„Schule in der Kultur der Digitalität – digitale Bildungstechnologie und Künstliche Intelligenz im Unterrichtsprozess und im Rahmen individueller Förderung“

Forschungsgutachten

Zur Vorlage bei der Enquetekommission I „Chancengleichheit in der Bildung”
des Landtags Nordrhein-Westfalen

02. Oktober 2024

LANDTAG
NORDRHEIN-WESTFALEN
18. WAHLPERIODE

INFORMATION
18/192

A40

Prof. Dr. Nikol Rummel
Center for Advanced Internet Studies &
Ruhr-Universität Bochum

Ann-Christin Falhs
Center for Advanced Internet Studies &
Ruhr-Universität Bochum

Dr. Astrid Wichmann
Center for Advanced Internet Studies &
Ruhr-Universität Bochum

Prof. Dr. Vincent Aleven
Carnegie Mellon Universität, Pittsburgh, USA

Prof. Dr. Nikol Rummel

Prof. Dr. Nikol Rummel ist Professorin für Pädagogische Psychologie und Bildungstechnologie der Ruhr-Universität Bochum. Als Forschungsprofessorin am Center for Advanced Internet Studies (CAIS) und assoziierte Professorin an der Carnegie Mellon University erforscht sie, wie Lehr- und Lernprozesse in technologiegestützten Umgebungen unterstützt und verbessert werden können. Dabei blickt sie aus einer pädagogisch-psychologischen Perspektive auf Lehr- und Lernprozesse und beleuchtet die mögliche Rolle, die Bildungstechnologien dabei einnehmen können. Sie war Mitglied der vom CAIS initiierten Taskforce „Künstliche Intelligenz im Bildungswesen“, die Handlungsempfehlungen für die Landesregierung des Landes Nordrhein-Westfalen erarbeitet hat.

Ann-Christin Falhs, M.A.

Ann-Christin Falhs ist Doktorandin und wissenschaftliche Mitarbeiterin am Lehrstuhl für Pädagogische Psychologie und Bildungstechnologie der Ruhr-Universität Bochum sowie im Forschungsprogramm „Bildungstechnologien und Künstliche Intelligenz” von Prof. Dr. Nikol Rummel am Center for Advanced Internet Studies (CAIS). Im Handlungsfeld der Komplementarität zwischen Lehrkräften und Technologie forscht sie im Rahmen ihrer Dissertation zum Einsatz digitaler Lernsoftware im Unterricht mit einem Fokus auf intelligente tutorielle Systeme (ITS).

Dr. Astrid Wichmann

Dr. Astrid Wichmann ist Wissenschaftsmanagerin am Lehrstuhl für Pädagogische Psychologie und Bildungstechnologie der Ruhr-Universität Bochum sowie im Forschungsprogramm „Bildungstechnologien und Künstliche Intelligenz” von Prof. Dr. Nikol Rummel am Center for Advanced Internet Studies (CAIS).

Prof. Dr. Vincent Aleven

Prof. Dr. Vincent Aleven ist Professor für Mensch-Computer-Interaktion am Institut für Human-Computer Interaction der Carnegie Mellon University, Pittsburgh. Seine Forschung konzentriert sich auf das intelligente Klassenzimmer der Zukunft, einschließlich Intelligenter Tutoring-Systeme (ITS). Er ist ein führender Experte in der Erforschung des Einsatzes von KI-gestützten Technologien in Lernkontexten und Mitbegründer von Carnegie Learning, Inc.

Dank

Die Autor:innen bedanken sich bei Anna Radtke (CAIS) für ihre Unterstützung bei der Literaturrecherche und -sichtung, bei Dr. Matthias Begenat (CAIS) und Dr. Valentina Nachtigall (RUB) für ihre redaktionelle Arbeit, bei Sarah Akyürek für ihre Hilfe bei der Gestaltung sowie bei Dr. Sebastian Strauß (RUB) für seine Unterstützung bei der Formatierung.

Anmerkung: Die in diesem Gutachten genannten kommerziellen Produkte (z.B. Bettermarks, MATHia, KhanMigo, ChatGPT) werden in diesem Text als Beispiel für die ihnen zugrundeliegende Technologie bzw. deren Anwendungsfelder genannt, damit ist weder eine implizite noch explizite Bewerbung oder Bewertung dieser Produkte verbunden.

Inhalt

1. Einleitung 3
2. Bildungspolitischer Rahmen 5
   2.1. Individuelle Förderung und Inklusion 5
   2.2. Nutzung digitaler Bildungstechnologien 7
3. Zentrale Desiderata für individuelle Förderung 8
4. Systematisierung von KI-Systemen in der Bildung 10
5. Analyse des Potenzials von KI-Technologien für individuelle Förderung 14
   5.1. Cluster 1: Systeme zur Lerndatenanalyse 14
      Exemplarische Anwendungen 15
      Empirische Befunde 18
      Einordnung 19
   5.2. Cluster 2: Systeme zur adaptiven Lernunterstützung 20
      Exemplarische Anwendungen 21
      Empirische Befunde 24
      Einordnung 25
   5.3. Cluster 3: Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe 26
      Exemplarische Anwendungen 26
      Empirische Befunde 29
      Einordnung 30
   5.4. Fazit der Potenzialanalyse 30
6. Handlungsempfehlungen 32
7. Ausblick 37
   7.1. Zentrale Herausforderungen 37
      Umgang mit Daten 37
      Nutzung von generativer KI im Bildungsbereich 39
   7.2. Neue Entwicklungsrichtungen 40
      Entwicklungsaufwand für Intelligente Tutoring-Systeme reduzieren 40
      Stärken symbolischer und generativer KI kombinieren 40
      Stärken von Lehrkräften und KI kombinieren 40

Referenzen 42

1. Einleitung

In der heutigen Kultur der Digitalität nehmen (künstlich-intelligente) Technologien eine zentrale Rolle im gesellschaftlichen Leben ein. Folglich ist das Bildungswesen aufgerufen, im Rahmen seines Erziehungs- und Bildungsauftrags entsprechende Bildungsmöglichkeiten bereitzustellen. Gleichzeitig bietet die Digitalisierung die Chance, bei einer zunehmend heterogenen Schülerschaft dem individuellen Unterstützungsbedarf von Schüler:innen Rechnung zu tragen, um Bildungsungleichheiten abzubauen und Chancengerechtigkeit zu fördern.

Es ist zu untersuchen, inwiefern der Einsatz von Künstlicher Intelligenz (KI) im Bildungssektor neue Wege eröffnet, Schüler:innen individuell zu fördern und Lehrkräfte zu entlasten. Daher lautet die zentrale Frage des Gutachtens: Kann KI insbesondere in heterogenen Lerngruppen, in denen Lernende mit unterschiedlichen Fähigkeiten, Stärken und Schwächen zusammenkommen, dazu beitragen, den Bedürfnissen aller Lernenden gerecht zu werden? Um diese Frage zu beantworten, werden in diesem Gutachten Bildungstechnologien und insbesondere künstlich-intelligente (KI) Systeme dahingehend untersucht, inwiefern sie das Potenzial bieten, individuelles Lernen zu gestalten, Binnendifferenzierung zu fördern sowie Unterstützung bei Schwächen und Förderbedarfen zu ermöglichen.

Zunächst werden die bildungspolitischen Rahmenbedingungen umrissen, die den Hintergrund der Diskussion bilden. Dabei wird aufgezeigt, welche Schritte das Land Nordrhein-Westfalen bislang gegangen ist, um einerseits individuelle Förderung und schulische Inklusion zu gestalten sowie andererseits den Einsatz digitaler Bildungstechnologien und KI im Unterrichtsprozess zu fördern. Es werden relevante Veröffentlichungen der Kultusministerkonferenz (KMK) und der Ständigen Wissenschaftlichen Kommission (SWK) sowie Richtlinien, Empfehlungen und Konzepte betrachtet, die den Einsatz digitaler Bildungstechnologien und KI im Unterrichtsprozess und bei der individuellen Förderung vorsehen. Im Anschluss werden zentrale Desiderata für die individuelle Förderung von Lernen abgeleitet.

Zur Einordnung werden grundlegende Begrifflichkeiten im Bereich künstlich-intelligenter (KI) Systeme erklärt und Gruppen KI-gestützter Lehr- und Lernanwendungen systematisiert. In der Potenzialanalyse werden drei Cluster von Systemen fokussiert, die für die adaptive, individuelle Lernunterstützung eingesetzt werden können: Systeme zur Lerndatenanalyse, Systeme zur adaptiven Lernunterstützung und Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe. Für diese Cluster werden jeweils exemplarische Anwendungen beschrieben und dahingehend untersucht, inwiefern sie bereits KI-Unterstützung nutzen. Wissenschaftliche Studien und empirische Befunde werden betrachtet, um den Kenntnisstand zu den einzelnen Clustern zu erörtern. Es wird festgehalten, welche Erkenntnisse hinsichtlich des Potenzials für bedarfsangepasste Unterstützung und individuelle Förderung bereits bestehen. Die Potenzialanalyse bildet den Ausgangspunkt für Handlungsempfehlungen. Im abschließenden Ausblick werden zentrale Herausforderungen aufgezeigt und neue Entwicklungsrichtungen skizziert.

2. Bildungspolitischer Rahmen

2.1. Individuelle Förderung und Inklusion

Paragraph 1 des Schulgesetzes Nordrhein-Westfalen legt den Auftrag der Schule fest. Dort wird nicht nur das Recht auf Bildung und Erziehung, sondern auch das Recht auf individuelle Förderung explizit benannt:

> (1) Jeder junge Mensch hat ohne Rücksicht auf seine wirtschaftliche Lage und Herkunft und sein Geschlecht ein Recht auf schulische Bildung, Erziehung und individuelle Förderung. Dieses Recht wird nach Maßgabe dieses Gesetzes gewährleistet.
>
> (2) Die Fähigkeiten und Neigungen des jungen Menschen sowie der Wille der Eltern bestimmen seinen Bildungsweg. Der Zugang zur schulischen Bildung steht jeder Schülerin und jedem Schüler nach Lernbereitschaft und Leistungsfähigkeit offen.
>
> (Schulgesetz für das Land Nordrhein-Westfalen [SchulG NRW], 2024)

Dem Willen des Gesetzgebers gemäß kommt also dem Recht auf individuelle Förderung der gleiche Stellenwert zu wie dem Recht auf Bildung und Erziehung. Damit werden die einzelnen Schüler:innen in den Fokus genommen. Die Heterogenität in den Lerngruppen und damit die Vielfältigkeit der Lernbedürfnisse bilden jedoch zentrale Herausforderungen für die individuell zugeschnittene Vermittlung von Lernerfahrungen, insbesondere bei Kindern mit besonderem Förderbedarf. Jedem Kind Lernerfahrungen zu vermitteln, die auf eben dieses Kind zugeschnitten sind, ist eine große Herausforderung - gerade bei den Kindern, die einen besonderen Förderbedarf haben.

Im Jahr 2009 verpflichtete sich Deutschland mit der Ratifizierung der UN-Behindertenrechtskonvention von 2006 dazu, ein Schulsystem zu ermöglichen, welches die Idee der Inklusion verwirklicht. Die Umsetzung liegt „nach der Kompetenzordnung des Grundgesetzes" (Kultusministerkonferenz [KMK], 2010, S. 2) in der Verantwortung der Länder und Kommunen. Dabei hat die Kultusministerkonferenz in den letzten Jahren durch Beschlüsse, Empfehlungen und Vereinbarungen den Rahmen definiert, durch den inklusive Bildung an Schulen gestaltet werden kann. Mit dem Beschluss vom 18. November 2010 wurde initiiert, dass Schüler:innen mit und ohne sonderpädagogischen Förderbedarf gemeinsam zielgleich oder zieldifferent in allgemeinen Schulen lernen können (Kultusministerkonferenz [KMK], 2010). Ziel ist eine bestmögliche Entfaltung sowie ein „höchstmögliches Maß an Aktivität und gleichberechtigter Teilhabe" (S. 4) durch die Förderung individueller Kompetenzen und Fähigkeiten aller Schüler:innen. Darüber hinaus soll nicht nur das Lernen, sondern auch die „Lernstands- und Leistungsmessung sowie die Leistungsbewertung [...] in Bezug auf diese individuellen Ziele erfolgen" (KMK, 2010, S. 4). Der Beschluss stellt heraus, dass die Umsetzung des Übereinkommens einer progressiven Realisierung unterliegt, die es immer wieder erforderlich macht, Weiterentwicklungen zu prüfen und erforderliche Maßnahmen zu veranlassen. So auch in Bezug auf die Möglichkeiten, welche sich durch den Einbezug KI-gestützter Technologien ergeben.

Mit dem Beschluss vom 20. Oktober 2011 hat die Kultusministerkonferenz die Empfehlungen zur inklusiven Bildung konkretisiert. Daraus geht hervor, dass inklusiver Unterricht Fähigkeiten auf Seiten der Lehrenden erfordert, die u.a. die Lern- und Entwicklungsbegleitung, die Individualisierung des Lernens auf diagnostischer Grundlage sowie die Planung und Differenzierung von Unterrichtsprozessen betreffen. Dieser Beschluss stellt den notwendigen Rahmen dar, damit Lehrkräfte die für inklusive Bildung nötigen Kompetenzen erwerben können. Dafür wurden auch die „Standards für die Lehrerbildung: Bildungswissenschaften“ (KMK, 2004) sowie die „Länder-gemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und Fach-didaktiken in der Lehrerbildung“ (KMK, 2024) überarbeitet, indem u.a. „Diversität und Heterogenität als Bedingungen von Schule und Unterricht” (KMK, 2019, S. 5) anerkannt und „Differenzierung, Integration und Förderung” (KMK, 2019, S. 5) somit zu curricularen Ausbildungsschwerpunkten werden. Dies bedeutet, dass Unterricht nach unterschied-lichen Lernvoraussetzungen und Entwicklungsprozessen geplant werden soll. Lehrkräfte haben damit etwaige Benachteiligungen, Beeinträchtigungen und Barrieren der Schüler:innen zu berücksichtigen.

In Nordrhein-Westfalen bildet das am 1. August 2014 in Kraft getretene 9. Schulrechts-änderungsgesetz die Grundlage zur Umsetzung inklusiver schulischer Bildung. Durch die Änderung des Schulrechtgesetzes ist geregelt, dass „das gemeinsame zielgleiche und zieldifferente Lernen von Schülerinnen und Schülern mit und ohne Behinderung in der allgemeinen Schule ermöglicht werden” soll (Bereinigte Amtliche Sammlung der Schulvorschriften NRW [BASS], 2013). Somit wurde inklusive Bildung in allgemeinen Schulen im Bundesland Nordrhein-Westfalen zum gesetzlich verankerten Regelfall. Das bedeutet auch, dass Schüler:innen mit einem Bedarf an sonderpädagogischer Unterstützung allgemeine Schulen besuchen können[1]. Dabei ist der Anteil der Schüler:innen mit sonderpädagogischer Förderung in allgemeinen Schulen bundesweit seit Jahren steigend (genaue Zahlen können der Statistischen Veröffentlichung der Kultusministerkonferenz „Sonderpädagogische Förderung in Schulen 2013 bis 2022” entnommen werden; siehe Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024). In Nordrhein-Westfalen beträgt die Förderquote, also der Anteil der Schüler:innen mit Förderbedarf von allen Schüler:innen mit Vollzeitschulpflicht, im Schuljahr 2021/22 8,9 %, wovon 4 % der Schüler:innen inklusiv in allgemeinen Schulen unterrichtet wurden (Klemm et al., 2023). Für diese Schüler:innen muss eine entsprechende Förderung im Unterricht stattfinden. Dabei nimmt auch die Diagnostik von Lernprozessen und Lernverläufen einen wichtigen Stellenwert ein. In diesem Kontext sind die Bereitstellung und Analyse förderrelevanter Informationen wichtig, sowohl um unmittelbar situationsadäquate Entscheidungen zur Unterstützung

[1] „Das 9. Schulrechtsänderungsgesetz umfasst (noch) nicht das Recht auf den Besuch einer von den Eltern gewünschten konkreten allgemeinen Schule, sondern betrifft zunächst das Recht, überhaupt ein Angebot des Gemeinsamen Lernens zu erhalten. Dieses Recht ist aber noch nicht schrankenlos. Die Schulaufsicht kann die Förderschule statt der allgemeinen Schule oder die allgemeine Schule statt der Förderschule festlegen, wenn „die personellen und sächlichen Voraussetzungen am gewählten Förderort nicht erfüllt sind und auch nicht mit vertretbarem Aufwand erfüllt werden können“ (§ 20 Absatz 4 SchulG).” (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW], 2024b)

der Schüler:innen zu treffen, als auch, um längerfristig auf Lernziele hinzuarbeiten und diese zu erreichen.

2.2. Nutzung digitaler Bildungstechnologien

Die KMK hat im Jahr 2016 mit dem Strategiepapier „Bildung in der digitalen Welt“ ein grundlegendes Handlungskonzept für die zukünftige Entwicklung der Bildung in Deutschland vorgelegt. Bereits dort wird als eines von zwei Zielen für Schulen genannt, dass die durch digitale Bildungstechnologien neu zur Verfügung stehenden Möglichkeiten zur Umsetzung individueller Förderung im Unterricht genutzt werden sollen (KMK, 2016). Und weiter:

> Mit zunehmender Digitalisierung entwickelt sich auch die Rolle der Lehrkräfte weiter. Die lernbegleitenden Funktionen der Lehrkräfte gewinnen an Gewicht. Gerade die zunehmende Heterogenität von Lerngruppen, auch im Hinblick auf die inklusive Bildung, macht es erforderlich, individualisierte Lernarrangements zu entwickeln und verfügbar zu machen. Digitale Lernumgebungen können hier die notwendigen Freiräume schaffen; allerdings bedarf es einer Neuausrichtung der bisherigen Unterrichtskonzepte, um die Potenziale digitaler Lernumgebungen wirksam werden zu lassen.
>
> (KMK, 2016, S. 13)

Im Jahr 2021 folgte eine auf dem Strategiepapier aufbauende ergänzende Empfehlung der KMK „Lehren und Lernen in der digitalen Welt“, welche einzelne Aspekte der Strategie für den Bildungsbereich Schule vertieft und das Potenzial für das Lehren und Lernen benennt, indem einerseits die Unterstützung von Barrierefreiheit und der Zugang zu Bildungsinhalten für Personen mit individuellen Beeinträchtigungen sowie andererseits individuelles zeitnahes Feedback sowie Rückmeldungen zu Lernprozessen und Lernergebnissen hervorgehoben werden. Dies umfasst auch Rückmeldungen aus dem Einsatz diagnostischer Instrumente, welche als zentral für erfolgreiches Lernen angesehen werden (KMK, 2021). Dabei können „Computergestützte, interaktive und sanktionsfreie Rückmeldungen [...] durch Adaptivität das individualisierte Üben gezielt unterstützen” (KMK, 2021, S. 10).

In Bezug auf den Einsatz von KI in Schulen gibt es bisher keine einheitlichen Regularien[2], wobei der *AI Act*, welcher am 1. August 2024 in Kraft getreten ist, nun sukzessive Anwendung findet. Bis dahin müssen die weiteren Zuständigkeiten noch im nationalen Recht geregelt werden (Landesbeauftragte für Datenschutz und Informationsfreiheit Nordrhein-Westfalen [LDI NRW], 2024).

