bildungskontexte und unterrepräsentierte sprachliche Gemeinschaften hinweg. Viertens, Entwicklung institutioneller Anbieterbewertungsrahmen, die Methodentransparenz, regulatorische Kooperation und Investition in Sicherheitsforschung als vergleichbare, prüfbare institutionelle Entscheidungsinputs operationalisieren.

Drei weitere methodologische Prioritäten verdienen Hervorhebung. Erstens stützt sich die in Kapitel 3 besprochene empirische Literatur stark auf kurzfristige experimentelle Designs, oft Essay-Aufgaben in einer Einzelsitzung unter kontrollierten Bedingungen. Die kognitiven Routinen, auf die das 4D-Rahmenwerk zielt, bilden sich über Jahre der Praxis hinweg, nicht in Minuten der Laborbeschäftigung. Forschungsdesigns, die rahmenrelevante Effekte über entwicklungsbezogene Zeithorizonte hinweg erkennen können, sind longitudinal, kostspielig und methodologisch anspruchsvoll. Sie sind auch notwendig. Ohne sie wird das Feld weiterhin kurzfristige Effekte der Aufgabenausführung mit langfristigen Effekten der Routinebildung verwechseln, mit ernststen Folgen für institutionelle Entscheidungen. Zweitens hat das Feld zu wenig in qualitative und ethnografische Arbeit investiert, die dokumentiert, wie KI-Nutzung im Klassenraum tatsächlich aussieht, sowohl wie sie angeleitet als auch wie sie praktiziert wird. Die ethnografische Aufzeichnung wird zu wichtigem historischen Material werden, während sich die Technologiegeneration wandelt; wir sollten sie jetzt aufbauen. Drittens ist die empirische Literatur derzeit zugunsten von Interventionsstudien gegenüber Nicht-Interventionsstudien gewichtet. Wir wissen mehr darüber, was geschieht, wenn Lernenden beigebracht wird, KI auf eine spezifische Weise zu nutzen, als darüber, was geschieht, wenn sie es nicht gelehrt bekommen. Die Kontrollbedingung (ungeführte Standardnutzung) ist die dominante Lernendenerfahrung, und sie ist methodologisch unterforscht.

Abschließend eine Anmerkung zur Positionalität des Autors. Das hier angebotene Rahmenwerk wurde von einem einzelnen Autor mit operativer Lehrerfahrung über anglo-amerikanische, tschechische und US-amerikanisch affilierte tertiäre Institutionen hinweg und Beratungserfahrung über mehrere KI-einsetzende Branchen hinweg entwickelt. Es ist nicht das Produkt einer Forschungsgruppe, einer multi-institutionellen Zusammenarbeit oder eines finanzierten Forschungsprogramms. Seine Stärken sind Kohärenz, operative Spezifität und Aktualität. Seine Schwächen sind die Beschränkung einer einzelnen Perspektive, das Fehlen formaler Peer-Review in dieser Phase und die spezifischen kulturellen und institutionellen Verzerrungen der beruflichen Trajektorie des Autors. Die Leser sollten die Empfehlungen entsprechend gewichten. Das Rahmenwerk wird als ernsthafter Beitrag zur Arbeitsdiskussion des Feldes angeboten, nicht als fertiges Forschungsprodukt. Konstruktive Auseinandersetzung, methodologische Kritik und empirische Prüfung der Behauptungen des Rahmenwerks sind ausdrücklich willkommen.

13. Schluss

Der Übergang, der in der Bildung jetzt ansteht, ist primär kein Übergang der Technologie. Er betrifft die Frage, ob sich Bildungsinstitutionen schnell, intelligent und menschlich genug entwickeln können, um Institutionen zu bleiben, durch die die nächste Generation die kognitiven Kapazitäten formt, die eine freie Gesellschaft erfordert. Grenzgängige generative KI ist in einer Lesart der mächtigste kognitive Partner, den unsere Spezies je erfunden hat. In einer anderen ist sie der mächtigste kognitive Ersatz. Die pädagogische Frage lautet, welches von beiden sie für das Kind wird, das jetzt in der

Grundschule ist, für den Jugendlichen, der jetzt in der Sekundarstufe ist, für den jungen Erwachsenen, der jetzt im Tertiärbereich ist. Die Antwort hängt davon ab, was wir ihnen darüber beibringen, in welcher Reihenfolge, mit welchem Schutz, im Dienst welcher kognitiven Kapazität.

Das 4D-Rahmenwerk (Delegation, Beschreibung, Urteilsvermögen, Sorgfalt) bietet eine strukturierte Möglichkeit, KI-Flüssigkeit als gestufte Kompetenz und nicht als undifferenzierte Erlaubnis zu vermitteln. Die hier vorgeschlagene entwicklungsbezogene Staffelung (schützende KI-nahe Bildung in der Primarbildung, gerüstgestützte KI-Kompetenz mit expliziter Reibung in der Sekundarbildung, operative KI-Flüssigkeit in der Tertiärbildung) bietet eine strukturierte Möglichkeit, diese Kompetenz mit dem sich entwickelnden Gehirn in Einklang zu bringen. Die Modellauswahlkriterien (Methodentransparenz, regulatorische Kooperation, Investition in Sicherheitsforschung) bieten eine strukturierte Möglichkeit für Institutionen, treuhänderisches Urteilsvermögen auszuüben. Die internationalen Politikrahmen, die nun konvergieren (EU-KI-Verordnung, UNESCO-Kompetenzrahmen, EU-OECD-AILit, PISA 2029 MAIL), liefern die Architektur der Kompetenzergebnisse, an der die institutionelle Reform kalibriert werden kann.