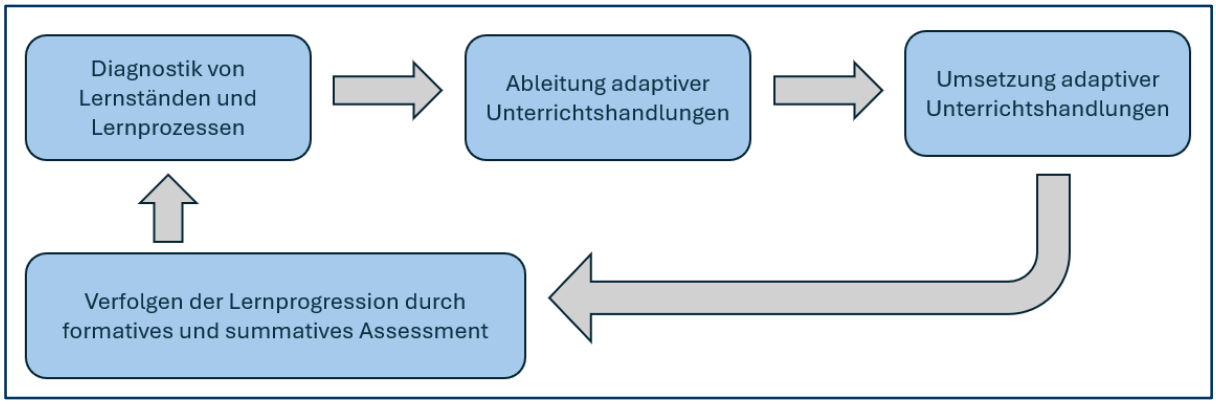
[2] Das Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen stellt Lehrkräften und Schulleitungen über die Webseite www.lernen-digital.nrw ein Unterstützungsangebot zur Verfügung. Dort werden Angebote zum sogenannten Handlungswissen (Lernen über KI) sowie Best Practices und fachbezogene Praxisbeispiele zum lernförderlichen Einsatz von KI genannt. Auch finden sich dort die handlungsleitenden Hinweise und Empfehlungen wie den Handlungsleitfaden KI „Umgang mit textgenerierenden KI-Systemen“ (MSB NRW, 2024).

Der hier skizzierte bildungspolitische Rahmen unterstreicht die Chancen, aber auch die Notwendigkeit, Potenziale von KI-Technologien im schulischen Bereich einzubinden.

3. Zentrale Desiderata für individuelle Förderung

Um individualisierten und differenzierten Unterricht zu ermöglichen, bedarf es verschiedener Schritte (siehe Abbildung 1).

Abbildung 1: Schritte individualisierten und differenzierten Unterrichts (in Anlehnung an Vock & Gronostaj, 2017)



Eine entscheidende Voraussetzung für eine individuelle Lernförderung sind die Verfolgung der Lernfortschritte sowie die Diagnostik von Lernständen und -prozessen. Eine präzise Diagnose[3] und kontinuierliche Überwachung des Lernstandes sind besonders im inklusiven Bildungsumfeld unerlässlich, um auf die individuellen Bedürfnisse der Lernenden einzugehen und personalisierte Fördermaßnahmen zu entwickeln. Dies stellt eine große Herausforderung dar, da solche Analysen zeitaufwendig sind und oft nur begrenzt auf einzelne Schüler:innen zugeschnitten werden können. Neben klassischen papierbasierten Formen werden bereits seit längerer Zeit computergestützte Testverfahren eingesetzt. Computergestützte Testverfahren erlauben eine Optimierung der Testung, indem die vorgegebenen Testitems aufgrund bereits gegebener Antworten adaptiv ausgewählt werden. So kann eine möglichst genaue Diagnostik mit möglichst wenigen Testitems erreicht werden. Mittlerweile ist mit computergestützten Testverfahren nicht nur eine Lernstandsdiagnostik aufgrund gegebener Testantworten, sondern zunehmend auch eine Lernverlaufsdiagnostik möglich (Goldhammer, Scherer & Greiff, 2019). Der Einsatz von KI bietet weiterführende Möglichkeiten, um Lernverhalten zu analysieren, Kompetenzen und Schwächen zu identifizieren und, darauf abgestimmt, sinnvoll und gezielt Fördermöglichkeiten abzuleiten. Beispielsweise werden in künstlich-intelligenten Systemen, denen Wissens- und Lernermodelle zugrunde liegen, die Lernaktivitäten der Schüler:innen detailliert erfasst, um auf Abweichungen reagieren zu können. Weitere Möglichkeiten bieten

[3] Mit Diagnostik sind in diesem Gutachten nicht die Anamnese oder klinische Diagnosestellung gemeint, sondern das Sammeln und Analysieren von Lerndaten sowie die Interpretation der Ergebnisse im Rahmen der Lernstands- und Lernprozessdiagnostik.

Verfahren der automatischen Analyse und Auswertung von Daten mittels maschineller Lernverfahren. Auf beide Arten von künstlich-intelligenten Systemen wird weiter unten noch eingegangen.

Wie in Abbildung 1 dargestellt, sind das Ableiten und Umsetzen adaptiver Maßnahmen weitere zentrale Schritte für individualisierten und differenzierten Unterricht. Datenbasierte Erkenntnisse über den Lernfortschritt bzw. den Lernstand bilden die Grundlage für eine bedarfsangepasste Unterstützung der Lernenden. Passende Vorschläge für Lernmaterialien, Interventionen oder andere Fördermaßnahmen können entweder direkt an die Lehrkraft weitergegeben werden, um den Unterricht entsprechend anzupassen, oder die Lernenden werden durch adaptive digitale Systeme gezielt unterstützt. Dabei ist die Entwicklung von Lerntechnologien, die sich adaptiv an den Unterstützungsbedarf der Lernenden anpassen, nicht neu. Bereits seit den 1970er Jahren beschäftigt sich das internationale Forschungsgebiet „Artificial Intelligence in Education" (AIED, Deutsch: Künstliche Intelligenz in der Bildung; Du Boulay, Mitrovic & Yacef, 2023) mit der Entwicklung und Beforschung adaptiver Systeme.

Unter Adaptivität soll hier breit gefasst die Eigenschaft der Anpassung eines Systems an die spezifischen Bedürfnisse der Lernenden verstanden werden. Der Einsatz von KI eröffnet die Möglichkeit, personalisierte Lernpfade und differenziertes, fein-granulares Feedback für eine individuelle Unterstützung zu erreichen. Dabei gibt es unterschiedliche Möglichkeiten, wie sich ein intelligentes System an die Eigenschaften der Lernenden anpassen kann. Adaptive Lernumgebungen können verschiedene Parameter an die unterschiedlichen Lernvoraussetzungen und -fortschritte der Schüler:innen anpassen, z.B. die Lernzeit, den Instruktionsumfang sowie die Aufgabensequenz oder -schwierigkeit, (Leutner, 2002). Sie können dies zudem in unterschiedlicher Art und Weise tun: etwa auf unterschiedlichen Zeitskalen oder bezogen auf unterschiedliche Unterrichtsphasen. Dementsprechend liegt eine unüberschaubare Anzahl von Klassifikationen von Adaptionsmaßnahmen und Taxonomien adaptiver Methoden vor. Eine grundlegende und für dieses Gutachten hilfreiche Einordnung treffen Aleven und Kolleg:innen (2016). Sie unterscheiden drei Arten von Adaptivität, die auf unterschiedlichen Zeitskalen liegen: Es gibt adaptive Instruktionen auf einer sehr fein-granularen Ebene, die sich auf einzelne Problemlöseschritte beziehen, beispielsweise Hinweise oder Rückmeldungen (*step-loop adaptivity).* Etwas gröbere Adaptivität geschieht bei der Auswahl der jeweils nachfolgenden Aufgabe in Anpassung an den individuellen Lernstand (*task-loop adaptivity*). Hierbei wird auch erfasst, ob der Lernende ein bestimmtes Lernziel erreicht hat und somit zur nächsten Einheit fortschreiten kann (*mastery learning*). Schließlich nennen Aleven et al. (2016) noch die sogenannte *design-loop adaptivity*, die sich auf Änderungen in der Gestaltung des Systems (durch die Entwickler) auf der Grundlage der Analyse von Nutzungsdaten bezieht. Durch diese Form der Adaptivität wird sichergestellt, dass ein System an Schwierigkeiten angepasst wird, die viele Lernende haben. Eine weitere zentrale Frage ist, auf welche Merkmale der Lernenden die Lernumgebung bzw. die Instruktion angepasst werden soll. Hier werden fünf große Kategorien von relevanten Variablen genannt, die als Grundlage für eine

Anpassung der Lernunterstützung verwendet werden können: Wissen, Strategien und Fehler, affektive und motivationale Variablen, Selbstregulation sowie Lernstil. Um sich an diese Lernendenvariablen anpassen zu können, muss das adaptive System Zugang zu entsprechenden Daten haben und sie im zeitlichen Lernverlauf diagnostisch verfolgen. Hier schließt sich der Kreis zum Verfolgen der Lernaktivitäten und zur Diagnostik von Lernständen und Lernprozessen (siehe Abbildung 1).

Künstlich-intelligente Systeme bieten vielversprechende Möglichkeiten zur Unterstützung und Automatisierung sowohl im Hinblick auf die Analyse von Lernfortschritten, als auch darauf aufbauend zur adaptiven Unterstützung von Lernenden in Anpassung an ihren individuellen Lernstand.

4. Systematisierung von KI-Systemen in der Bildung

Eine aktuelle Definition von künstlich-intelligenten Systemen findet sich im Europäischen *AI Act*:

> 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.[4]
>
> (AI Act, 2024, Artikel 3 (1))

Diese Definition umfasst die zentralen Charakteristika künstlich-intelligenter Systeme: Eingaben, beispielsweise Problemlöseschritte beim Bearbeiten einer Aufgabe, werden vor dem Hintergrund von Zielen so interpretiert, dass Ausgaben entstehen. Mit Ausgaben sind beispielsweise systemseitige Rückmeldungen an die Lernenden zu ihren Lösungsschritten oder adaptiv ausgewählte Lernmaterialien gemeint. Dies geschieht mit einer gewissen Autonomie und adaptiv, also nicht nach einem festgelegten Schema, sondern abhängig von den erfolgenden Eingaben.

Stark vereinfacht liegen den verschiedenen KI-Anwendungen zwei grundsätzlich verschiedene Formen von KI zugrunde. **Symbolische oder wissensbasierte KI** nutzt Regeln und logische Verknüpfungen, um Entscheidungen zu treffen oder Probleme zu lösen. Wissen wird hier in Form von abstrakten Symbolen (z.B. Zahlen, Wörtern) repräsentiert. Diese Symbole können von Menschen geschrieben und gelesen werden, sind aber auch maschinenlesbar. Aufgrund der zugrundeliegenden Wissensmodellierung sind Systeme, die auf symbolischer KI basieren, inhalts- und fachspezifisch und sie sind aufwändig in der Entwicklung. Die Entscheidungen symbolischer KI-Systeme sind nachvollziehbar und vorhersagbar. **Statistische oder neuronale KI** hingegen nutzt

[4] „KI-System" ist ein maschinengestütztes System, das so konzipiert ist, dass es mit unterschiedlichen Graden an Autonomie operiert und sich entsprechend der Entwicklung adaptiv anpassen kann, und das, für explizite oder implizite Ziele aus den Eingaben, die es erhält, ableitet, wie es Ausgaben wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, die physische oder virtuelle Umgebungen beeinflussen können. (Übersetzung der Autor:innen)

statistische Verfahren, die bottom-up Muster aus Daten erkennen. Dazu werden verschiedene Techniken (z.B. *Deep Learning*) verwendet, um Modelle mit teils sehr großen Datensätzen zu trainieren, die dann auf Basis von statistischen Wahrscheinlichkeiten Ausgaben generieren. Dabei werden die Daten „unreflektiert" verwertet, d.h. jede Verzerrung in den Daten wird von dem maschinellen Lernprogramm übernommen. Zudem sind die Ergebnisse von Deep-Learning-Verfahren oft intransparent, da die Ergebnisse der Anwendung des gewählten Algorithmus nicht nachvollziehbar sind. Derartige statistische KI-Methoden liegen den aktuellen boomenden generativen KI-Technologien zugrunde. Generative KI bezeichnet KI-Systeme, die in der Lage sind, neue Inhalte zu erstellen. Dies können sowohl Texte, aber auch Bilder, Videos oder Musik sein. Large Language Modelle sind eine Form generativer KI-Systeme, die natürlichsprachliche Texte generieren. Sie werden mit großen Mengen von Textdaten aus dem Internet trainiert und können aufgrund einer komplexen Verarbeitung von Abhängigkeiten zwischen Wörtern Texte verstehen und generieren.

Sowohl symbolische, als auch neuronale KI-Verfahren werden in Technologien genutzt, die im Bildungskontext Anwendung finden. Die zugrundeliegenden KI-Verfahren haben wichtige Konsequenzen für die Funktionsweise der jeweiligen Bildungstechnologien. Damit einher gehen unterschiedliche Möglichkeiten für die Nutzung zur individuellen Förderung und adaptiven Unterstützung Lernender. Ebenso sind unterschiedliche Herausforderungen mit der Nutzung und mit der Entwicklung dieser Bildungstechnologien verbunden. Für die nachfolgende Potenzialanalyse ist daher die Systematisierung von KI-Technologien, die in der Schule Anwendung finden, eine wichtige Grundlage. Tabelle 1 zeigt sieben Technologiegruppen, (a)-(g), sowie ihre Kernaufgaben in der Bildung (DFKI, 2023). In der Tabelle wird auch illustriert, in welcher Weise Schüler:innen und/oder Lehrkräfte durch die jeweiligen Systeme im Lehr-Lernkontext unterstützt werden (in Anlehnung an Holmes & Tuomi, 2022).

Tabelle 1: Technologiegruppen und ihre Aufgaben sowie Nutzungsmöglichkeiten (in Anlehnung an DFKI, 2023; Holmes & Tuomi, 2022).

Technologiegruppen & Kernaufgaben in der Bildung	**Nutzungsmöglichkeiten von** Schüler:innen	Schüler: innen & Lehrkraft	Lehrkraft
(a) **Prüfungsunterstützende Systeme** Systeme, die für die Generierung und Korrektur von Aufgaben und Prüfungen eingesetzt werden.			Generierung von Übungs- und Prüfungsaufgaben und Korrektur/ Plagiatsprüfung
(b) **Learning Analytics & Educational Data Mining** Analyse von Lernverhalten, Vorhersagen oder frühzeitige Warnungen (z.B. zum Lernverlauf und -erfolg) und Evaluation von Bildungsprozessen.		Erzeugung automatischer (Lernstands-)Berichte und Diagnosen (Förderempfehlungen)	Lernstandsverfolgung und -dokumentation mittels Teacher Dashboards/ Identifizierung von Förderbedarfen
(c) **Bildungs- und unterrichts-organisierende Systeme** Systeme, die Administration und Organisation des Unterrichtsgeschehens erleichtern oder für die Planung und Management von Bildungsprozessen eingesetzt werden.			Unterstützung der Unterrichtsabläufe/ Automatisierung administrativer Aufgaben (Kurs-, Stunden- und Terminplangestaltung, Evaluationen)/ Intelligente Zusammenstellung von Lernmaterialien
(d) **Intelligente Tutoring- und Empfehlungssysteme** Wissensbasierte Intelligente Tutoring- und Empfehlungssysteme, die individuelles Feedback zu Lern- und Testleistungen geben oder basierend auf verschiedenen Merkmalen der Lernenden die Reihenfolge von Lerninhalten personalisiert anpassen bzw. empfehlen.	Personalisierte Anpassung der Reihenfolge von Lerninhalten oder Empfehlungen/ Individuelles Feedback zu Lern- und Testleistungen geben		
(e) **Text-to-Speech- und Speech-to-Text-Systeme** Umwandlung von Text in Sprache und umgekehrt, z.B. für Lernende mit Seh- oder Höreinschränkungen oder für die Erstellung von Sitzungsprotokollen im Kollegium.		Vorlesen von Text: Anpassung von Stimmen, Sprachgeschwindigkeit, Tonhöhe u.ä./ Erzeugen Text: Untertitel, Transkription, u.ä.	
(f) **Textgenerierende und -übersetzende Systeme** Generieren von Lehr- und Lerntexten.	Beantwortung von Fragen (sokratischer Dialogpartner)/ Feedback	Zusammenfassung, Vervollständigung, Paraphrasierung oder Übersetzung	Erstellung von Lernmaterial Differenzierung (nach Leistungsstufen oder Verständnisniveaus)
(g) **Material- und Bildgenerierende Systeme** Generieren von Unterrichtsmaterialien, z.B. Bilder, Videos, Präsentationen oder Arbeitsblätter		Erzeugung text- oder bildhaften Materials, das individuell nach eigenen Präferenzen oder im Hinblick auf die Schüler:innen individuell angepasst werden kann	

Da der Fokus dieses Gutachtens primär auf dem Potenzial von KI für Lehr- und Lernkontexte liegt, wird an dieser Stelle nicht näher auf prüfungsunterstützende Systeme (Technologiegruppe a) sowie den Bereich der Bildungsdatenanalyse für institutionelle und damit verbundene administrative Zwecke eingegangen (ein Teilbereich der Technologiegruppe c).

Stattdessen werden digitale (künstlich-intelligente) Technologien fokussiert, die für die adaptive, individuelle Lernunterstützung eingesetzt werden können. Dies umfasst auch die Förderung bei bestimmten Lernschwächen wie Lese-Rechtschreibschwäche (Legasthenie) oder Rechenschwäche (Dyskalkulie). Somit werden in der folgenden Potentialanalyse drei Schwerpunkte gesetzt:

Im Hinblick auf das Desideratum der Lernverlaufsanalyse und Diagnostik des Lernstands werden in **Cluster 1** zunächst Systeme betrachtet, die **Lerndaten automatisch analysieren**. Dieses Cluster umfasst somit im Kern Anwendungen aus der Technologiegruppe b. Große Datenmengen fallen während des Lernprozesses mit digitalen Technologien an und können wertvolle Einblicke in die Lerngewohnheiten, Fortschritte und Schwierigkeiten von Lernenden bieten. Die Technologien in diesem Cluster basieren daher zumeist auf neuronalen KI-Verfahren. Die KI-gestützte Lerndatenanalyse ist eine wichtige Voraussetzung für eine adaptive, individuelle Lernunterstützung, wie sie in Intelligenten Tutoring-Systeme (siehe Technologiegruppe d) bereitgestellt wird. Informationen zu Lernverlauf und Lernstand können aber auch als Input für Systeme zur Unterstützung der Unterrichtsabläufe (siehe Technologiegruppe c) genutzt werden. So können Lehrkräfte durch diese bei der Diagnostik, beim Bereitstellen von Lernressourcen und bei der Unterrichtsorganisation unterstützt werden.

Mit Blick auf das Desideratum der Ableitung und Umsetzung adaptiver Unterrichtshandlungen werden nachfolgend in **Cluster 2** Systeme fokussiert, durch die Schüler:innen **systembasierte adaptive Lernunterstützung** erhalten. Hiermit sind einerseits Anwendungen angesprochen, die ein gewisses Maß an Adaptivität ermöglichen, aber noch keine KI nutzen. Schwerpunktmäßig werden in diesem Cluster Intelligente Tutoring-Systeme betrachtet (siehe Technologiegruppe d). Intelligente Tutoring-Systeme nutzen symbolische KI, um eine wissensbasierte Anpassung der Lernunterstützung zu erreichen und Lernprozesse effektiv und effizient zu unterstützen. Diese Technologiegruppe ist von zentraler Bedeutung für dieses Gutachten, da die bedarfsangepasste, adaptive Unterstützung auf Stärken und Schwächen der Schüler:innen eingeht und sich daher besonders für Lernende mit Förderbedarfen eignet.

Zuletzt werden in **Cluster 3** Systeme zusammengefasst, die neuronale und speziell generative KI nutzen, um eine **Anpassung der Lernumgebung an individuelle Lernbedarfe** zu unterstützen. Das Cluster fokussiert auf Anwendungen aus den Technologiegruppen f und g. Textgenerierende und -übersetzende sowie Material- und Bildgenerierende Systeme können sowohl von Schüler:innen, als auch von Lehrkräften genutzt werden. Sie ermöglichen es, Materialien differenziert und angepasst an

verschiedene Lernausgangslagen zu gestalten und so individuelle Lernunterstützung einfacher umzusetzen. textgenerierende Systeme können auch genutzt werden, um Schüler:innen individuell Feedback zu geben oder als „sokratischer Dialogpartner" bei der interaktiven Erarbeitung von Wissen zu unterstützen. In das Cluster fallen auch assistive Technologien (siehe Technologiegruppe e), die Benachteiligungen einzelner Lernender ausgleichen können, um gleichberechtigten Zugang zu Lernformen und -systemen zu ermöglichen. Text-to-Speech- und Speech-to-Text-Systeme können insbesondere für Lernende mit entsprechenden Seh- oder Höreinschränkungen als Assistenzsysteme beim Lernen dienen. Sie können so auch die Bedienung und Benutzung von Intelligenten Tutoring-Systemen unterstützen.

5. Analyse des Potenzials von KI-Technologien für individuelle Förderung

Die drei Cluster werden im Folgenden im Hinblick auf ihr Potenzial für eine bedarfsgerechte Lernunterstützung und individuelle Förderung näher beschrieben und beurteilt. Dabei werden exemplarische Anwendungen näher beleuchtet und dahingehend untersucht, inwiefern sie bereits KI-Unterstützung nutzen. Die Auswahl und Betrachtung der Beispiele orientiert sich an folgenden Leitfragen: Welche Systeme gibt es in Deutschland bzw. im deutschsprachigen Raum? Werden sie bereits in der Schulpraxis eingesetzt? Gibt es international besonders herausragende Beispiele, aus denen wir Implikationen für die deutsche Schullandschaft ableiten können? Anschließend wird für jedes Cluster anhand der bislang vorliegenden wissenschaftlichen Evidenz aufgezeigt, in welchem Ausmaß die Systeme das Lernen fördern und Lehrende unterstützen können. Bei der Potenzialanalyse wird ein besonderes Augenmerk auf die individuelle Förderung gelegt.

5.1. Cluster 1: Systeme zur Lerndatenanalyse

Die in diesem Cluster betrachteten Systeme zur Lerndatenanalyse eint, dass sie Lernprozesse oder Lernstände erfassen und analysieren. KI-gestützte Systeme zur Lerndatenanalyse nutzen sogenannte *Learning Analytics*. Diese Methoden werden wie folgt beschrieben: „Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs."[5] (1st International Conference on Learning Analytics and Knowledge, zitiert nach Siemens & Long, 2011, S.34). Für die Lerndatenanalyse mittels *Learning Analytics* können verschiedene Datenquellen herangezogen werden wie zum Beispiel Verhaltens- und Logdaten aus der Interaktion mit einer digitalen Lernumgebung, Selbstauskunftsdaten oder Beiträge in Diskussionsforen (Börte et al., 2023). Durch die Analyse dieser Daten ermöglichen es die entsprechenden Systeme, Förderbedarfe frühzeitig zu diagnostizieren und Lernverläufe, auch über längere Zeiträume, individuell zu verfolgen. Die Systeme können Lehrkräften

[5] „*Learning Analytics* ist die Messung, Sammlung, Analyse und Berichterstattung von Daten über Lernende und ihren Kontext, um das Lernen zu verstehen und zu verbessern" (Übersetzung der Autor:innen)

helfen, selbst oder automatisiert datenbasierte Entscheidungen über passende Fördermaßnahmen zu treffen. Insbesondere in multiprofessionellen Teams, wie sie in inklusiven Lehr-Lernkontexten häufig vorkommen, kann so die Lernbegleitung kooperativ und effektiv gestaltet werden.

Eine wichtige Voraussetzung für die Nutzung der systemseitig analysierten Daten ist die Visualisierung der Analyseergebnisse. Sogenannte *Dashboards* stellen detaillierte Informationen über das Lernen der Schüler:innen bereit, etwa über ihre Leistungen, den Lernfortschritt oder ihr Verhalten innerhalb der Lernumgebung (Börte et al., 2023). Die visuelle Rückmeldung kann sowohl den Schüler:innen, als auch den Lehrkräften zur Verfügung stehen. Schüler:innen können Informationen über gelöste Aufgaben einsehen und so gezielt weitere Übungsgelegenheiten suchen. Lehrkräfte werden mittels eines *Dashboards* über den Lernstand einzelner Lernenden, von Lerngruppen oder ganzer Klassen informiert und erhalten so in Echtzeit Informationen zum Leistungsniveau und Bearbeitungsstand der Aufgaben. Auf diese Weise können Lehrkräfte eine bessere Vorstellung davon erhalten, wem sie helfen sollten. Insbesondere in inklusiven Lehr-Lernsettings scheint es sinnvoll, diese Art des formativen Assessments in Echtzeit zu ermöglichen. Lehrkräfte können so dabei unterstützt werden, den Überblick zu behalten und ggf. gezielt pädagogisch-didaktisch zu intervenieren, indem sie situative Anpassungen im Hinblick auf die Bedürfnisse der Schüler:innen vornehmen.

Exemplarische Anwendungen

Eine computergestützte Anwendung, die in Deutschland zur Diagnostik von Leseverständnisleistung in der ersten bis zum Beginn der siebten Klasse eingesetzt werden kann, ist der Lesetest *ELFE II*[6], der unter anderem bei der *IGLU-Studie* 2017 eingesetzt wurde (Lenhard et al., 2018). Dieser Test erfasst sowohl das Leseverständnis, als auch die Leseflüssigkeit. Die Software nutzt keine KI, sondern spezifische Untertests zum Wort-, Satz- und Textverständnis sowie psychometrische Verfahren, um individuelle Stärken und Schwächen zu identifizieren. Zudem findet eine Analyse der Anzahl an richtig gelösten und insgesamt bearbeiteten Aufgaben, der Genauigkeit von Antworten und der Arbeitsgeschwindigkeit statt. Für die Auswertung steht eine repräsentative Norm wie auch eine adaptierte Norm für Kinder mit Migrationshintergrund zur Verfügung. *ELFE II* kann auch von Lehrkräften eingesetzt werden, z.B. bei Verdacht auf eine Lese-Rechtschreibstörung.

Zum Schuljahresbeginn 2024/25 wurde den Schulen in Nordrhein-Westfalen ein neues computergestütztes Testverfahren kostenfrei zur Verfügung gestellt, mit dem die individuellen Lernvoraussetzungen von Kindern bei der Grundschulanmeldung erhoben werden. Dafür wird im Herbst 2024 das Verfahren *ILEA-T*[7] (**I**ndividuelle **L**ern-**E**ntwicklungs-**A**nalyse in der **T**ransition) genutzt (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW], 2024c). *ILEA-T* ist ein breit angelegtes Erhebungsinstrument, das die Erfassung und Dokumentation der bio-psycho-sozialen

[6] https://www.psychometrica.de/elfe2.html

[7] https://ilea-t.reha.uni-halle.de/

Gesamtsituation und der individuellen Lernvoraussetzungen von Kindern im Bereich schriftsprachlicher und mathematischer Grundkompetenzen ermöglicht. Es wurde von der Universität Halle-Wittenberg entwickelt und mittels einer Validierungsstudie wissenschaftlich erprobt (Berger, 2013). *ILEA-T* ist auf inklusive Settings ausgerichtet und zielt darauf ab, einen inklusionsorientierten Übergang in die Grundschule zu ermöglichen. Dies geschieht durch die Bereitstellung differenzierter, domänenspezifischer Angebote, die auf der vorhergehenden Diagnostik basieren (Geiling et al., 2012). Als digitales Screening-Instrument wird *ILEA-T* zunächst an 130 Grundschulen erprobt und soll bei Erfolg ab Herbst 2025 allen Grundschulen in Nordrhein-Westfalen zur Verfügung stehen (MSB NRW, 2024 c).

Eine weitere Anwendung für die Diagnostik von Lesekompetenzen ist der Lautlesetutor (*LaLeTu*)[8] von Klett. Dabei ergänzen sich der *Lautlesetutor Diagnostik* und der *Lautlesetutor Lesen* gegenseitig. Der Lesetutor nutzt eine sprachbasierte KI-Komponente. Die Leseleistung wird nach Kriterien wie Leseflüssigkeit, Lesetempo, Satzbetonung oder Lesefehler analysiert und die Lehrkraft erhält anschließend eine Auswertung der Diagnose- und Förderaufgaben für einzelne Lernende oder für die ganze Lerngruppe. Es gibt drei Teiltests, um den Fortschritt zu drei Zeitpunkten über das Schuljahr hinweg zu messen. Angaben bzw. Nutzerzahlen zum Einsatz der Tutoren im Unterricht sind nicht verfügbar.

Einen Analyseansatz, bei dem unterschiedliche Datenquellen miteinander kombiniert und abgeglichen werden, stellt das online zur Verfügung stehende Tool *CleverR*[9] dar. Dieses Tool kann zur Diagnose und Förderung des selbstregulierten Lernens von Schüler:innen eingesetzt werden. Schüler:innen schätzen zur Diagnose des Ist-Zustands zunächst selbst ihre Stärken und Schwächen in Bezug auf acht relevante Aspekte des selbstregulierten Lernens ein. Dieses Wissen sowie die Nutzung von Selbstregulations- und Lernstrategien werden mittels verschiedener Aufgaben in einem Wissenstest erfasst und durch Selbstberichte zur Strategienutzung ergänzt. Diese Informationen und Ergebnisse werden dann miteinander abgeglichen. So entstehen individuelle Kompetenzprofile, die es Schüler:innen und Lehrkräften ermöglichen, datengestützt zu überprüfen, inwiefern ihre Einschätzungen akkurat sind und welche Förderbereiche bestehen. So können entsprechende Maßnahmen geplant werden. Um die Entwicklung der Kompetenzen zu verfolgen, wird ein wiederholtes Ausfüllen nach drei Monaten empfohlen. Das Tool wurde bislang pilotiert, im Rahmen einer groß angelegten Hauptstudie in Deutschland und der Schweiz erprobt sowie anschließend für die Nutzung in der Sekundarstufe I zur Verfügung gestellt (Bäuerlein et al., 2022).

Eine Online-Plattform zur Diagnostik und Förderung von Kindern mit Lernstörungen ist *LONDI*[10]. Die Plattform bietet unter anderem ein Screening Instrument sowie ein Hilfesystem, welches weitere Empfehlungen hinsichtlich Testverfahren für vertiefte

[8] https://www.laletu.de/home

[9] https://www.clever-selbstreguliert-lernen.ch/index

[10] https://www.londi.de/

Diagnostik bietet. Mit der Screening-Anwendung werden mittels bewährter Tests und wissenschaftlich normierter Testverfahren die Fertigkeiten in den Bereichen Lesen, Schreiben und Rechnen eingeschätzt. Die ermittelten Ergebnisse werden von der Lehrkraft in das bereitgestellte Hilfesystem eingegeben, wodurch eine auf den einzelnen Lernenden abgestimmte Diagnostikempfehlung erstellt wird. Nach Durchführung der Diagnostik können die Ergebnisse erneut in das Hilfesystem eingefügt werden. Daraufhin wird automatisch eine Auswahl zugeschnittener Förderempfehlungen generiert und ausgegeben.

Die bisher genannten Beispiele nutzen zwar computergestützte Testverfahren und erlauben eine differenzierte Diagnostik von Lernständen und ggf. Förderbedarfen. Sie nutzen hierzu jedoch keine KI-gestützten Verfahren auf der Basis von *Learning Analytics*. Dadurch ist die Möglichkeit einer auf die Diagnostik aufbauenden individuellen Unterstützung durch ein adaptives System sehr eingeschränkt. Es geht hier eher darum, Empfehlungen zu geben, die Lehrkräfte zur Anpassung der Lern- und Förderangebote nutzen können.

Ein einschlägiges Beispiel aus den USA für ein System, das KI-Methoden einsetzt, ist *ASSISTments*[11]. Es handelt sich hierbei um eine technologiebasierte, formative Bewertungsplattform, die drei Produkte umfasst: *ASSISTments Teacher*, *ASSISTments Advantage* und *ASSISTments Tutor*. Das System erfasst während der Aufgabenbearbeitung Lerndaten zu Aufgaben, die den Schüler:innen zuvor von der Lehrkraft zugewiesen wurden. Lehrkräfte können am Lehrplan orientiert Materialien nutzen, welche auf der Plattform angeboten werden, oder eigene Aufgaben erstellen und diese ihren Schüler:innen zuweisen. Basierend auf einer automatisch erfolgenden Lerndatenanalyse gibt das System den Schüler:innen Schritt-für-Schritt-Feedback (*step-loop adaptivity*, s.o.). Den Lehrkräften stehen die analysierten Daten des formativen Assessments des Lernprozesses und der Leistungen auf Lernendenebene zur Verfügung (Heffernan & Heffernan, 2014). Dank der lehrplannahen Inhalte wird *ASSISTments* vielfach genutzt, wodurch umfangreiche Lerndaten auch über einen längeren Zeitraum hinweg erfasst und analysiert werden können (Heffernan & Heffernan, 2014). Dies erlaubt auf einer detaillierten Ebene die Verfolgung des Lernstands der Schüler:innen.

Ein weiteres Beispiel für ein KI-gestütztes System zur Lerndatenanalyse ist die Plattform *ViLLE*[12], betrieben vom Zentrum für Learning Analytics der Universität von Turku, Finnland. Diese Plattform bietet den datenschutzkonformen Einsatz von *Learning Analytics*. International wird die Plattform unter dem Namen *Eduten Playground*[13] angeboten. *ViLLE* bietet unter anderem automatisch bewertete Übungen, Prüfungen, aber auch die Möglichkeit, mit anderen zu zweit oder als Lerngruppe zusammenzuarbeiten. Lehrkräfte können dabei den Lernfortschritt verfolgen. Das

[11] https://new.assistments.org/
[12] https://en.learninganalytics.fi/ville
[13] https://eduten.com/

Zentrum für Learning Analytics entwickelt dabei nicht nur das System weiter. Vielmehr erforscht das Zentrum auch umfassend die Lehr- und Lernanalytik sowie die gesellschaftlichen Auswirkungen des Einsatzes digitaler Tools und Analysen. Diese breite Ausrichtung ist beispielhaft für die Entwicklung von *Learning Analytics*, insbesondere, wenn es sich um Daten inklusiver Lehr-Lernkontexte handelt.

Empirische Befunde

Das Feld der *Learning Analytics* ist noch jung. Daher liegen bislang nur wenige praktische Erfahrungen in Lehr- und Lernkontexten vor (Peters & Bovenschulte, 2020). Zudem bestehen insbesondere im schulischen Kontext ethische und datenschutzrechtliche Herausforderungen, die es zu berücksichtigen gilt und die den Einsatz solcher Systeme erschweren.

Wie der Einsatz gefördert und eine weite Verbreitung einer solchen Plattform erreicht werden kann, untersucht ein auf den bisherigen Entwicklungsarbeiten zur Plattform *LONDI*[14] aufbauendes Nachfolgeprojekt (*Londi II*[15]). Dabei steht die effektive und nachhaltig wirksame Nutzung durch verschiedene Nutzergruppen im Vordergrund. Dass zusätzlich zu dem Entwicklungsprojekt *LONDI* ein zweites Projekt durchgeführt wird, in dem die Nutzung und Wirksamkeit erforscht wird, zeigt, wie aufwändig und zeitintensiv die Entwicklung von effektiven und verlässlichen Systemen zur Lerndatenanalyse ist. Die Ergebnisse des Projekts liegen bisher nicht öffentlich vor. Im Zusammenhang mit dem Einsatz der Plattform *quop*[16] untersuchten Schmitterer und Kolleg:innen (2023) die Effektivität eines Systems zur Lernverlaufsdiagnostik des Leseverständnisses. Sie konnten zeigen, dass besonders Schüler:innen mit einem anfangs geringen Leseverständnis von der Lernverlaufsdiagnostik profitierten.

In Bezug auf die Plattform *ASSISTments*[17] konnten Heffernan und Heffernan (2014) feststellen, dass eigene Gestaltungsmöglichkeiten für Lehrkräfte ein wichtiger Faktor sind, der die Nutzung der Plattform positiv beeinflusst. Die Plattform bietet Lehrkräften und Forschenden ein Autorentool („*authoring tool*") mit dessen Hilfe sie Inhalte erstellen können. Forschende können damit die Plattform auch nutzen, um kontrollierte Studien durchzuführen.

Insgesamt liegen bislang nur wenig Befunde dazu vor, ob Systeme zur Lerndatenanalyse Lernende individuell fördern können. Eine Ausnahme ist die Forschung zur Plattform *ASSISTments*. Der Einsatz von *ASSISTments* wurde mehrfach hinsichtlich seiner Effektivität untersucht und eine Lernwirksamkeit konnte nachgewiesen werden, insbesondere für Schüler:innen mit Förderbedarf (Koedinger et al., 2010). Dabei weisen Heffernan & Heffernan (2014) darauf hin, dass nicht die Plattform selbst zu einem besseren Lernen führt, sondern die lernförderliche Wirkung stark von den verwendeten

[14] https://www.londi.de/

[15] https://www.dipf.de/de/forschung/projekte/londi-2-implementation-und-evaluation-der-londi-online-plattform-zur-diagnostik-und-foerderung-bei-lernstoerungen#0

[16] https://www.quop.de/de/start/

[17] https://new.assistments.org/

Inhalten sowie der Reihenfolge der Präsentation derselben abhängig ist. Dennoch konnten Koedinger und Kolleg:innen (2010) zeigen, dass schon allein der Einsatz von *ASSISTments* lernwirksam ist. Schüler:innen in der Gruppe der Lehrkräfte, die *ASSISTments* einsetzten, schnitten auch dann besser ab, wenn sie selbst das System nur wenig oder gar nicht nutzten. Dies deutet darauf hin, dass die Lehrkräfte den Unterricht für die gesamte Klasse auf Grundlage der zur Verfügung stehenden Daten angepasst haben und die Schüler:innen dadurch lernwirksam profitieren konnten. Die Studienergebnisse von Feng und Kolleg:innen (2023) deuten ebenfalls darauf hin.

Aslan und Kolleg:innen (2019) konnten zeigen, dass der Einsatz eines *Dashboards*, welches Echtzeit-Informationen über das Engagement der Schüler:innen anzeigt, einen signifikanten Einfluss auf die Unterrichtspraktiken der Lehrkräfte sowie das Engagement der Schüler:innen hatte. Letztere scheinen durch die dynamische Bewertung des individuellen Lernfortschritts lernförderlich zu profitieren (Tetzlaff et al., 2021). Die Meta-Analyse von Jung und Kolleg:innen (2018) zeigt, dass dies umso mehr für Schüler:innen mit besonderen Lernbedürfnissen gilt.

Andererseits fanden Molenaar & Knoop-van Campen (2016) aber, dass Lehrkräfte bei der Nutzung von *Learning Analytics* unterstützt werden müssen und dass es weiterer Forschung bedarf, um die menschlichen Stärken und die Vorteile der Systeme zur Lerndatenanalyse zu verbinden. Die Nutzung von *Learning Analytics* setzt eine visuell schnell zu verstehende und einfach zu interpretierende Darstellung der Analysen voraus. Ein futuristisch anmutendes Forschungsprojekt aus den USA zeigt eine intuitiv zugängliche Darstellung von Lerndatenanalysen. In *Lumilo* tragen die Lehrkräfte ein Smart-Glasses-System wie eine Brille (Holstein, 2018). Lumilo ermöglicht durch transparente Gläser den Klassenraum und die Schüler:innen im Blick zu haben. Zusätzlich werden in der Brillenansicht Informationen über das Lernverhalten (*Learning Analytics*) der Schüler:innen durch Symbole über den Köpfen der Schüler:innen repräsentiert. So können die Lehrkräfte den Lernfortschritt der Schüler:innen verfolgen und ggf. eingreifen. In der Studie von Holstein (2018) wurde untersucht, ob Schüler:innen besser lernen, wenn Lehrkräfte nicht nur den Lernfortschritt über *Lumilo* verfolgen, sondern auch Zugang zu *Learning Analytics* haben. Die Ergebnisse zeigen, dass der Lernzuwachs der Schüler:innen in der Gruppe der Lehrkräfte, die über *Lumilo* zusätzlich Zugang zu *Learning Analytics* hatten, am höchsten war.