Keiner dieser Vorschläge ist endgültig. Alle sind Arbeitshypothesen, gedacht zur Revision, während sich Evidenz ansammelt und die Technologie sich weiterentwickelt. Ich bin überzeugt, dass die gegenwärtigen dominanten Haltungen (Verbot, unregulierte Übernahme, überwachungsbasierte Kontrolle) jeweils auf ihre eigene Weise entwicklungsbezogen ungebildet sind. Sie können das kommende Jahrzehnt nicht überstehen. Institutionen, die jetzt mit operativer Reform beginnen, mit gezieltem Curriculumsneudesign, Lehrkräftevorbereitung, Bewertungsneudesign und Schutz-durch-Voreinstellung-Infrastruktur, werden positioniert sein, ihren Lernenden zu dienen. Institutionen, die Reformen aufschieben, werden steigende Aufholkosten unter wachsendem externen Druck zahlen. Die nächste Generation mit KI zu denken zu lehren statt um sie herum, erfordert Lehrkräfte, die bereit sind, dasselbe zu tun.

Literatur

- AI Literacy Institute. (2026, 3. März). AI Literacy Review – March 3, 2026.
<https://ailiteracy.institute/ai-literacy-review-march-3-2026/>
- Anthropic. (2025). AI Fluency for Educators [Kurs]. Anthropic. <https://www.anthropic.com/learn>
- Anthropic. (2026). Responsible Scaling Policy, Version 3.0. Anthropic.
<https://www.anthropic.com/responsible-scaling-policy>
- Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., McKinnon, C., Chen, C., Olsson, C., Olah, C., Hernandez, D., Drain, D., Ganguli, D., Li, D., Tran-Johnson, E., Perez, E., ... Kaplan, J. (2022). Constitutional AI: Harmlessness from AI feedback. arXiv preprint arXiv:2212.08073. <https://doi.org/10.48550/arXiv.2212.08073>
- California Department of Education. (2025). Artificial Intelligence in Education: Guidance for K-12 Schools. California Department of Education. <https://www.cde.ca.gov/ci/pl/aiincalifornia.asp>
- Chew, H. E., Soon, C., & Kaur, H. (2026). Unified AI Literacy Framework for Singapore. Institute of Policy Studies, National University of Singapore.
- Cunningham, C. J. L., et al. (2025). One year in the classroom with ChatGPT: empirical insights and transformative impacts. *Frontiers in Education*, 10, 1574477.
<https://doi.org/10.3389/feduc.2025.1574477>
- Deng, R., Wang, X., & Zhang, Y. (2026). ChatGPT's impact on student learning outcomes: A meta-analysis of 35 experimental studies. *Humanities and Social Sciences Communications*, advance online publication. <https://doi.org/10.1057/s41599-026-07019-z>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
<https://doi.org/10.1146/annurev-psych-113011-143750>
- Digital Promise. (2024). Revealing an AI Literacy Framework for Learners and Educators. Digital Promise. <https://digitalpromise.org/2024/02/21/revealing-an-ai-literacy-framework-for-learners-and-educators/>
- Ergene, Ö., & Ergene, B. Ç. (2025). Cognitive offloading and over-dependence on generative AI in higher education. *Education and Information Technologies*, advance online publication.
- Europäische Kommission, & OECD. (2025). Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education [Entwurf zur Begutachtung]. Europäische Kommission & OECD. <https://ailiteracyframework.org/>
- Europäische Union. (2024). Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (KI-Verordnung). Amtsblatt der Europäischen Union, L 1689/1.
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530.
- Gascón-Cánovas, J. J., et al. (2025). Executive function in adolescents with problematic smartphone use. *International Journal of Adolescence and Youth*, 30(1), Artikel 2562950.
<https://doi.org/10.1080/02673843.2025.2562950>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- Gilbert, S. J. (2024). Cognitive offloading is value-based decision making: Modelling cognitive effort and the expected value of memory. *Cognition*, 247, 105783.
- Hinze, V., Blakemore, S.-J., Dalgleish, T., Ford, T., Mansfield, K. L., Ukoumunne, O. C., Kuyken, W., & Montero-Marin, J. (2025). The longitudinal relationship between self-reported executive function and mental health in early adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry – Open*, advance online publication.
<https://doi.org/10.1016/j.jaacop.2025.07.003>
- Hutton, J. S., Dudley, J., Horowitz-Kraus, T., et al. (2022). Associations between screen-based media use and brain white matter integrity in preschool-aged children. *JAMA Pediatrics*, advance online publication.