Einordnung

Durch *Learning Analytics* können aus großen Mengen an Lerndaten wichtige Erkenntnisse gewonnen werden, die Lehrkräften sonst möglicherweise entgehen würden. Durch die KI-gestützte Datenanalyse werden üblicherweise nicht zugängliche Informationen sichtbar, z.B. über den Lernfortschritt oder Problemfelder, die ohne diese Technik unbemerkt bleiben könnten. Subjektive Eindrücke von Lehrkräften können so datenbasiert, gestärkt oder sogar korrigiert werden. Allerdings bestehen ethische Bedenken, und es sind noch viele Fragen ungeklärt, etwa welche Daten zu welchem Zweck erfasst werden, wer Zugang zu diesen Daten hat und welche Regulierungen durch die zuständigen Instanzen notwendig sind. Daher gibt es bislang auch in Deutschland

keine implementierten, breit eingesetzten Systeme, die langfristig Daten erfassen und KI-gestützt analysieren.

Durch *Learning Analytics* können nicht nur Lernstände diagnostiziert, sondern diese Erkenntnisse auch in nachfolgend sinnvolle adaptive Hilfestellungen übertragen werden (z.B. Lawrence et al., 2022; Yang et al., 2022). Eine Besonderheit in vielen Systemen dieses Clusters ist die Verschränkung mehrerer Anwendungen. So werden in mehreren der genannten Beispiele zwar primär Anwendungen genutzt, die den Lernstand diagnostizieren, darüber hinaus aber weitere, die zur gezielten Förderung eingesetzt werden können. Hinzu kommen Komponenten, die es Lehrkräften ermöglichen, über *Dashboards* auf Lerndaten zuzugreifen und diese für die Zusammenstellung von Lerngruppen oder für einen Wechsel zwischen den Unterrichtsphasen zu nutzen. So bietet die *ASSISTments* Plattform drei Produkte, die einzeln oder miteinander kombiniert eingesetzt werden können.

Learning Analytics können Lehrkräfte zwar unterstützen, indem Hilfebedarfe während der Unterrichtssituationen erkannt werden. Es bedarf jedoch einer sorgsamen Aufbereitung der analysierten Daten, damit sie für Lernende und Lehrkräfte verständlich und interpretierbar sind sowie das selbstregulierte Lernen und Reflexionsprozesse unterstützen können (Holstein, 2018; Molenaar & Knoop-van Campen, 2016; DFKI, 2023). Lehrkräfte brauchen technologische sowie diagnostische Kompetenzen, um die Ergebnisse zu hinterfragen und ggf. auch kritisch einordnen zu können (DFKI, 2023).

Ein vielversprechender neuer Ansatz für das Lernen in inklusiven Lernsettings in den USA ist das sogenannte *Human-AI-Tutoring*. Dabei werden Informationen auf einem *Dashboard* bereitgestellt, um Lehrkräfte und Pädagogen zu unterstützen, indem sie gezielt entscheiden können, welche Schüler:innen Hilfe benötigen. Das Ziel ist es, die Leistungsfähigkeit bestehender KI-Systeme zu nutzen, um den Lehrkräften und Pädagog:innen, die in der sonderpädagogischen Förderung tätig sind, mehr Freiraum zu geben, sich auf den Beziehungsaufbau und die sozial-emotionale Unterstützung der Schüler:innen zu konzentrieren. Erste Studienergebnisse deuten darauf hin, dass die Kombination von menschlichem Tutoring und KI-Unterstützung positive Auswirkungen sowohl auf den Lernprozess, als auch auf die Lernergebnisse insbesondere von Schüler:innen mit Förderbedarf hat (Thomas et al., 2024).

In Deutschland kommen KI-gestützte Systeme zur Lerndatenanalyse bislang kaum zum Einsatz, daher sind die mit diesem Cluster verbundenen Potenziale eher theoretischer Natur. Es bedarf übergeordneter Regularien zum Einsatz solcher Systeme sowie datenschutzkonformer Lösungen, um weitergehende Evidenz zu sammeln.

5.2. Cluster 2: Systeme zur adaptiven Lernunterstützung

An individuelle Bedarfe angepasste Lernunterstützung zu ermöglichen, ist ein zentrales Ziel moderner digitaler Lehr-Lernsysteme. „Die Frage nach der Adaptivität multimedialer Lehr- und Informationssysteme bezieht sich [...] auf die Frage, inwieweit ein System selbst in der Lage ist, den Unterstützungsbedarf der Lernenden zu diagnostizieren und

das Ergebnis der Diagnose in geeignete angepasste Lehrtätigkeiten umzusetzen“, wie Leutner (2002) treffend zusammenfasst (S. 118).

KI-gestützt wird adaptive Lernunterstützung im Rahmen sogenannter Intelligenter Tutoring-Systeme bereitgestellt. Unter Intelligenten Tutoring-Systemen versteht man Softwareprogramme, die eine Vielzahl an Aufgaben bestimmter (fach-)spezifischer Wissensdomänen umfassen und diese in Form von kleinschrittigen Wissenskomponenten modellieren. Intelligente Tutoring-Systeme ähneln in ihrer Funktionsweise der 1:1 Lernbegleitung durch eine:n menschliche:n Tutor:in. Diese Systeme helfen Schüler:innen, das Lösen bestimmter Aufgaben zu üben und die zugrundeliegenden Wissenskomponenten zu erwerben. Typischerweise handelt es sich um Aufgaben mit mehreren Lösungsschritten und mehreren potenziell möglichen, aber verschiedenen Lösungswegen. Intelligente Tutoring-Systeme bieten adaptive, individualisierte Schritt-für-Schritt-Unterstützung (*step-loop adaptivity*, s.o.) innerhalb der Aufgabenbearbeitungen. Es wird also nicht nur am Ende einer Aufgabe Feedback zur Lösung gegeben, sondern die Schüler:innen erhalten zusätzlich Hinweise während der Aufgabenbearbeitung. Die Hinweise orientieren sich an den Lösungsstrategien und Fehlern der Schüler:innen. Die Hilfestellungen werden daher mit zunehmender Kompetenz weniger. Das System gibt sowohl automatische Hilfestellungen, die zeigen, wie eine Aufgabe korrekt zu lösen ist oder wie Fehler behoben werden können; die Schüler:innen können die Hilfe aber auch selbst anfragen. Daneben bieten Intelligente Tutoring-Systeme eine adaptive Aufgabenauswahl (*task-loop adaptivity*, s.o.). Die Aufgabenauswahl erfolgt dabei nicht anhand einer vorab ermittelten Einschätzung der Lernenden, sondern anhand von lernbegleitenden Messungen des Wissenszuwachs der einzelnen Schüler:innen. Diese Messungen erfolgen kontinuierlich auf fein-granularer Ebene durch das Intelligente Tutoring-System, während die Schüler:innen in dem System Aufgaben bearbeiten. Aufgrund der automatischen Datenanalyse wird die Aufgabenschwierigkeit angepasst. Durch das hohe Adaptionsvermögen ergeben sich gute Möglichkeiten, insbesondere auf Leistungsunterschiede sowie besonderen Förderbedarf in einer heterogenen Schulklasse einzugehen.

Intelligente Tutoring-Systeme nutzen üblicherweise symbolische KI. Sie arbeiten auf der Grundlage verschiedener, meist durch Regeln miteinander verknüpfter Wissensmodelle: einer umfassenden Modellierung des jeweiligen Inhaltswissens in einem Domänenmodell und einem Lernermodell, das auf den Lösungen der Aufgaben und gegebenenfalls Fehleranalysen jedes Lernenden basiert. Manche Intelligente Tutoring-Systeme haben zudem ein pädagogisches Modell (Tutormodell), das die Art der Rückmeldung und adaptiven instruktionalen Unterstützung bestimmt. In manchen Intelligenten Tutoring-Systemen, insbesondere im Bereich des Sprachenlernens, wird das wissensbasierte Modellieren mit statistischen Modellen kombiniert (Leifheit et al., 2024).

Exemplarische Anwendungen

Im Bereich des digitalen inklusiven Lernens wurde in den letzten Jahren in Deutschland bereits ein umfassendes Angebot an digitalen Materialien, Lernsoftware und interaktiven Lernprogrammen entwickelt, die den Anspruch haben, die Lernenden adaptiv und

individuell zu fördern. Der Bundesverband für Legasthenie und Dyskalkulie e.V. hat dem Fokus des Verbands entsprechend in einem online zur Verfügung gestellten Dokument Lernsoftware zusammengestellt, die sich zur ergänzenden Förderung bei Rechenschwäche und Lese-Rechtschreibschwäche bewährt haben[18]. Die dort aufgeführten Programme dienen in der Regel der Förderung von Basiskompetenzen in den unteren Klassenstufen, können jedoch bei älteren Kindern sowie höheren Klassenstufen auch bei entsprechenden Förderbedarfen eingesetzt werden. Zur Förderung mathematischer Konzepte und Regeln für die erste bis dritte Grundschulklasse werden beispielsweise die *Rechenspiele mit Elfe und Mathis I* genutzt, welche sich an den nationalen Bildungsstandards der Kultusministerkonferenz orientieren und lernpsychologischen Prinzipien folgen. Für die erfolgreiche Bewältigung eines Spiels muss eine festgelegte Lösungsquote erreicht werden, die individuell verändert werden kann (Lenhard et al., 2011). Eine Software, die von Kognitions- und Entwicklungspsycholog:innen ausdrücklich als computerbasiertes Trainingsprogramm für Kinder mit Lese-Rechtschreibschwäche konzipiert und entwickelt wurde, ist *Lautarium* (Konerding et al., 2020). Es basiert auf dem internationalen Forschungsstand zu Ursachen sowie Interventions- und Präventionsmöglichkeiten bei Lese-Rechtschreibschwäche. Das Programm kann als Ergänzung zu einer professionellen Legasthenie-Therapie bei Dritt- und Viertklässlern sowie als Präventionsprogramm bei Erst- und Zweitklässlern mit Problemen beim Schriftspracherwerb eingesetzt werden. Die Instruktionen sind interaktiv in den Programmablauf integriert (Konerding et al., 2020). Die hier genannten Lernprogramme erlauben eine gewisse Anpassung, erreichen allerdings nicht den Grad an Adaptivität, die Intelligente Tutoring-Systeme aufgrund der zugrundeliegenden Wissensmodelle ermöglichen können.

In Deutschland gibt es bislang nur wenige Intelligente Tutoring-Systeme, die in Schulen eingesetzt werden. Im Bereich Mathematik wird das intelligente Tutoring-System *Bettermarks* in den Jahrgängen vier bis einschließlich der Oberstufe eingesetzt. Es ist „das führende Intelligente Tutorielle System für Mathematik in Deutschland"[19] und durch Landeslizenzen mit verschiedenen Bundesländern an deutschen Schulen relativ verbreitet (DFKI, 2023). In Bettermarks können Schüler:innen mit verschiedenen Eingabewerkzeugen mathematische Inhalte üben. Bei der Aufgabenbearbeitung erhalten sie einerseits formative Schritt-für-Schritt-Unterstützung (*step-loop adaptivity*, s.o.) beim Bearbeiten von Aufgaben, als auch eine adaptiv auf den Lernstand zugeschnittene Aufgabenauswahl (*task-loop adaptivity*, s.o.). Dem System liegen differenzierte Wissensmodelle zugrunde. Zurzeit werden die Unterstützungsmöglichkeiten, die das System im Fach Mathematik bietet, durch die Integration sprachgenerativer KI-Komponenten erweitert und verfeinert. Zudem ist eine Erweiterung auf andere Fächer geplant (C. Speroni, persönliche Kommunikation, 27. September 2024).

[18] http://www.alpha-fundsachen.de/wp-content/uploads/Lernsoftware_Dyskalkulie_und_Legasthenie.pdf
[19] https://de.bettermarks.com/

Im Bereich des Fremd- und Sprachenlernens wird das Programm *FeedBook*[20] (Meurers et al., 2019) eingesetzt. Mit *FeedBook* kann Englisch ab der siebten Klasse gelernt werden, wobei das System über den Browser zugänglich ist. Es kombiniert einen regelbasierten mit einem datengetriebenen Ansatz und ist eins der wenigen Intelligenten Tutoring-Systeme zum Einsatz im schulischen Fremdsprachenunterricht in Deutschland (Meurers et al., 2018). Im Englischunterricht eingesetzt, bietet es Schüler:innen schrittweises Feedback, indem es Tipps zum eigenständigen Finden der korrekten Lösung gibt.

Zwei in der Schweiz entwickelte und für Schulen in Deutschland sprachlich angepasste Systeme sind *Dybuster Calcularis*[21] und *Grafari*[22]. Das Intelligente Tutoring-System *Calcularis*[23] wurde ursprünglich für Kinder mit Rechenschwäche entwickelt, kann aber in den unteren Klassenstufen zur Festigung der Basiskompetenzen eingesetzt werden (Käser et al., 2012; 2013). Bei der Entwicklung haben verschiedene Fachdisziplinen zusammengearbeitet, um gemeinsam und unter Berücksichtigung der jeweiligen Expertise an der Entwicklung der Software mitzuwirken. So haben beispielsweise Neuropsycholog:innen mit entsprechender Expertise bezüglich Dyskalkulie und Informatiker:innen mit entsprechender Expertise im maschinellen Lernen an der Entwicklung und Evaluation zusammengearbeitet. Die Software kann als Ergänzung zu einer Therapie zur Förderung der Grundkompetenzen und Grundrechenarten eingesetzt werden. *Dybuster Grafari*[24] ist als Software-Programm für den Lese- und Rechtschreiberwerb ähnlich ausgerichtet, indem es bei jüngeren Schüler:innen den Erwerb und den Ausbau der Basiskompetenzen fördert und bei älteren Lernenden auch für die Förderung bei entsprechenden Schwächen eingesetzt werden kann (Kast et al., 2011).

Im angloamerikanischen Raum haben verschiedene Intelligente Tutoring-Systeme den Schritt in die Schulpraxis geschafft, beispielsweise in den USA die Systeme *MATHia*[25] und ALEKS[26]. *MATHia* ist eine von der Firma Carnegie Learning vertriebene Lernsoftware, die vor dem Hintergrund jahrzehntelanger Forschung zu Intelligenten Tutoring-Systemen an der Carnegie Mellon University (USA) entstanden ist. Die sogenannten Kognitiven Tutoren[27], zu denen auch Intelligente Tutoring-Systeme zählen, wurden für viele Bereiche entwickelt, wobei *MATHia* speziell im Mathematikunterricht eingesetzt werden kann. Erkenntnisse aus Psychologie und Kognitionswissenschaft sowie aus der Informatik wurden genutzt, um ein Intelligentes Tutoring-System zu entwickeln, welches sich adaptiv an die Lernfortschritte und Bedürfnisse der Schüler:innen anpasst. Durch den Abgleich mit Wissensmodellen erkennt *MATHia* Fehler und Misskonzepte der Schüler:innen und unterstützt diese durch Echtzeit-Feedback und

[20] http://feedbook.schule
[21] https://constructor.tech/solutions/schools
[22] https://constructor.tech/solutions/schools
[23] https://constructor.tech/de/products/learning/calcularis
[24] https://constructor.tech/products/learning/grafari-o
[25] https://www.carnegielearning.com/solutions/math/mathia/
[26] https://www.aleks.com/?_s=3279768772623512
[27] Kognitive Tutoren sind nicht bereichsspezifisch und wurden für viele Fächer entwickelt. Die Technologie bezieht die kognitive Theorie des menschlichen Problemlösens mit ein (Anderson et al., 1995).

adaptive Aufgabenauswahl. Lehrkräfte können das Lernen ihrer Schüler:innen verfolgen, indem die Software entsprechende Daten über Reports bereitstellt. Die Inhalte sind curriculumsbezogen und decken mittlerweile die Klassenstufen 6-12 des amerikanischen Schulsystems ab. So lernen etwa 600.000 Schüler:innen jedes Jahr mit der Software, die in circa 40% der Unterrichtszeit genutzt wird. Die Nutzung ist curricular eingebettet in das fachdidaktische Unterrichtskonzept.

Empirische Befunde

Da Intelligente Tutoring-Systeme bereits seit mehreren Jahrzehnten entwickelt und beforscht werden, liegen zu ihnen umfangreiche Befunde vor. Diese stammen allerdings hauptsächlich aus dem angloamerikanischen Raum. Dabei zeigt eine Vielzahl an Studien, dass sich besonders die adaptive Anpassung der instruktionalen Unterstützung und der Aufgabenauswahl (*task-loop adaptivity*) an den Wissensstand der Schüler:innen sowie die Anpassung an ihre Strategien und Fehler (*step-loop adaptivity*) als sehr lernförderlich erwiesen haben (Kulik & Fletcher, 2016; Ma et al., 2014; Steenbergen-Hu & Cooper, 2013; 2014; Xu et al., 2019).

In Deutschland werden Intelligente Tutoring-Systemen bislang vergleichsweise selten eingesetzt. Zudem besteht mitunter die Schwierigkeit, adaptive Lernsoftware, die von Unternehmen entwickelt wurde, zu evaluieren und zu beurteilen, da die entsprechenden Daten nur den Unternehmen zugänglich sind. Auf der Webseite von Bettermarks finden sich Hinweise auf eine Lernwirksamkeit[28], zitiert wird dazu eine ältere Studie von Scharnagl und Kolleg:innen (2009). Für die exemplarisch genannte Software *Calcularis* konnten mittels Pilot-, Evaluierungs- und Kontrollstudien nicht nur die Wirksamkeit des Trainingsprogramms untersucht, sondern auch die zugrundeliegenden Lernermodelle sowie die Anpassungsfähigkeit der Lernumgebung bewertet und überarbeitet werden (Käser et al., 2013; Kohn et al., 2020). Diese iterativen Verbesserungsverfahren sind als vorbildlich im Hinblick auf die Softwareentwicklung einzuschätzen.

Für den Einsatz des oben vorgestellten *FeedBook*, das über ein Jahr lang wissenschaftlich begleitet und untersucht wurde, konnte ein positiver Einfluss auf den Lernerfolg nachgewiesen werden. Es wurde an Schulen in Baden-Württemberg eingesetzt, die in Zusammenarbeit mit der Universität Tübingen an der Erprobung teilnahmen (Meurers et al., 2019).

In den USA hingegen werden Intelligente Tutoring-Systeme nicht nur punktuell eingesetzt, sondern auch curricular eingebettet. Dadurch kann der Einsatz auch über einen längeren Zeitraum evaluiert werden. Ein Intelligentes Tutoring-System im Bereich Mathematik untersuchten etwa Pane und Kolleg:innen (2014). In dieser Evaluationsstudie, an der 17.000 Schüler teilnahmen, zeigte sich, dass die Gruppe, die mit dem *Cognitive Tutor Algebra I* lernte, innerhalb eines Jahres im Vergleich zu einer Kontrollgruppe, die nach dem üblichen Lernplan unterrichtet wurde, doppelt so viel lernte. Aufgrund der langjährigen Forschung zu Intelligenten Tutoring-Systemen in den USA

[28] https://de.bettermarks.com/wirkung/

können für die Bewertung der Effektivität auch Meta-Analysen und Meta-Reviews herangezogen werden, die Ergebnisse von Einzelstudien zusammenfassen. Eine Meta-Analyse von 107 Einzelstudien zum Einfluss von Intelligenten Tutoring-Systemen auf die Lernleistung von Schüler:innen zeigte, dass insbesondere im Fach Mathematik positive Effekte durch Intelligente Tutoring-Systeme bestehen (Ma et al., 2014). Weitere Studien, in denen untersucht wurde, ob bestimmte Formen der Adaptivität das Lernen der Schüler:innen verbessern, weisen deutlich darauf hin, dass die Anpassung an das Wissen und den Wissenszuwachs der Schüler:innen sowie die Anpassung an ihre Strategien und Fehler sehr effektiv sind. Schüler:innen lernen effektiver, wenn sie mit einem Intelligenten Tutoring-System üben können (Kulik & Fletcher, 2015; Ma et al., 2014; Steenbergen-Hu & Cooper, 2013; 2014; Xu et al., 2019). VanLehns Meta-Studie (2011) konnte zeigen, dass Intelligente Tutoring-Systeme nahezu so effektiv sind, wie das 1:1 Tutoring durch Menschen. Dabei ist schrittweises Tutoring effektiver als Tutoring, welches nur am Ende der Aufgabenbearbeitung eine Rückmeldung gibt. Diese empirischen Befunde lassen darauf schließen, dass Intelligente Tutoring-Systeme im regulären Unterricht effektiv genutzt werden können, um Lernende individuell zu fördern.

Bislang ist nur wenig untersucht worden, ob die Lernwirksamkeit von Intelligenten Tutoring-Systemen auch für Lernende mit besonderen Förderbedarfen gilt. Ergebnisse deuten darauf hin, dass insbesondere Leistungsschwächere durch den Einsatz eines intelligenten Tutoring-Systems profitieren (Beal et al., 2013; Holstein et al., 2018; Huang et al., 2016). Weitere Arbeiten legen nahe, dass die Kombination einer Lernbegleitung durch die Lehrkraft und der Einsatz von KI vielversprechend ist (u.a. Chine et al., 2022).

Einordnung

Symbolische, wissensbasierte KI, wie sie in Intelligenten Tutoring-Systemen Anwendung findet, ermöglicht ein sehr hohes Maß an adaptiver Unterstützung. Durch die automatische Analyse der Lösungsschritte und darauf aufbauende schrittweise Anleitungen eignen sich diese besonders für den Erwerb grundlegender Kenntnisse und Fertigkeiten wie die der Rechtschreibung oder der mathematischen Grundlagen. In Bezug auf die Unterstützung von Schüler:innen mit spezifischen Förderbedarfen wie Lese-Rechtschreib- und Rechenschwäche eignen sich Intelligente Tutoring-Systeme, da sie darauf abzielen, kleinschrittig modellierte Kompetenzbereiche zu üben und somit längerfristig zu automatisieren (Jenni, 2024). Dabei sollte berücksichtigt werden, dass die Lernziele für Schüler:innen mit Benachteiligungen und Schwächen zu hoch sein können. Daher kann es sinnvoll sein, weitergehend zu untersuchen, ob eine Anpassung bzw. Veränderung der festgelegten zu fördernden Wissenskomponenten für Schüler:innen mit Lese-Rechtschreibschwäche bzw. Rechenschwäche förderlicher ist.

Vor dem Hintergrund der mittlerweile über 40-jährigen Forschung zu Intelligenten Tutoring-Systemen ist festzustellen, dass trotz dieses langen Zeitraums kein flächendeckender Einsatz dieser Systeme erreicht werden konnte. Allerdings gilt für die USA, wo Systeme wie *MATHia* verbreitet in Schulen genutzt werden, dass dort immer mehr Anwendungen, die in Schulen eingesetzt werden, Eigenschaften intelligenter Tutoring-Software aufweisen.

Um intelligente Tutoring-Systeme zu entwickeln, die auf individuelle Förderung und Binnendifferenzierung abzielen, braucht es komplexe Wissensmodelle, die dem System zugrunde liegen. Dies stellt einen nicht zu unterschätzenden Aufwand dar (Meurers et al., 2018). Dieser Aufwand kann allerdings durch die Nutzung von sogenannten *authoring tools* reduziert werden. Damit wird die Erstellung von Lerninhalten für Intelligente Tutoring-Systeme auch für Personen ohne Programmierkenntnisse möglich (Aleven et al., 2016). In Anbetracht der Lernwirksamkeit ist das Potenzial dieser Systeme trotz der Entwicklungskosten als positiv zu bewerten.

5.3. Cluster 3: Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe

Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe, die in diesem Cluster betrachtet werden, generieren, basierend auf Wahrscheinlichkeiten, Inhalte (generative KI) oder wandeln Inhalte um (Text-to-Speech und Speech-to-Text). Sie können sowohl von Schüler:innen, als auch von Lehrkräften genutzt werden. Abgrenzend zu Intelligenten Tutoring-Systemen erfolgt in diesen Systemen der Impuls der Anpassung an individuelle Unterstützungs- und Förderbedarfe durch die Lehrkraft oder die Lernenden selbst und nicht systembasiert. Lehrkräfte können mit diesen Systemen Materialien differenziert und angepasst an verschiedene Lernausgangslagen generieren. Weitere Nutzungsmöglichkeiten bestehen in der Unterstützung der Lehrkräfte bei der Unterrichtsplanung und -gestaltung sowie bei Textüberarbeitung und -übersetzung. Systeme, die generative KI nutzen, können aber auch als Dialogpartner für Schüler:innen eingesetzt werden und diesen individuelles Feedback geben. Assistive Systeme können eingesetzt werden, um Benachteiligung einzelner Lernender auszugleichen und so gleichberechtigten Zugang zu Lernformen und -systemen zu ermöglichen. Dies kann auch bedeuten, dass diese Systeme beispielsweise eingesetzt werden können, um adaptive Lernsysteme nutzen zu können. Durch die rasche Entwicklung neuer generativer KI-Anwendungen ergeben sich auch stetig neue Nutzungsmöglichkeiten für den Bildungskontext. Es ist noch nicht abzusehen, welche weiteren Einsatzmöglichkeiten zur Anpassung der Lernumgebung an individuelle Lernbedarfe in naher Zukunft entstehen werden.

Exemplarische Anwendungen

Lehrkräfte können generative KI-basierte Anwendungen unter anderem nutzen, um Unterricht zu planen und Unterrichtsmaterial zu erstellen. Eine Anwendung zum Erstellen von Erklärvideos ist *simpleshow*[29]. Lehrkräfte schreiben ein Skript für das Video und bildgenerierende KI erstellt daraus ein entsprechendes Video, welches noch angepasst werden kann. Auf sprachgenerativer KI basierende Anwendungen können genutzt werden, um benachteiligten Schüler:innen Zugang zu Lernformen und -systemen zu ermöglichen. So können beispielsweise Texte überarbeitet oder übersetzt werden, um in heterogenen Klassen Texte an das sprachliche Niveau anzupassen oder Sprach-

[29] https://simpleshow.com/de/

barrieren abzubauen. *Summ AI*[30] oder *Capito.ai*[31] sind Beispiele solcher Tools, die es ermöglichen, Texte zu analysieren und vollautomatisch in verschiedene Sprachstufen (z.B. A1, A2, B1; Capito.ai) umzuwandeln. Bei unterschiedlichen Lernausgangslagen ist dies sehr hilfreich, um die Schüler:innen abzuholen und nicht zu überfordern.

Speech-to-Text Systeme wie *Whisper*[32] ermöglichen es, automatisch transkribierte Untertitel zu Lernvideos o.ä. zu erstellen. Dies kann es höreingeschränkten Schüler:innen erleichtern, die Inhalte zu verstehen. Text-to-Speech-Systeme wie *Prizmo Go* können Schüler:innen helfen, indem Texte gescannt und vorgelesen werden. Systeme wie *Seeing AI*[33], *Lookout*[34] oder *Be My AI* (noch in der Betaphase[35]) funktionieren ähnlich, wobei nicht Texte, sondern Bilder gescannt und in Sprache umgesetzt werden, so dass z.B. Sehbeeinträchtigungen ausgeglichen werden. Die speziell für Lese-Rechtschreibschwäche entwickelte Anwendung *Ridit*[36] überführt Fotos von gedruckten oder handgeschriebenen Texten u.a. in Legasthenie-freundliche Schriftarten (siehe auch *OpenDyslexic*[37]) und anpassbare Farbkontraste und erfüllt so eine assistive Funktion. In der inklusiven Praxis gestaltet sich der Einsatz aus Gründen des Datenschutzes schwierig.

Schüler:innen können mithilfe sprachgenerativer KI personalisiertes Feedback zu ihren Arbeitsergebnissen erhalten. Die Anwendungen können, je nachdem wie sie gepromptet werden, unterschiedliche Rückmeldungen geben: motivierend, korrigierend oder mit Empfehlungen einhergehend. Ein System, welches mithilfe von sprachgenerativer KI individualisiertes Feedback im Schreibprozess bietet, ist *PEER*[38]; ein System, das beim Schreiben und Verbessern von Aufsätzen konstruktives Feedback gibt. Ein kommerzielles Produkt zur Förderung von Schreibkompetenzen ist *fiete.ai*[39]. Lehrkräfte wählen für ihre Schüler:innen Aufgabenformate aus oder erstellen diese selbst und legen Kriterien fest. Anhand dieser Kriterien erhalten die Schüler:innen nach der Aufgabenbearbeitung eine Rückmeldung, wie gut sie dieses Kriterium erfüllt haben, und detailliertes Feedback mit Verbesserungsvorschlägen. Die von den Lehrkräften getroffene Auswahl stellt sicher, dass die Aufgaben und festgelegten Kriterien auf die Schüler:innen und ihren jeweiligen Leistungsstand abgestimmt sind. Grundsätzlich erscheint die Möglichkeit sinnvoll, Lernende durch KI-generiertes Feedback zu unterstützen, da die Wirksamkeit von Feedback gut belegt ist (Hattie & Timberley, 2007). Das lernförderliche Potenzial solcher Systeme kann daher hoch eingeschätzt werden, da Schüler:innen auf ihre Arbeitsergebnisse zeitnah individuelle Rückmeldung erhalten, die Lehrkräfte für einzelne Schüler:innen nicht leisten können. Allerdings kann die Art und

[30] https://summ-ai.com/
[31] https://www.capito.eu/
[32] https://openai.com/index/whisper/
[33] https://www.seeingai.com/
[34] https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal&hl=de&gl=US
[35] https://www.inklusive-medienarbeit.de/be-my-ai/
[36] https://apps.apple.com/us/app/ridit/id1442148798
[37] https://opendyslexic.org/
[38] https://peer-ai-tutor.streamlit.app/
[39] https://www.fiete.ai/

Präsentation des Feedbacks an die Schüler:innen entscheidenden Einfluss auf die Lernförderlichkeit haben (z.B. Jansen et al., 2024; Wisniewski et al., 2020). So macht es einen Unterschied, ob die Rückmeldung inhaltlich die Aufgabenbearbeitung aufgreift oder allgemein motivational formuliert ist. Ebenso kann Feedback zu verschiedenen Zeitpunkten, während oder nach der Aufgabenbearbeitung, gegeben werden.

Generative KI als eine Art Dialogpartner kann Schüler:innen dazu anregen, Themen zu explorieren und Fragen zu stellen, die im regulären Unterrichtsgeschehen möglicherweise nicht aufkommen würden. Insbesondere Schüler:innen mit Lernschwächen erhalten so die Möglichkeit, Fragen zu stellen und Antworten zu erhalten, die für sie persönlich relevant sind. Bekannte Anwendungen von Large Language Modellen sind beispielsweise Allzweck-Chatbots, die auf immens große Bestände an frei verfügbarem Textmaterial und Literatur zugreifen und aufgrund der damit trainierten Sprachmodelle eine scheinbar „intelligente" Konversation in natürlicher Sprache über nahezu jedes Thema führen können (z.B. *ChatGPT*, *Claude*). Über *Fobizz* und die dort angebotene *KI für Personen*[40] haben Schüler:innen die Möglichkeit, mit berühmten Menschen zu chatten. Die KI nimmt dabei die Rolle der berühmten Person ein und Schüler:innen können diese über ihr Wissen und ihre Erfahrungen ausfragen.

Eine weitere interessante Nutzung von Chatbots ergibt sich, wenn diese als „sokratische Dialogpartner" fungieren. Im Rahmen des sokratischen Dialogs gibt der Chatbot keine Antworten, sondern stellt gezielte Fragen, die die Schüler:innen anregen sollen, intensiver und kritisch reflektierend über ein Thema nachzudenken und eigene Schlussfolgerungen zu ziehen. Im inklusiven Kontext kann dies relevant sein, da der Dialog an das Niveau der Lernenden angepasst stattfindet und der Chatbot gezielt aufklärende Fragen stellen kann, um das Verständnis zu fördern.

Ein bekanntes Beispiel aus den USA ist *KhanMigo*[41] (Anand, 2023), das von der Khan Academy auf der Grundlage von *ChatGPT* entwickelt wurde. Auch diesem Chatbot können die Schüler:innen Fragen zu verschiedenen Fachgebieten stellen. Zurzeit werden sowohl in der Forschung, als auch auf Seiten der Unternehmen eine ganze Reihe weiterer Anwendungen für das Lernen entwickelt, die Large Language Modelle nutzen. Insgesamt lässt sich feststellen, dass sich der Bereich der Large Language Modelle sehr schnell weiterentwickelt. Es werden häufig neue Modelle erstellt; ebenso werden regelmäßig neue Versionen bestehender Modelle mit verbesserter Leistung produziert. Large Language Modelle werden bislang allerdings meist von Unternehmen ohne Bezug zum Bildungssektor entwickelt (z.B. *OpenAI*, *Google*, *Microsoft*, *Meta*). Dies ist vor allem dadurch bedingt, dass die Entwicklung sehr viele Ressourcen erfordert.

Die Anzahl an generativen KI-Anwendungen, auf die Lehrkräfte zurückgreifen können, ist mittlerweile nahezu unüberschaubar. Um diese Systeme im schulischen Feld einsetzen zu können, braucht es allerdings eine gewisse Qualitätskontrolle sowie einen datenschutzkonformen Zugang über entsprechende DSGVO-konforme Schnittstellen.

[40] https://tools.fobizz.com/ai/personas/info
[41] https://www.khanmigo.ai/

Seit Mitte 2018 gibt es den Anbieter *Fobizz*[42], der über eine Plattform Lehrkräften die Möglichkeit bietet, einen datenschutzkonformen Zugang zu generativen KI-Werkzeugen zu erhalten und diese für ihren Unterricht zu nutzen. Ein weiterer bekannter Anbieter ist *schulKI*[43].

Empirische Befunde

Die meisten bisher vorliegenden Forschungsergebnisse beziehen sich auf den Einsatz von generativer KI und dabei vor allem auf den Einsatz von Anwendungen, die Schüler:innen Feedback geben. Dai und Kolleg:innen (2023) konnten zeigen, dass *ChatGPT* in der Lage ist, den Lernenden ein konsistentes, detailliertes und individualisiertes Feedback auf Basis vordefinierter Bewertungskriterien (sogenannter Rubriken) zu geben. Dabei kann das KI-generierte Feedback eine hohe Übereinstimmung mit dem Feedback menschlicher Tutor:innen haben und im Vergleich zu menschlichem Feedback qualitativ ähnlich gut sein (Steiss et al., 2024). Meyer und Kolleg:innen (2024) konnten empirisch nachweisen, dass sich KI-generiertes Feedback im Vergleich zu keinem Feedback positiv auf die Lernleistung, aber auch auf motivationale und emotionale Aspekte des Lernens auswirken kann. Die Ergebnisse einer Studie von Jansen und Kolleg:innen (2024) legen nahe, dass Lernende bessere Leistungen bei der Überarbeitung ihrer Texte erzielen, wenn sie leistungsbasiertes Feedback erhalten, welches anhand einer automatisierten Textevaluierung generiert wurde.

Eine lernförderliche Wirkung sprachgenerativer KI-Anwendungen in Lehr- und Lernsettings scheint insbesondere gegeben zu sein, wenn diese als eine Art Dialogpartner eingesetzt werden. In einer Studie von Guo und Kolleg:innen (2023) konnte gezeigt werden, dass der Einsatz von Chatbots in Debatten im Schulunterricht die Argumentationsfähigkeit sowie das Engagement der Lernenden steigern kann.

Aktuelle Nutzerzahlen (z.B. Fobizz.com) zeigen, dass Lehrkräfte sehr stark auf KI-Werkzeuge zurückgreifen. Viele Lehrkräfte sehen den Einsatz generativer KI als Chance zur beruflichen Weiterentwicklung und zur Reduktion von Stress im schulischen Alltag (Collie & Martin, 2024). Es fehlen jedoch Studien, die sich mit den Einstellungen und Wahrnehmungen von Lehrkräften gegenüber generativer KI befassen, da diese Faktoren die Nutzung solcher Technologien maßgeblich beeinflussen können (z.B. Kizilcek, 2024).

Insgesamt liegen im Bereich Lernen und dem Einsatz generativer KI bislang nur begrenzte Forschungsergebnisse vor, wobei die meisten Studien laborexperimentell durchgeführt wurden und sich auf spezifische Detailfragen konzentrieren. Es besteht ein dringender Bedarf an Feld- sowie Replikationsstudien, insbesondere im deutschsprachigen Raum, um fundierte Aussagen über die Effektivität generativer KI-Anwendungen zur Lernunterstützung im Unterricht treffen zu können.

[42] https://plattform.fobizz.com/users/sign_in

[43] https://schulki.de/

Einordnung
Generative KI, die zur Anpassung der Lernumgebung an individuelle Lernbedarfe genutzt wird, ermöglicht es Lehrkräften, in kurzer Zeit Unterrichtsmaterial zu differenzieren (z.B. angepasst an das Sprachniveau, übersetzt in andere Sprachen, durch Anpassung des Umfangs etc.) und somit an die Kompetenzen und den Wissensstand einzelner Schüler:innen oder Schülergruppen angepasstes Material zu generieren. Dies erleichtert den Umgang mit heterogenen Lerngruppen und kann eine zeitliche Entlastung schaffen. Darüber hinaus können beispielsweise material- und bildgenerierende Systeme eingesetzt werden, um den Unterricht inklusiv zu gestalten und Nachteile auszugleichen. Schüler:innen scheinen insbesondere durch KI-generiertes Feedback zu profitieren, wobei es weiterer Grundlagenforschung bedarf.

Es ist wichtig zu betonen, dass generative KI-Anwendungen keine Wissensmodelle zugrunde liegen und daher nicht sichergestellt ist, ob generierte Inhalte wie Zusammenfassungen, Feedback, Aufgabenmaterial oder vorgeschlagene Lösungswege korrekt und sinnvoll sind. Hier bedarf es der Überprüfung durch Menschen.

5.4. Fazit der Potenzialanalyse

Über die Cluster hinweg wird deutlich, dass digitale Technologien und vor allem KI-Systeme gute Möglichkeiten für adaptive Lernunterstützung und individuelle Förderung bieten. Ebenso zeigt sich aber, dass die Potenziale bisher noch nicht ausgeschöpft werden. Aus den genannten Beispielen und Befunden geht hervor:

Cluster 1: Systeme zur Lerndatenanalyse

- Die systematische Erfassung von Lernständen bietet die Möglichkeit, Lernverläufe abzubilden und Lernverlaufsentwicklung abzuschätzen.
- In Deutschland werden noch kaum Systeme im Regelbetrieb in den Schulen eingesetzt. Dies gilt umso mehr für Systeme mit KI-Unterstützung.
- Existierende Analysesysteme werden bislang am ehesten punktuell zur Diagnostik verwendet. Diese Diagnostik ist aber nicht systematisch mit darauf aufbauender Intervention verbunden. Eine Verschränkung von Systemen zur Lerndatenanalyse und Systemen zur KI-gestützten adaptiven Förderung ist zurzeit nur bedingt realisiert.
- Ein systematisches Verfolgen von Lernverläufen ist zumindest in Deutschland in Schulen noch nicht möglich. Es fehlen regulative Voraussetzungen im Sinne des Datenschutzes.