- Illingworth, S. (2026). What is critical AI literacy? AI Literacy Institute Working Papers.
- Kharbach, M. (2026). UNESCO's AI Competency Framework for Students: What every educator should know. medkharbach.com.
- Kooli, C., et al. (2025). Generative AI Addiction Syndrome: A conceptual framework. *Computers in Human Behavior Reports*, advance online publication.
- Kosmyna, N., Hauptmann, S. C., Yuan, T., Situ, Y., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks. MIT Media Lab Technical Report. <https://doi.org/10.48550/arXiv.2506.08872>
- Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2024). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 7, 100101.
- Lee, J., & Low, V. (2024). Generative AI in academic writing: Trade-offs between efficiency and originality. *Computers & Education: Artificial Intelligence*, 7, 100268.
- León-Domínguez, U. (2024). Potential cognitive risks of generative transformer-based AI chatbots on higher-order executive functions. *Neuropsychology*, 38(4), 293–308. <https://doi.org/10.1037/neu0000948>
- Lewin, K. M., et al. (2025). The longitudinal impact of screen media activities on brain function, architecture and mental health in early adolescence. *International Journal of Clinical and Health Psychology*, 25(2).
- Lin, Z., Dai, Y., & Ng, O.-L. (2025). Constructionism in K-12 AI literacy education: A systematic review of pedagogical designs, student outcomes, and learning mechanisms. *Journal of Educational Computing Research*, advance online publication. <https://doi.org/10.1177/07356331251360442>
- Martin, F., Bacak, J., Polly, D., & Wang, W. (2024). A systematic review of research on K-12 AI literacy. *Computers and Education: Artificial Intelligence*, 7, 100295.
- Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO.
- OECD. (2026). PISA 2029 Media and Artificial Intelligence Literacy (MAIL) Assessment Framework: Navigating an Evolving Digital World [Entwurf]. OECD Publishing. <https://www.oecd.org/en/about/projects/pisa-2029-media-and-artificial-intelligence-literacy.html>
- Perrigo, B. (2026, 24. Februar). Exclusive: Anthropic drops flagship safety pledge. TIME. <https://time.com/7380854/exclusive-anthropic-drops-flagship-safety-pledge/>
- Raptopoulou, A. (2025). ChatGPT in higher education: Supporting academic literacy through ChatGPT-based activities. *European Journal of Education*, e70131. <https://doi.org/10.1111/ejed.70131>
- Reggente, N., et al. (2026). The impact of smartphone use on brain function in adolescence: A scoping review. *Pediatric Reports*, 18(2), 43. <https://doi.org/10.3390/pediatric18020043>
- Roe, J., Furze, L., & Perkins, M. (2026). Digital plastic: A metaphorical framework for critical AI literacy in the multiliteracies era. *Postdigital Science and Education*, advance online publication.
- Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, 28(1), 78–106. <https://doi.org/10.1016/j.dr.2007.08.002>
- Su, J., Ng, D. T. K., & Chu, S. K. W. (2025). Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 education from a systematic review. *Interactive Learning Environments*, 33(9), 5766–5788. <https://doi.org/10.1080/10494820.2025.2487538>
- Thomas, J. W. (2000). A review of research on project-based learning. Autodesk Foundation.
- UNESCO. (2024). AI competency framework for students and teachers. UNESCO Publishing.
- Welker, M. (2025). Global AI literacy readiness assessment 2025. UNESCO/UNICEF Working Paper.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>

- Xing, B., et al. (2025). ChatGPT in K-12 education: Effects on student learning outcomes. Journal of Educational Computing Research, advance online publication.
- Zhang, S., Xiao, R., Botelho, A. F., Liao, G., Chiu, T. K. F., Stamper, J., & Koedinger, K. R. (2026). How to assess AI literacy: Misalignment between self-reported and objective-based measures. arXiv preprint arXiv:2601.06101.

Hinweis zur Übersetzung. Dies ist eine deutsche Übersetzung des englischen Originals "Teaching With, Not Around" (Berky, 2026). Die Übersetzung wurde vom Autor mit KI-Unterstützung und anschließender Autorenrevison erstellt. Eine muttersprachliche akademische Überprüfung wird vor der Veröffentlichung auf Zenodo empfohlen. Die kanonische Zitierform bleibt das englische Original. Die Literaturangaben sind in der Sprache der ursprünglichen Veröffentlichung belassen.

© 2026 N.J. Berky. Dieses Arbeitspapier wird unter der Creative Commons Attribution 4.0 International Lizenz (CC BY 4.0) veröffentlicht. Sie dürfen es mit angemessener Zuschreibung teilen und anpassen. Persistenter Identifikator · ORCID 0009-0001-2227-215X. Empfohlene Zitierform · Berky, N.J. (2026). Mit KI unterrichten, nicht um sie herum: Ein entwicklungsorientiertes Rahmenwerk für die Integration generativer KI in der Primar-, Sekundar- und Tertiärbildung (2023–2026). Arbeitspapier, Version 2.0 (deutsche Übersetzung). NJB Berky Advisory.