Cluster 2: Systeme zur adaptiven Lernunterstützung

- Die Wissensbasierung dieser Systeme bedingt ihre Lernförderlichkeit. Dies ist in Abgrenzung zu Systemen, die rein auf generativer KI beruhen, zu betonen.
- Die Entwicklung von adaptiven Systemen ist fach- bzw. inhaltsspezifisch und somit nur bedingt auf andere Inhalte übertragbar.
- Die Entwicklung adaptiver Systeme ist zeit- und kostenintensiv.
- Forschung findet bislang schwerpunktmäßig in den USA statt. Dort werden diese Systeme auch seit vielen Jahren entwickelt.
- In Deutschland werden bislang kaum Systeme im Regelbetrieb in den Schulen eingesetzt. Dies gilt umso mehr für Systeme mit wissensbasierter KI-Unterstützung.
- Es braucht Wege von der Forschung in die Praxis, um in der Forschung positiv evaluierte Systeme in Schulen zu implementieren.
- Eine curriculare Einbindung und pädagogische Gesamtkonzepte für die Nutzung im Unterricht werden benötigt.
- Es fehlen regulative Voraussetzungen. Insbesondere muss der *AI Act* berücksichtigt werden, da Lerndaten verarbeitet werden und durch adaptive Anpassungen nicht alle Schüler:innen die gleichen Lerngelegenheiten bekommen.

Cluster 3: Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe

- Unterschiedlichste Anwendungen auf Grundlage generativer KI werden eingesetzt, aber nicht systematisch. Es fehlen pädagogische Konzepte.
- Die Forschung in diesem Cluster ist im Anfangsstadium und noch sehr grundlagenorientiert (kontrollierte Studien zu spezifischen Forschungsfragen).
- Es besteht Bedarf an Wissen über die Grenzen und Herausforderungen dieser KI-Systeme.
- Es fehlen regulative Voraussetzungen; insbesondere sind Fragen des Datenschutzes zu klären.

Übergreifend für die drei betrachteten Cluster:

- KI-gestützte Systeme und Anwendungen bieten die Möglichkeit, gezielte Förderung zu unterstützen und individuelles Lernen zu verbessern.
- Um dieses Potenzial entfalten zu können, ist es notwendig:
    - regulative Voraussetzungen zu klären,
    - die Systeme unter Beteiligung wissenschaftlicher Forschung weiterzuentwickeln,
    - die Anwendungen curricular einzubinden und in pädagogische Konzepte zu integrieren.

6. Handlungsempfehlungen

Die Handlungsempfehlungen zu den drei Clustern umfassen jeweils Aspekte, die vor einem Einsatz der KI-Systeme zu klären, zu erarbeiten oder zu prüfen sind. Der zweite Punkt bezieht sich auf notwendige oder wünschenswerte Weiterentwicklungen. Der dritte Punkt fasst clusterspezifische Aspekte der Einbindung und Unterstützung von Lehrkräften zusammen. Über diese spezifischen Handlungsempfehlungen für jedes der drei Cluster hinaus, können übergreifende Empfehlungen formuliert werden.

Cluster 1: Systeme zur Lerndatenanalyse

Qualitätsstandards sind notwendig, um zu gewährleisten, dass eine gute Diagnostik von Lernständen und Lernprozessen (siehe Abbildung 1 oben) stattfindet. Eine Erarbeitung von Richtlinien durch Expert:innen zur Unterstützung bei der Auswahl passender Systeme wäre hilfreich. Die Einführung von Systemen sollte systematisch durch die Landesministerien initiiert und koordiniert werden, da hier übergeordnete Bildungsziele eine Rolle spielen. Schulen und Lehrkräfte dürfen bei der Auswahl nicht sich selbst überlassen werden.

Eine Weiterentwicklung mit dem Ziel einer gezielten **Integration von Analyse- und Unterstützungssystemen** sollte angestrebt werden. Lerndatenanalysen bieten Ansatzpunkte für individuell zugeschnittene Unterstützung und Lernförderung der Schüler:innen. Dieses Potenzial sollte besser genutzt werden. Lehrkräfte sollten bei der Umsetzung begleitet werden, indem Konzepte zur Umsetzung von Diagnosen in passende Interventionen bereitgestellt werden.

Lehrkräfteschulungen sind zentral, um sicherzustellen, dass Lehrkräfte verstehen, wie die Systeme diagnostisch vorgehen und wie die repräsentierten *Learning*

Analytics zu interpretieren sind. Dadurch gewinnen die Lehrkräfte an Handlungssicherheit und sie können fundierte Entscheidungen treffen.

Cluster 2: Systeme zur adaptiven Lernunterstützung

Um geeignete Systeme zur individuellen Lernunterstützung identifizieren zu können, sollte eine **Prüfung angebotener Software** durchgeführt werden. Viele Systeme versprechen eine adaptive Anpassung an den Lernstand und ggf. den Lernfortschritt der Schüler:innen. Es ist wichtig zu verstehen, welche Form der Adaptivität vorliegt und welche Regeln dieser zugrunde liegen. Ein Leitfaden, siehe Anhang 1, bietet eine Orientierung. Dieser Leitfaden erlaubt Nicht-Experten anhand bestimmter Kriterien und Leitfragen das Ausmaß der Adaptivität eines Systems einzuschätzen.

Ein **systematischer Auf- und Ausbau von Systemen** ist anzustreben. Bereits evaluierte, genutzte und empfohlene Software sollte als Ressource für die Entwicklung von KI-gestützter adaptiver Lernunterstützung dienen. Auf diesen Systemen aufbauend, sollten Wege gefunden werden, um mithilfe von KI eine individuell zugeschnittene Unterstützung und Lernförderung zu ermöglichen. Eine schrittweise curriculumsnahe Ausweitung, wie am Beispiel von Carnegie Learning verdeutlicht, ermöglicht, dass Anpassungen und Verbesserungen durch iterative Entwicklungs- und Evaluationszyklen stattfinden (*design loop adaptivity*). Dadurch kann ein ressourcenschonender Ausbau stattfinden, der auch dynamische Entwicklungen zulässt.

Zudem kann durch **Schulungen** und die partizipative Einbindung der Lehrkräfte eine Integration und Akzeptanz der Systeme gefördert werden.

Cluster 3: Systeme zur Anpassung der Lernumgebung an individuelle Lernbedarfe

Es muss eine **Prüfung der Einsatzmöglichkeiten** von generativen KI-Systemen erfolgen, um eine klare Ausrichtung auf das Ziel einer individuell zugeschnittenen Unterstützung und Lernförderung der Schüler:innen zu gewährleisten. Auf dieser Grundlage können Anwendungsfelder für den Einsatz von generativer KI bestimmt werden, die spezifische Lernende bzw. Lernendengruppen ansprechen und für diese lernförderlich wirken. Nützlich wäre eine zentrale Bereitstellung geeigneter Systeme, um ausgewählte Angebote zu bündeln und Lehrkräften zu helfen, einen Überblick zu erhalten.

Auch ist die **Erarbeitung von pädagogisch-didaktischen Konzepten** und Weiterentwicklung der Curricula notwendig, um Schüler:innen zu helfen, generative KI produktiv als Werkzeug zu benutzen. Für Lehrkräfte und die Bewertung von Lernprozessen bedeutet dies ein Umdenken, weg von der Bewertung von Lernergebnissen, hin zu einer Bewertung von Lernprozessen.

Es ist zentral, **Lehrkräfte zu schulen**, damit sie für die Einschränkungen von generativer KI sensibilisiert sind und diese in ihrem Unterricht entsprechend berücksichtigen können. Lehrkräfte müssen ihr Wissen über KI an ihre Schüler:innen weitergeben. Nur so erhalten diese die Möglichkeit, langfristig einen angemessen kritischen Umgang mit KI-Systemen zu erlernen und eigenständig die Einsatzmöglichkeiten und Unterstützungsleistung einzuschätzen.

Handlungsempfehlungen über die drei betrachteten Cluster hinweg:

Evidenz einfordern

Der Einsatz von KI-Systemen sollte nicht aufgrund der Innovation erfolgen. Systeme sollten nicht nur den Anspruch erheben, adaptiv zu sein, sondern die individuell zugeschnittene Unterstützung und Lernförderung der Schüler:innen sollte nachprüfbar sein. Dabei sollten sie eine lernförderliche Wirkung nachweisen können und bessere Ergebnisse in der Wissens- und Kompetenzvermittlung erzielen als alternative Methoden.

Daten nutzen

Als Datengrundlage für differenzierte Erkenntnisse zur Lernförderlichkeit von KI-Systemen können Daten aus Pilotprojekten dienen, die im Rahmen der Umstellung auf digitalisierte Formate erhoben werden. Ebenso werden zunehmend länderspezifische Lernstandserhebungen digitalisiert. Darüber hinaus erwerben Bundesländer landesweite Lizenzen für Systeme, die Schulen zur Verfügung stehen. Diese großen Datenmengen könnten genutzt werden, um die Lernförderlichkeit von KI-Systemen im Feld zu evaluieren. Idealerweise würden die wissenschaftlichen Erkenntnisse aus diesen Studien wiederum in die Entscheidungen der Bundesländer über den Einsatz von Software in Schulen einfließen.

Kontinuierliche Beratung durch Wissenschaftler:innen

Um lehr- und lernförderliche Systeme einordnen zu können, ist es wichtig, Wissenschaftler:innen mit entsprechender fachlicher Expertise beratend in Entscheidungen mit einzubeziehen. So können informierte Entscheidungen über die Auswahl und den Einsatz von KI-Systemen mit einem fundierten Entscheidungsrationale getroffen werden.

Regulative Voraussetzungen schaffen

Es gilt, eine Balance zu schaffen zwischen den Chancen der individuellen Förderung durch KI und der Wahrung von Datenschutz und Persönlichkeitsrechten von Schüler:innen. Insbesondere muss der *AI Act* berücksichtigt werden, da Lerndaten verarbeitet werden und durch adaptive Anpassungen nicht alle Schüler:innen die gleichen Lerngelegenheiten bekommen.

Lehrkräfte unterstützen und schulen

Ohne die Lehrkräfte können KI-Technologien ihre Potenziale zur individuellen Unterstützung und Förderung der Schüler:innen nicht entfalten. Lehrkräfte müssen durch rahmengebende Regulation abgesichert und durch Richtlinien und Umsetzungskonzepte unterstützt werden. Die verschiedenen in den drei Clustern angesprochenen KI-Anwendungen erfordern zudem sehr unterschiedliches Wissen und unterschiedliche Kompetenzen auf Seiten der Lehrkräfte. Daher müssen in allen drei Phasen der Lehrkräftebildung die dort vermittelten Inhalte um das Thema KI erweitert und kontinuierlich weiterentwickelt werden. Lehrkräfte brauchen ausreichend technologisches sowie technologisch-pädagogisches Wissen sowie entsprechende Kompetenzen, um reflektierte und informierte Entscheidungen bezüglich des Einsatzes von KI-Systemen in ihrem Unterricht treffen können.

Systematisch Forschungsevidenz aufbauen

Um das Potenzial von KI-Technologien angesichts der raschen Entwicklungen neuer Technologien auch zukünftig sinnvoll einschätzen zu können, bedarf es dringend weiterer Grundlagenforschung. Damit sind nicht Nutzerstudien für einzelne Systeme gemeint, sondern ein systematischer Aufbau von empirischer Evidenz zu den Prinzipien dessen, was unter Einbezug von KI-Systemen lehr- und lernförderlich wirksam ist.

Forschungsergebnisse differenziert betrachten

Das Generalisieren von Forschungsergebnissen auf die Gesamtheit der Schüler:innen differenziert nicht ausreichend. Es ist sinnvoll, für Lerngruppen differenzierende Studienanalysen durchzuführen. Manche Systeme wirken nicht auf alle Lernendengruppen gleichermaßen effektiv; sie müssen unterschiedlich bewertet werden, je nach Leistungsstärke und besonderen Bedürfnissen der Schüler:innen. Nur durch differenzierte Studienanalysen lässt sich Heterogenität in Lernsettings angemessen adressieren.

Implementation ins Feld wissenschaftlich begleitet vorantreiben

Grundlagenforschung benötigt Zeit, was kaum vereinbar ist mit der rasanten Entwicklung neuer Systeme. Zudem besteht auch aus bildungspolitischer Sicht die Forderung, KI-Systeme in die Schulpraxis zu integrieren. Somit entsteht ein Bedarf

auf Seiten der Schulpraktiker:innen, bestehende KI-Systeme in den Unterricht zu integrieren. Folglich müssen neben Grundlagenforschung vorhandene Systeme genutzt werden, deren Evaluation noch aussteht. Dieser Einsatz sollte jedoch wissenschaftlich begleitet in Implementations- und Evaluationsstudien stattfinden. Es ist wichtig, dass diese nicht nur unter kontrollierten Bedingungen stattfinden, sondern auch im Feld, sprich unter realen Bedingungen in Klassenzimmern mit Lehrkräften und Schüler:innen durchgeführt werden. Die durch den Einsatz von Systemen gewonnenen Daten sollten auch für die Grundlagenforschung zur Verfügung stehen.

Internationale Beispiele berücksichtigen

Internationale Forschungserkenntnisse und Best Practice Beispiele einzubeziehen kann hilfreich sein. Anhand der Entwicklung von Intelligenten Tutoring-Systemen kann veranschaulicht werden, wie der Weg ausgehend von der Grundlagen-forschung, über experimentelle Entwicklung und empirische begleitete Überarbeitungszyklen und Evaluationen, bis hin zur Implementation in Schulen und die Ausgründung in ein Unternehmen aussehen kann.

Einbindung relevanter Stakeholder

Die Partizipation relevanter Beteiligter, wie Lehrkräften, Schüler:innen, Schul-leitungen, Eltern, bildungspolitischen Entscheider:innen u.a., bei der nutzer-zentrierten Entwicklung, Erforschung und Evaluation von KI- Systemen ist von großer Relevanz. Dies kann, der raschen Entwicklung von neuen Systemen geschuldet, zyklisch verlaufen, was zu einer stetigen Neubewertungen führt.

Verzahnung von Forschung, Entwicklung und Nutzung

Bereits in den ergänzenden Empfehlungen der KMK (2021) wurde empfohlen, dass Strategien zwischen Forschung, Bildungsmedienanbietern und der Praxis entwickelt werden sollen, um die Potenziale des Marktes aufzugreifen sowie die zielgerichtete Anbindung von digitalen Lernumgebungen in die Lernplattformen der Schüler:innen zu ermöglichen. Vermarktung, aber auch Instandhaltung, Wartung, Updates und weiteres sind wichtige Aufgaben, die durch Forschungsprojekte nicht gewährleistet werden können, die aber relevante Faktoren sind, um eine breite Anwendung und langfristige Implementation in Schulen zu gewährleisten. Es gilt jedoch stets, dass die Bildungsziele nicht kommerziellen Interessen unterzuordnen sind (Varsik & Vosberg, 2024).

Schaffen von technischen Voraussetzungen

Eine ausreichende technische Infrastruktur und deren kontinuierliche Aktualisierung und Wartung ist an Schulen vorauszusetzen und sicherzustellen, um Unterschiede hinsichtlich der Zugangsmöglichkeiten zu KI-Anwendungen zu nivellieren. Die in diesem Gutachten beschriebenen Potenziale hängen maßgeblich davon ab, ob

entsprechende technische Voraussetzungen sowohl in den Schulen, als auch bei den Schüler:innen vorhanden sind.

7. Ausblick

Abschließend werden Herausforderungen beleuchtet, die bei der Nutzung der vorgestellten KI-gestützten Technologien zur individuellen Förderung im Bildungskontext adressiert werden müssen. Auch sollen zukunftsweisende Überlegungen zur Weiterentwicklung dieser Technologien erörtert werden.

7.1. Zentrale Herausforderungen

Umgang mit Daten

Ein wesentlicher Aspekt beim Einsatz von KI in Bildungskontexten ist der Umgang mit Daten. Intelligente Tutoring-Systeme sind darauf angewiesen, Daten zu sammeln und zu speichern, um individuelles Feedback zu geben und adaptiv Aufgaben zuzuweisen. Entscheidend ist dabei, wo diese Daten gespeichert und wie sie anonymisiert oder pseudonymisiert werden. Es muss sichergestellt sein, dass keine Rückschlüsse auf einzelne Schüler:innen durch Unternehmen möglich sind. Generative KI-Systeme sind auf umfangreiche Datensätze angewiesen, auf denen sie trainiert werden. Dies birgt das Risiko, dass sensible Inhalte in Trainingsdaten aufgenommen und dort gespeichert bzw. bei Anfragen aus den Trainingsdaten reproduziert werden.

Grundsätzlich ist beim Einsatz von KI im Bildungsbereich der geltende Rechtsrahmen zu beachten, insbesondere die europäische Verordnung über Künstliche Intelligenz, der EU *AI Act* (Verordnung (EU), 2024). Diese Verordnung wurde am 12. Juli 2024 im Amtsblatt der Europäischen Union veröffentlicht und trat am 1. August 2024 in Kraft. Der EU *AI Act* teilt Technologien in drei Risikogruppen ein: minimales, hohes und unannehmbares Risiko. Anhang III der Verordnung beschreibt Anwendungsfälle im schulischen Kontext. Unter folgenden Bedingungen gelten Technologien als hochriskant:

> (a) KI-Systeme, die dazu dienen, den Zugang oder die Zulassung zu Einrichtungen der allgemeinen und beruflichen Bildung auf allen Ebenen zu bestimmen oder natürliche Personen zuzuweisen;
>
> (b) KI-Systeme, die zur Bewertung von Lernergebnissen eingesetzt werden sollen, auch wenn diese Ergebnisse zur Steuerung des Lernprozesses natürlicher Personen in Einrichtungen der allgemeinen und beruflichen Bildung auf allen Ebenen verwendet werden;
>
> (c) KI-Systeme, die zur Bewertung des angemessenen Bildungsniveaus, das eine Person erhalten wird oder zu dem sie Zugang haben wird, im Rahmen von oder innerhalb von Einrichtungen der allgemeinen und beruflichen Bildung auf allen Ebenen eingesetzt werden sollen;

> (d) KI-Systeme, die zur Überwachung und Erkennung von unzulässigem Verhalten von Schülern bei Prüfungen im Rahmen von oder in Einrichtungen der allgemeinen und beruflichen Bildung auf allen Ebenen eingesetzt werden sollen.
>
> (Verordnung (EU), 2024; Anhang III, 3.)

Hochriskant bedeutet nicht, dass der Einsatz der Technologien verboten ist. Vielmehr werden hohe Anforderungen, beispielsweise hinsichtlich der Datenverwaltung, der technischen Dokumentation oder der Überwachung durch (menschliche) Aufsichtsbehörden formuliert (Verordnung (EU), 2024; Kapitel III, Abschnitt 2). Für die unter Anhang III aufgeführten Systeme gilt die Verordnung spätestens 24 Monate nach ihrem Inkrafttreten (Future of Life Institute, 2024). Zudem gilt die Datenschutz-Grundverordnung (DSGVO), nach der die Verarbeitung von Schüler:innen- oder Lehrer:innendaten entweder durch eine konkrete Rechtsgrundlage oder durch Einwilligung der Betroffenen gestattet sein muss (bzw. bei minderjährigen Schüler:innen in der Regel durch die Eltern, SchulG NRW, 2024,§§ 120 bis 122). Dabei regelt der Datenschutz nicht nur den Umgang mit personenbezogenen Daten, wie die Erfassung von Noten, Lern- und Förderbedarfen, sondern auch die Übermittlung von Daten an externe Dienstleister. Zudem ist die Leistungsbewertung, wie auch schon im *AI Act* angeklungen, eine pädagogische und hoheitliche Aufgabe, die ausschließlich von Lehrkräften erfüllt werden kann. Das Profilingverbot nach Artikel 22 DSGVO (Datenschutz-Grundverordnung, 2016) regelt, dass keine Entscheidung ausschließlich automatisiert erfolgen darf und somit immer eine menschliche Letztentscheidung gewährleistet sein muss. In den einzelnen Bundesländern bestehen unterschiedliche Richtlinien. Hinsichtlich datenschutzrelevanter Vorgaben in Nordrhein-Westfalen bieten die Internetseiten des Ministeriums für Schule und Bildung NRW eine Orientierung (MSB NRW, 2024a). Neben dem Datenschutz sind auch ethische Aspekte zu reflektieren. Die „Ethische[n] Leitlinien für Lehrkräfte über die Nutzung von KI und Daten für Lehr- und Lernzwecke“ (Europäische Kommission, 2022) sowie die Ethik-Leitlinien für eine vertrauenswürdige KI der unabhängigen Expertengruppe für Künstliche Intelligenz (Europäische Kommission, 2019), geben Impulse zu diesen Aspekten. Der Deutsche Ethikrat (2013) weist in diesem Zusammenhang auf zu bedenkende Aspekte hin und spricht Empfehlungen für den Bildungsbereich aus.

Insgesamt wird deutlich: Der grundsätzliche Einsatz bestimmter KI-Technologien sowie Datenschutzaspekte von Anwendungen muss an übergeordneter Stelle geklärt werden (Denker et al., 2021). Ein Schritt in diese Richtung wurde vom Ministerium für Schule und Bildung NRW unternommen, als im Februar 2023 der Handlungsleitfaden zum Umgang mit textgenerierenden KI-Systemen veröffentlicht wurde (MSB NRW, 2023). Im Januar 2024 folgte das Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz zu Large Language Modellen und deren Potenziale im Bildungssystem (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz [SWK], 2023). In Nordrhein-Westfalen hat die Taskforce „Künstliche Intelligenz im Bildungswesen“ (2024) umfassender und nicht nur auf generative KI bezogene Empfehlungen herausgegeben, wie eine verantwortungsvolle Integration von KI in Lehr-

und Lernprozesse entlang der Bildungskette nachhaltig gefördert werden kann. Für den Herbst 2024 wird die Veröffentlichung eines Themenkatalogs zum Umgang mit KI an Schulen erwartet, den die Bundesländer aktuell gemeinsam erarbeiten. Zudem befassen sich verschiedene Projekte mit dem Thema Datenschutz mit dem Ziel, Lehrkräfte und Schulträger entsprechend zu unterstützen (z.B. *eduCheck digital*[44], *DIRECTIONS*[45] sowie *VIDIS*[46].

Nutzung von generativer KI im Bildungsbereich

Seit der Vorstellung von *ChatGPT* im November 2022 haben sich Anwendungen generativer KI, die auf Large Language Modellen basieren, rasant entwickelt. Im Bildungsbereich wird die Nutzung kritisch diskutiert. Obwohl sprachgenerative KI in der Lage ist, ähnlich wie ein menschliches Gegenüber zu reagieren, basiert der Output der Modelle lediglich auf einer sophistizierten Berechnung wahrscheinlicher Wortfolgen für den gegebenen Kontext aufgrund von umfangreichen Trainingsdaten und beruht nicht auf Wissensmodellen. Dadurch ergeben sich eine Reihe von Schwierigkeiten, die im Folgenden kurz zusammengefasst werden. Diese Problematiken gelten grundsätzlich für alle Arten generativer KI. Aufgrund des hohen Bekanntheitsgrades von sprach-generativen Anwendungen, fokussieren die Beispiele auf Large Language Modelle.

Eine bekannte Einschränkung von Large Language Modellen ist das Auftreten sogenannter „Halluzinationen". Damit ist das Generieren falscher oder erfundener Informationen durch die KI gemeint. Eine KI generiert Informationen grundsätzlich ohne Hinweis auf ihre faktische Korrektheit. Es bedarf entsprechendem Fachwissen, um diese Inhalte als richtig oder falsch zu identifizieren. Halluzinationen treten insbesondere dann auf, wenn ein Modell komplexe oder spezifische Sachinformationen generieren soll. Größere und fortschrittlichere Modelle sind zwar besser in der Lage, diese Fehler zu minimieren, sie produzieren jedoch immer noch falsche Ergebnisse.

Bei generativen KI-Systemen, wie Large Language Modellen, ist es auch für Entwickler, z.B. von Technologieunternehmen, die diese in ihren Anwendungen implementieren, nicht möglich, die Generierung spezifischer Ergebnisse vollständig zu verstehen. Diese mangelnde Transparenz von Funktionsweisen und Prozessen erschwert den Aufbau von Vertrauen in die Ergebnisse der KI-Systeme und kann sich negativ auf die Akzeptanz bei Lehrkräften auswirken.

Large Language Modelle können unbeabsichtigt Verzerrungen und Vorurteile übernehmen, die in den Daten vorhanden sind, auf denen sie trainiert wurden, und diese aufrechterhalten und sogar verstärken. Da die Modelle auf großen, vielfältigen Datensätzen aus dem Internet trainiert werden, können sie gesellschaftliche Vorurteile in Bezug auf Ethnie, Geschlecht, Kultur und ähnliche Aspekte erlernen. Dies kann zu Ergebnissen führen, die gegenüber Minderheiten voreingenommen, diskriminierend oder schädlich sind. Dieses Trainieren voreingenommener Haltungen kann auch absichtlich

[44] https://educheck.schule/
[45] https://directions-cert.de/
[46] https://www.vidis.schule/

geschehen, was ein erhebliches Missbrauchspotenzial mit sich bringt. Es bedarf des Fine-Tunings von Large Language Modellen, um Voreingenommenheiten abzutrainieren.

Aus den beiden vorangehenden Punkten ergeben sich zwangsläufig ethische Bedenken für den Einsatz generativer KI. Es besteht die Möglichkeit des Missbrauchs von Large Language Modellen zur Erzeugung von Fehlinformationen, Nachahmungen oder anderen schädlichen Inhalten. Das Missbrauchspotenzial ist erheblich, insbesondere da die Ergebnisse dieser Modelle derzeit oft nicht von Inhalten, die von Menschen erstellt wurden, zu unterscheiden sind. Daher ist es von großer Wichtigkeit, ein adäquates Bewusstsein bei Lehrkräften und Schüler:innen zu fördern und die durch Large Language Modelle generierten Inhalte kritisch zu reflektieren.

7.2. Neue Entwicklungsrichtungen

Entwicklungsaufwand für Intelligente Tutoring-Systeme reduzieren
Generative KI kann auch als Werkzeug für die Entwicklung anderer künstlich-intelligenter Bildungstechnologien dienen. Beispielsweise können KI-Anwendungen auf der Basis von Large Language Modellen genutzt werden, um Quizzes, Aufgaben, Texte, Bilder oder Videos sowie Hinweise und Rückmeldungen zu generieren. Diese Inhalte könnten dann in Intelligente Tutoring-Systeme integriert werden. Auf diese Weise ließe sich möglicherweise der hohe Entwicklungsaufwand reduzieren, der u.a. bislang die breite Skalierung Intelligenter Tutoring-Systeme eingeschränkt hat. Dadurch könnten die Potenziale von Intelligenten Tutoring-Systeme, die transparente, akkurate und erprobte KI-Systeme darstellen, für die individuelle Förderung und breitere Nutzung zugänglich gemacht werden.

Stärken symbolischer und generativer KI kombinieren
Eine Kombination von Intelligenten Tutoring-Systemen und generativer KI eröffnet neue Möglichkeiten, die Intelligenten Tutoring-Systeme für bestimmte Gruppen von Lernenden zugänglicher zu machen. Durch den Einsatz von Large Language Modellen bzw. Chatbots kann eine natürlich-sprachliche Interaktion ermöglicht werden, die in Kombination mit Intelligenten Tutoring-Systemen genutzt werden kann. So kann die Interaktion der Schüler:innen mit der Lernplattform intuitiver und barrierefreier gestaltet werden. Insbesondere Schüler:innen mit Einschränkungen und Schwächen können von der verbesserten Zugänglichkeit und der motivationalen bzw. emotionalen Unterstützung durch natürlichsprachliche Interaktion profitieren. Solche hybriden KI-Systeme könnten eine Lernumgebung schaffen, die sowohl kognitive als auch motivationale und emotionale Bedürfnisse stärker berücksichtigt.

Stärken von Lehrkräften und KI kombinieren
Bei der Frage, ob und in welchem Maße KI das Lernen fördern kann, ist es wichtig, die Rolle der Technologie im Kontext des Unterrichts zu betrachten und ein erweitertes Rollenverständnis von Lehrkräften zu berücksichtigen. Schon zu Beginn der Entwicklung des computerbasierten Lernens wurde die Frage aufgeworfen, ob ein Computer eine

Lehrkraft ersetzen kann. Auch heute gibt es Bemühungen um „lehrerlosen" Unterricht[47]. Sinnvoller ist es, wie bereits mit dem Strategiepapier der KMK (2016) zum Ausdruck gebracht, KI-Systeme als Ergänzung und Unterstützung für Lehrende und Lernende in bestimmten Phasen des Unterrichts zu sehen. KI kann Routinearbeiten übernehmen, Übungsmöglichkeiten schaffen, Monitoring erleichtern oder auf typische Fehler der Lernenden reagieren. Allerdings kann die KI nicht die Fähigkeit menschlicher Flexibilität ersetzen (Holstein & Aleven, 2022). So können beim hybriden Mensch-KI-Tutoring Schüler:innen im Unterricht mit KI-basierten Tutoring-Systemen arbeiten. Zusätzlich erhalten diejenigen, die zusätzliche Unterstützung und Motivation benötigen, gezielte Interaktionen mit einem menschlichen Tutor bzw. einer menschlichen Tutorin. Die Interaktion mit menschlichen Tutor:innen kann die Schüler:innen zur Weiterarbeit motivieren (Thomas et al., 2024). Die Rolle der KI kann also variieren, hat aber klar den Auftrag zu unterstützen, zu helfen und zu fördern. Die Wirksamkeit dieser Rollenverteilung mit Blick auf die Frage, wie lernförderlich sich diese auf einzelne Schüler:innen auswirkt, kann besonders in heterogenen Lehr-Lernkontexten davon abhängen, wie Schüler:innen diese empfinden und annehmen können. Herauszufinden, inwiefern Lehrkräfte durch die Komplementarität mit KI entlastet und unterstützt werden können, ist und bleibt auch zukünftig eine Fragestellung, der weiter nachgegangen werden sollte.

Zusammenfassend kann festgehalten werden: Individuelle Förderung ist eine der zentralen Herausforderungen in Zeiten von Lehrkräftemangel und großen Klassen mit zunehmend heterogener Schülerschaft. Das Gutachten hat deutlich gemacht, dass KI-Technologien Lehrkräfte und Lernende dabei unterstützen können, dieser Herausforderung zu begegnen. Deutlich ist aber auch geworden, dass KI-Technologien keine schnell verfügbaren und günstigen Lösungen bieten. KI-Systeme können bestimmte Aufgaben übernehmen und Lehrkräften unterstützend zur Seite stehen. Um ihr Potenzial zum Tragen zu bringen, sind jedoch etliche Voraussetzungen zu erfüllen und zeitliche sowie personelle Ressourcen bereitzustellen. Letztlich ist es nicht die KI selbst, die lernförderlich wirksam ist, sondern der durch Lehrkräfte methodisch und didaktisch sinnvoll gestaltete Einsatz der KI im Unterricht. Lehrkräfte dazu zu befähigen, KI gezielt und effektiv einzusetzen, sollte ein entscheidungsleitendes Ziel sein. So verstanden, kann KI dazu beitragen, Schüler:innen, insbesondere denjenigen mit Lernschwächen, mehr individuelle Lernunterstützung zukommen zu lassen, so dass perspektivisch Bildungsungleichheiten abgebaut und Chancengerechtigkeit gefördert werden.

[47] https://news.sky.com/story/uks-first-teacherless-ai-classroom-set-to-open-in-london-13200637

Referenzen

Anand, P. (2023, March 16). Khan Academy creates GPT-4 based helper Khanmigo marking formal entry of AI into education. Dataquest. https://www.dqindia.com/khan-academy-creates-gpt-4-based-helper-khanmigo-marking-formal-entry-of-ai-into-education/

Aleven, V., Mclaughlin, E., Glenn, A., & Koedinger, K. (2016). Instruction based on adaptive learning technologies. In: Mayer, R.E., Alexander, P. (eds.) *Handbook of Research on Learning and Instruction*, pp. 522–560. Routledge, New York.

Artificial Intelligence Act (Regulation (EU) 2024/1689), Official Journal version of 13 June 2024'. ELI: http://data.europa.eu/eli/reg/2024/1689/oj

Autorengruppe Bildungsberichterstattung (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen.* Bertelsmann Verlag: Bielefeld. https://doi.org/10.3278/6001820dw

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, Ray. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, *4*(2), 167–207. https://doi.org/10.1207/s15327809jls0402_2

Aslan, S., Alyuz, N., Tanriover, C., Mete, S. E., Okur, E., D'Mello, S. K., & Arslan Esme, A. (2019). Investigating the impact of a real-time, multimodal student engagement analytics technology in authentic classrooms. In Proceedings of the Conference on Human Factors in Computing Systems (CHI '19). Paper 304, 1–12. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3290605.3300534

Bäuerlein, K., Etter, B., Brunner, S., & Karlen, Y. (März, 2022). CleveR – Ein Online-Diagnosetool zur Kompetenzdiagnostik im selbstregulierten Lernen. Poster präsentiert an der Tagung der Gesellschaft für empirische Bildungsforschung (GEBF), Bamberg (Online)

Beal, C.R. (2013). AnimalWatch: An intelligent tutoring system for algebra readiness. In: Azevedo, R., Aleven, V. (eds) International Handbook of Metacognition and Learning Technologies. *Springer International Handbooks of Education*, vol 28. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-5546-3_22

Bereinigte Amtliche Sammlung der Schulvorschriften NRW (2013). Zu BASS 1 - 1. Erstes Gesetz zur Umsetzung der VN-Behindertenrechtskonvention in den Schulen (9. Schulrechtsänderungsgesetz) vom 5. November 2013. GV. NRW. S. 618. https://www.schulministerium.nrw/sites/default/files/documents/NeuntesSchulrechtsaenderungsgesetz.pdf

Berger, M. (2013). Vergleichende Darstellung der Sichtweisen von Akteuren_innen im Transitionsprozess Kita-Schule - zusammenfassender Bericht zu den Ergebnissen der schriftlichen Befragungen. Als Download unter http://ilea-t.reha.uni-halle.de

Börte, K., Lillejord, S., Chan, J., Wasson, B., & Greiff, S. (2023). Prerequisites for teachers' technology use in formative assessment practices. A systematic review. *Educational Research Review, 41*, 100568. https://doi.org/10.1016/j.edurev.2023.100568

Bundesverband Legasthenie & Dyskalkulie e.V. (o.D.) *Lernsoftware*. http://www.alpha-fundsachen.de/wp-content/uploads/Lernsoftware_Dyskalkulie_und_Legasthenie.pdf

Chine, D. R., Brentley, C., Thomas-Browne, C., Richey, J. E., Gul, A., Carvalho, P., Branstetter, L., & Koedinger, K. (2022). Educational equity through combined human-AI personalization: A propensity matching evaluation. In M. M. Rodrigo, N. Matsuda, A. Cristea, & V. Dimitrova (Eds.), *Artificial Intelligence in Education* (pp. 366–377). Springer International Publishing. https://doi.org/10.1007/978-3-031-11644-5_30

Collie, R. J., & Martin, A. J. (2024). Teachers' motivation and engagement to harness generative AI for teaching and learning: The role of contextual, occupational, and background factors. *Computers and Education: Artificial Intelligence, 6*, 100224. http://dx.doi.org/10.1016/j.caeai.2024.100224

Dai, W., Lin, J., Jin, F., Li, T., Tsai, Y.-S., Gašević, D. & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. [Preprint from EdArXiv]. https://doi.org/10.35542/osf.io/hcgzj

Datenschutz-Grundverordnung (DSGVO) (2016). Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016. Zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). Art. 22, 1. Automatisierte Entscheidungen im Einzelfall einschließlich Profiling. https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679

Denker, B., Horn, N. & Vallée, T. (2021): Datenschutz und digitale Schule. Impulse zur Entlastung und Unterstützung von Schulen. Herausgegeben vom Forum Bildung Digitalisierung.

Deutscher Bundestag (2020). Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale. Bundestagsdrucksache 19/2978. https://dserver.bundestag.de/btd/19/237/1923700.pdf

Deutscher Ethikrat (2023). Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Stellungnahme. https://www.ethikrat.org/publikationen/stellungnahmen/mensch-und-maschine/

DFKI – Deutsches Forschungszentrum für Künstliche Intelligenz (2023). Schule und KI – Ein praxisorientierter Leitfaden. telekom-stiftung.de/sites/default/files/files/ Leitfaden-Schule-und-KI.pdf

du Boulay, B., Mitrovic, A., & Yacef, K. (Eds.). (2023). *Handbook of Artificial Intelligence in Education*. Cheltenham, UK: Edward Elgar Publishing. https://doi.org/10.4337/9781800375413

Europäische Kommission (2019). Ethics Guidelines for trustworthy AI. High-Level Expert Group on Artificial Intelligence. https://www.europarl.europa.eu/cmsdata/196377/AI%20HLEG_Ethics%20Guidelines%20for%20Trustworthy%20AI.pdf

Europäische Kommission (2022). *Ethische Leitlinien für Lehrkräfte über die Nutzung von KI und Daten für Lehr- und Lernzwecke.* https://publikationen.kmk-pad.org/ethische-leitlinien-fur-lehrkrafte-uber-die-nutzung-von-ki-und-daten/67644257

Feng, S., Magana, A.J., & Kao, D. (2021). A Systematic Review of Literature on the Effectiveness of Intelligent Tutoring Systems in STEM. *2021 IEEE Frontiers in Education Conference (FIE)*, 1-9. https://doi.org/10.1109/fie49875.2021.9637240

Feng, M., Heffernan, N., Collins, K., Heffernan, C., Murphy, R.F. (2023). Implementing and evaluating ASSISTments online math homework support at large scale over two years: findings and lessons learned. In: Wang, N., Rebolledo-Mendez, G., Matsuda, N., Santos, O.C., Dimitrova, V. (eds) *Artificial Intelligence in Education*. AIED 2023. Lecture Notes in Computer Science, 13916. Springer, Cham. https://doi.org/10.1007/978-3-031-36272-9_3

Fobizz. (o. J.). Eine Schule macht's vor: Digitales Spielzimmer für kreatives Lernen, fobizz – Online-Fortbildungen für Lehrkräfte. Zugriff am 26.09.2024. https://fobizz.com/digitale-spielzimmer-fur-kreatives-lernen/

Fobizz (2023, 27. November). *Ein Jahr ChatGPT. KI in der Schule: So nutzen Lehrkräfte die Technologie im Unterricht*. [Pressemitteilung]. https://drive.google.com/file/d/1O4XlX1-ZlArQdsdY7V6JFeQXOudey-Loo/view

Future of Life Institute (2024). Zeitplan für die Umsetzung. https://artificialintelligenceact.eu/de/implementation-timeline/

Geiling, U., Liebers, K., Prengel, A. (2012). ILEA T: Individuelle Lern-Entwicklungs-Analyse im Übergang/Transition – ein Beitrag für einen inklusionsorientierten Übergang von der Kita zur Grundschule?. In: Hellmich, F., Förster, S., Hoya, F. (eds) *Bedingungen des Lehrens und Lernens in der Grundschule. Jahrbuch Grundschulforschung*, (vol 16, S. 119–122). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-19137-9_18

Goldhammer, F., Scherer, R., & Greiff, S. (2019). Advancements in technology-based assessment. Emerging item formats, test designs, and data sources. *Frontiers in Psychology. 10.* 3047. https://doi.org/10.3389/978-2-88963-506-1

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education, 203*, 104862. https://doi.org/10.1016/j.compedu.2023.104862

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heffernan, N.T., Heffernan, C.L. (2014). The ASSISTments Ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24,* 470-497. https://doi.org/10.1007/s40593-014-0024-x

Hirunyasiri, D., Thomas,D. R. , Lin, J. ,Koedinger, K. R. & Aleven, V. (2023). Comparative analysis of GPT-4 and human graders in evaluating praise given to students in synthetic dialogues. CoRR, abs/2307.02018, 2023.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Holstein, K., & Aleven, V. (2022). Designing for human–AI complementarity in K-12 education. *AI Magazine 43*, 239–48. https://doi.org/10.1002/aaai.12058

Holstein, K., McLaren, B.M., Aleven, V. (2018). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. Proceedings of Artificial Intelligence in Education AIED 2018. *Lecture Notes in Computer Science*, vol 10947. Springer, Cham. https://doi.org/10.1007/978-3-319-93843-1_12

Huang, X., Craig, S. D., Xie, J., Graesser, A., & Hu, X. (2016). Intelligent tutoring systems work as a math gap reducer in 6th grade after-school program. *Learning and Individual Differences, 47*, 258-265. https://doi.org/10.1016/j.lindif.2016.01.012

Jenni, O. (2024). Herausforderungen beim Lernen – Lese-, Rechtschreib-, Rechenstörungen. In: *Entwicklungsstörungen verstehen*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-69223-3_11

Jansen, T., Meyer, J., Fleckenstein, J., Horbach, A., Keller, S., & Möller, J. (2024). Individualizing goal-setting interventions using automated writing evaluation to support secondary school students' text revisions. *Learning and Instruction, 89*, 101847. https://doi.org/10.1016/j.learninstruc.2023.101847.

Jung, P.-G., McMaster, K. L., Kunkel, A. K., Shin, J., & Stecker, P. M. (2018). Effects of data-based individualization for students with intensive learning needs: a meta-analysis. *Learning Disabilities Research & Practice, 33*, 144–155. https://doi.org/10.1111/ldrp.12172, 3.

Käser, T., Busetto, A. G., Baschera, G.-M., Kohn, J., Kucian, K., Aster, M., et al. (2012). Modelling and optimizing the process of learning mathematics. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds) *Intelligent Tutoring Systems. ITS 2012. Lecture Notes in Computer Science*, 7315. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-30950-2_50

Käser, T., Baschera, G., Kohn, J., Kucian, K., Richtmann, V., Grond, U., Gross, M. & Von Aster, M. (2013). Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00489

Kast, M., Baschera, G. M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia 2011, 61*(2), 177–200. https://doi.org/10.1007/S11881-011-0052-2

Kizilcec, R. F. (2024). To advance AI use in education, focus on understanding educators. *International Journal of Artificial Intelligence in Education. 34*, 12-19 https://doi.org/10.1007/s40593-023-00351-4

Klemm, K., Hollenbach-Biele, N. & Lepper, C. (2023). *Inklusion im deutschen Schulsystem Schuljahr 2021/2022. Factsheet.* Bertelsmann Stiftung: Gütersloh. https://doi.org/10.11586/2023050

Koedinger, K., McLaughlin, E., & Heffernan, N. (2010). A quasi-experimental evaluation of an online formative assessment and tutoring system. *Journal of Educational Computing Research Baywood Publishing, 4*, 489–510. https://doi.org/10.2190/EC.43.4.d

Kohn, J., Rauscher, L., Kucian, K., Käser, T., Wyschkon, A., Esser, G., & von Aster, M. (2020). Efficacy of a computer-based learning program in children with developmental dyscalculia. What influences individual responsiveness?. *Frontiers in Psychology, 11*, 1115. https://doi.org/10.3389/fpsyg.2020.01115

Konerding, M., Bergström, K., Lachmann, T. & Klatte, M. (2020). Lesen und Schreiben lernen mit Lautarium: Ein computerbasiertes Trainingsprogramm für Kinder mit Lese-Rechtschreibschwierigkeiten. *Praxis Schulpsychologie, 24*, 13.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., Gälli, M., Martin, E. & von Aster M (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, 57(3), 782-795. https://doi.org/10.1016/j.neuroimage.2011.01.070

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Kultusministerkonferenz (2004). Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/Beschluesse_Veroeffentlichungen/allg_Schulwesen/standards_lehrerbildung.pdf

Kultusmisterkonferenz (2010). Pädagogische und rechtliche Aspekte der Umsetzung des Übereinkommens der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (Behindertenrechtskonvention - VN-BRK) in der schulischen Bildung. Beschluss vom 18.11.2010. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2010/2010_11_18-Behindertenrechtkonvention.pdf

Kultusministerkonferenz (2011). Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen. Beschluss der Kultusministerkonferenz vom 20.10.2011. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf

Kultusministerkonferenz (2016). Bildung in der digitalen Welt - Strategie der Kultusministerkonferenz. Beschluss der Kultusministerkonferenz vom 08.12.2016 in der Fassung vom 07.12.2017. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf

Kultusministerkonferenz (2019). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz* vom 16.12.2004 i. d. F. vom 16.05.2019. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf

Kultusministerkonferenz (2021). Lehren und Lernen in der digitalen Welt. *Ergänzung zur Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“*. Beschluss der Kultusministerkonferenz vom 09.12.2021. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2021/2021_12_09-Lehren-und-Lernen-Digi.pdf

Kultusministerkonferenz (2024). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. Beschluss der Kultusministerkonferenz vom 16.10.2008 i. d. F. vom 08.02.2024. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2008/2008_10_6-Fachprofile-Lehrerbildung.pdf

Landesbeauftragte für Datenschutz und Informationsfreiheit Nordrhein-Westfalen (LDI NRW) (25.07.2024). *KI-Verordnung mit neuen Aufgaben.* https://www.ldi.nrw.de/ki-verordnung-mit-neuen-aufgaben

Lawrence, L, Guo, B., Yang, K. B., Echeverria, V., Kang, Z., Bathala, V., Li, C., Huang, W., Aleven, V., & Rummel, N. (2022). Co-designing AI-based orchestration tools to support dynamic transitions: Design narratives through conjecture mapping. *Proceedings of International Conference on Computer-Supported Collaborative Learning 2022* (pp. 139-146). International Society of the Learning Sciences.

Leifheit, L., Loefflad, D., Belschner, S., Beuttler, B., Winkelmann, J., Meurers, W. D., & Holz, H. (2024). KI im Unterricht: Entwicklung von Lehrveranstaltungen für Lehramtsstudierende der Sprach- und MINT-Fächer. *Ludwigsburger Beiträge Zur Medienpädagogik*, *24*, 1–19. https://doi.org/10.21240/lbzm/24/08

Lenhard, A., Lenhard, W., Schug, M. & Kowalski, A. (2011). Computerbasierte Mathematikförderung mit den „Rechenspielen mit Elfe und Mathis I" - Vorstellung und Evaluation eines Computerprogramms für Erst- bis Drittklässler. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 43* (2), 79-88. https://doi.org/10.1026/0049-8637/a000037

Lenhard, W., Schneider, W., Lenhard, A., & Schneider, W. (2018). ELFE II: ein Leseverständnistest für Erst-bis Siebtklässler-Version II. Göttingen: Hogrefe.

Leutner, D. (2002). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In L. J. Issing & P. Klimsa (Hrsg.), Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxis (3., vollst. überarb. Aufl, S. 115–125). Beltz PVU.

Ma, W., Adesope, O. O., Nesbit, J. C. & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal Of Educational Psychology*, *106*(4), 901–918. https://doi.org/10.1037/a0037123

Meurers, D., De Kuthy, K., Möller, V., Nuxoll, F., Rudzewitz, B. & Ziai, R. (2018). Digitale Differenzierung benötigt Informationen zu Sprache, Aufgabe und Lerner. Zur Generierung von individuellem Feedback in einem interaktiven Arbeitsheft. *FLuL – Fremdsprachen Lehren und Lernen, 47*(2), 64–82.

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B. & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics. 39.* 161-188. https://dx.doi.org/10.1017/S0267190519000126

Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence, 6*, 100199. https://doi.org/10.1016/j.caeai.2023.100199

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2017a). Arbeitshilfe: Gewährung von Nachteilsausgleichen für Schülerinnen und Schüler mit Behinderungen, Bedarf an sonderpädagogischer Unterstützung und/oder besonderen Auffälligkeiten in der Primarstufe – Eine Orientierungshilfe für Schulleitungen. https://www.schulministerium.nrw/sites/default/files/documents/1-Arbeitshilfe_Primarstufe.pdf

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2017b). *Arbeitshilfe: Gewährung von Nachteilsausgleichen für Schülerinnen und Schüler mit Behinderungen, Bedarf an sonderpädagogischer Unterstützung und/oder besonderen Auffälligkeiten für die Gymnasiale Oberstufe sowie für die Abiturprüfung – Eine Orientierungshilfe für Schulleitungen.* https://www.schulministerium.nrw/sites/default/files/documents/3-Arbeitshilfe_GymnasialeOberstufe-und-Abiturpruefung.pdf

Ministeriums für Schule und Bildung Nordrhein-Westfalen (MSB NRW) (2023). *Umgang mit textgenerierenden KI-Systemen. Ein Handlungsleitfaden.* https://www.schulministerium.nrw/system/files/media/document/file/handlungsleitfaden_ki_msb_nrw_230223.pdf

Ministeriums für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2024, 07/08) *Lehren und Lernen in der digitalen Welt: Webseite unterstützt Lehrkräfte – jetzt auch zu KI*. https://www.schulministerium.nrw/lehren-und-lernen-der-digitalen-welt-webseite-unterstuetzt-lehrkraefte-jetzt-auch-zu-ki

Ministeriums für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2024a). *Datenschutz im Schulbereich.* https://www.schulministerium.nrw/schule-bildung/recht/datenschutz-im-schulbereich

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2024b). *Inklusion - Recht*. Schulministerium.nrw. https://www.schulministerium.nrw/inklusion-recht

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (MSB NRW) (2024c). *Ministerin Feller: Wir fördern das Lesen, Schreiben und Rechnen konsequent weiter*. Pressemitteilung vom 15.08.2024. https://www.schulministerium.nrw/presse/pressemitteilungen/ministerin-feller-wir-foerdern-das-lesen-schreiben-und-rechnen-konsequent

Molenaar, I., & Knoop-van Campen, C. (2016). Learning analytics in practice: The effects of adaptive educational technology Snappet on students' arithmetic skills. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. Association for Computing Machinery, New York, NY, 538–539. https://doi.org/10.1145/2883851.2883892

Myneni, L. S., Narayanan, N. H., Rebello, S., Rouinfar, A. & Pumtambekar, S. (2013). An interactive and intelligent learning system for physics education. *IEEE Transactions On Learning Technologies*, *6*(3), 228–239. https://doi.org/10.1109/tlt.2013.26

Pane, J. F., Griffin, B. A., McCaffrey, D. F. & Karam, R. (2014). Effectiveness of cognitive Tutor Algebra I at scale. *Educational Evaluation And Policy Analysis*, *36*(2), 127–144. https://doi.org/10.3102/0162373713507480

Peters, R. & Bovenschulte, M. (2020). Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (Hrsg.) *Learning Analytics – Potenzial von KI-Systemen für Lehrende und Lernende.* Themenkurzprofil Nr. 42. https://www.bundestag.de/resource/blob/845926/d5a0554dd04097c1288776f10fc85e00/Themenkurzprofil-042.pdf

Pfeiffer, W. & Dögl, R. (1999). Das Technologie-Portfolio-Konzept zur Beherrschung der Schnittstelle Technik und Unternehmensstrategie. In D. Hahn & B. Taylor (Hrsg.). *Strategische Unternehmungsplanung. Strategische Unternehmungsführung* (S. 440-467). Heidelberg: Physica.

Scharnagl, S., Evanschitzky, P., Streb, J., Spitzer, M. & Hille, K. (2014). Sixth Graders benefit from educational software when learning about fractions. *A controlled classroom study. Numeracy, 7* (1), article 4. DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4

Scheuer, O., & McLaren, B. M. (2012). Educational data mining. *Encyclopedia of the Sciences of Learning* (pp. 1075–79). New York: Springer US.https://doi.org/10.1007/978-1-4419-1428-6_618

Schmitterer, A.M., Tetzlaff, L., Hasselhorn, M., & Brod, G. (2023). Who benefits from computerized learning progress assessment in reading education? Evidence from a two-cohort pre–post design. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12851

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2024). *Sonderpädagogische Förderung in Schulen 2013 bis 2022. Statistische Veröffentlichungen der Kultusministerkonferenz*. Dokumentation Nr. 240 – Februar 2024. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_240_SoPae_2022.pdf

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*, 31-40.

Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK) (2023): *Large Language Models und ihre Potenziale im Bildungssystem.* Impulspapier der Ständigen Wissenschaftlichen Kommission (SWK) der Kultusministerkonferenz. http://dx.doi.org/10.25656/01:28303

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology, 105*(4), 970–987. https://doi.org/10.1037/a0032447

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology, 106*(2), 331–347. https://doi.org/10.1037/a0034752

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, *91*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Taskforce „Künstliche Intelligenz im Bildungswesen" (2024). *Empfehlungen der Taskforce „Künstliche Intelligenz im Bildungswesen" für Schule, Hochschule und Weiterbildung.* https://www.cais-research.de/wp-content/uploads/Taskforce-KI-im-Bildungswesen-Empfehlungen-fuer-NRW.pdf

Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing personalized education: A dynamic framework. *Educational Psychology Review, 33*(3), 863-882. https://doi.org/10.1007/s10648-020-09570-w

Thomas, D.R., Gatz, E., Gupta, S., Aleven, V., Koedinger, K.R. (2024). The Neglected 15%: Positive effects of hybrid human-AI tutoring among students with disabilities. In: Olney, A.M., Chounta, IA., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds) Artificial Intelligence in Education. AIED 2024. Lecture Notes in Computer Science, vol 14829. Springer, Cham. https://doi.org/10.1007/978-3-031-64302-6_29

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Varsik, S. & L. Vosberg (2024), The potential impact of artificial intelligence on equity and inclusion in education, *OECD Artificial Intelligence Papers*, No. 23, OECD Publishing, Paris, https://doi.org/10.1787/15df715b-en

Verordnung (EU) (2024). Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) (Text von Bedeutung für den EWR). ELI: http://data.europa.eu/eli/reg/2024/1689/oj

Vock, M. & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht.* Schriftenreihe des Netzwerk Bildung. Friedrich Ebert Stiftung.

Von Aster, M. G. & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49*(11), 868–873. https://doi.org/10.1111/j.1469-8749.2007.00868.x

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. British *Journal of Educational Technology. 50*, 3119-3137. https://doi.org/10.1111/bjet.12758

Yang, K. B., Lu, Z., Echeverria, V., & Sewall, J., Lawrence, L., Rummel, N., & Aleven, V. (2022). Technology ecosystem for orchestrating dynamic transitions between individual and collaborative AI-tutored problem solving. *Proceedings of International Conference on Artificial Intelligence in Education (AIED' 22)* (pp. 673-678).

Yang, K.B., Echeverria, V., Lu, Z., Mao, H., Holstein,K., Rummel, N., Aleven, V. (2023). Pair-Up: Prototyping human-AI co-orchestration of dynamic transitions between individual and collaborative learning in the classroom. *Proceedings of the 2023 CHI conference on human factors in computing systems (CHI' 23)*.

Zheng, S. (2024). The effects of chatbot use on foreign language reading anxiety and reading performance among Chinese secondary school students. *Computers and Education: Artificial Intelligence, 7*, 100271. https://doi.org/10.1016/j.caeai.2024.100271

Anhang 1

Leitfaden zum Prüfen von Systemen hinsichtlich ihrer Adaptivität in Anlehnung an Aleven et al. (2016).

Die wichtigsten Kriterien für eine Einordnung als Intelligentes Tutoring-System sind:

1. Das System hilft den Schüler:innen, komplexe Aufgaben zu üben. Typischerweise sind dies mehrschrittige Aufgaben mit potenziell mehreren verschiedenen Lösungswegen.
2. Das System bietet eine individuelle, schrittweise adaptive Hilfestellung innerhalb dieser Aufgaben mit Hinweisen und Rückmeldungen, die dem Lösungsweg der Schüler:innen folgen und daran anpassen (*step-loop adaptivity*).
3. Das System unterstützt das individuelle Lernen, indem es den Wissenszuwachs jedes Lernenden verfolgt und Aufgaben auswählt, die noch zu übende Wissenskomponenten beinhalten (*task-loop adaptivity*). Alle Schüler:innen erhalten so das richtige Maß an Übung (*mastery learning*).

Die folgenden Fragen können dabei helfen zu beurteilen, inwieweit ein bestimmtes System diese Eigenschaften aufweist:

Leitfragen	✅	❌
Verwendet das System ein detailliertes Wissensmodell, um den Erwerb des zu lernenden Inhalts durch die Schüler:innen zu verfolgen?	☐	☐
Gibt es Beispiele für mehrstufige Aufgaben, die Schüler:innen mit dem System lösen?	☐	☐
Gibt es Beispiele für alternative Lösungen zu einem oder mehreren dieser mehrstufigen Aufgaben?	☐	☐
Bietet das System schrittweise Hilfestellungen in Form von Hinweisen und Rückmeldungen?	☐	☐
Sind die schrittweisen Hilfestellungen adaptiv? D.h., passen sie sich flexibel an die verschiedenen möglichen Lösungen und Schrittfolgen von Schüler:innen an?	☐	☐
Wie weist das System nachfolgende Aufgaben zu und passt sich die Aufgabenauswahl dem aktuellen Kenntnisstand der Lernenden an?	☐	☐
Verfolgt das System den Wissenszuwachs einzelner Schüler:innen über die Zeit (d. h. über aufeinanderfolgende Aufgaben hinweg)? Ist diese Information der Lehrkraft den individuellen Schüler:innen zugänglich?	☐	☐
Ist transparent, welche Methode bzw. welcher Algorithmus für die Lernstandsverfolgung verwendet wird?	☐	☐
Werden Kriterien verwendet, um festzustellen, ob Lernende die in einer Lerneinheit angestrebten Kenntnisse oder Fähigkeiten beherrschen? Sind diese der Lehrkraft und dem individuellen Lernenden zugänglich?	☐	☐

Darüber hinaus ist relevant zu fragen, welche wissenschaftlichen Studien durchgeführt wurden, um das System zu evaluieren sowie ob diese Studien zeigen konnten, dass das System lernförderlich ist. Folgende Leitfragen können helfen, Studien einzuordnen:

Leitfragen
Wie lange haben Schüler:innen im Rahmen der Studie mit dem System gearbeitet und wie häufig wurde es eingesetzt? *(Idealerweise sollte das System über mehrere Schulstunden oder ein ganzes Schuljahr verwendet worden sein)*
Wie viele Schüler:innen haben an der Studie teilgenommen? *(Idealerweise sollten dies mehrere Hundert oder sogar Tausend sein)*
Gab es eine Kontrollbedingung? Wenn ja, wie haben die Schüler:innen in der Kontrollbedingung gelernt? *(Idealerweise wird eine starke Kontrollbedingung verwendet, z. B. der reguläre Unterricht)*
Wie wurden die Schüler:innen den Bedingungen zugewiesen? *(Idealerweise nach dem Zufallsprinzip)*
Wie wurden Lernergebnisse oder Lernzuwächse der Schüler:innen gemessen? *(Idealerweise sowohl mit einem standardisierten Instrument, als auch mit einem von Wissenschaftler:innen entwickelten Instrument)*
Wurden die Lernzuwächse zwischen den Bedingungen verglichen, um die statistische Signifikanz der festgestellten Unterschiede zu ermitteln?
Wurden „Dosierungseffekte“ gemessen, d. h. haben die Wissenschaftler:innen untersucht, ob eine häufigere Nutzung des Systems mit einem größeren Lernzuwachs verbunden war?
Wurden die Ergebnisse der Studie in einer wissenschaftlichen Konferenz oder Zeitschrift mit Fachgutachtersystem veröffentlicht?

Je mehr dieser Fragen beantwortet werden können, umso fundiertere Aussagen lassen sich zu dem System treffen und umso besser ist es beurteilbar.