

KI im Unterricht – Entwicklung von Lehrveranstaltungen für Lehramtsstudierende der Sprach- und MINT-Fächer

Luzia Leifheit, Denise Loefflad*, Sina Belschner, Benedikt Beuttler, Jan Winkelmann, Detmar Meurers, Heiko Holz*

Zusammenfassung

In einer Zeit des rasanten technologischen Fortschritts und der digitalen Transformation wird die Notwendigkeit, Künstliche Intelligenz (KI) in den Unterricht zu integrieren, immer deutlicher. Lehrkräfte spielen dabei eine entscheidende Rolle, da sie im Kontext ihrer jeweiligen Fachbereiche in der Lage sein müssen, KI-Tools kritisch zu reflektieren und fachdidaktisch fundiert anzuwenden. Dazu entwickeln wir im Projekt „WoLKE“ fachlich und didaktisch passgenaue Lehr-Lernformate für den reflektierten Einsatz von KI-Tools in der Bildung, um bereits angehende Lehrkräfte systematisch auf die Herausforderungen und Chancen des didaktisch motivierten Einsatzes von KI-Tools vorzubereiten. Dies geschieht auf Basis einer computerlinguistisch sowie fachdidaktisch fundierten Analyse sowohl etablierter als auch aktueller KI-Methoden, wie z. B. textgenerierender KI. Ziel des Projekts ist es, Lehramtsstudierende mit den notwendigen Kenntnissen und Kompetenzen auszustatten, um KI-Tools verantwortungsvoll in die pädagogische Praxis zu integrieren. Der Fokus liegt dabei auf der Rolle von KI-Tools im Sprach- und MINT-Unterricht und umfasst sowohl die fachspezifischen Anforderungen an KI-Tools als auch ein methodisch und ethisch fundiertes Verständnis der Potenziale und Grenzen von KI-Methoden. Das Projekt verfolgt einen ko-konstruktiven Ansatz in einem interdisziplinären Verbund aus Computerlinguistik, Informatik, Sprach- und MINT-Didaktik. Darin werden etablierte und aktuelle textgenerierende KI-Methoden hinsichtlich Eignung, Reifegrad und Synergiepotenzial für den Einsatz im Fachunterricht analysiert. Die Konzeption der Lehrveranstaltungen folgt dem Ansatz des Design-based Research und wird mit qualitätssichernden Methoden der empirischen Bildungsforschung evaluiert. Begleitende Transferworkshops ermöglichen die Dissemination der Lehrveranstaltungen an andere Hochschulen. Das Projekt adressiert bestehende Desiderate, indem es die Frage aufgreift, wie Methoden der KI effektiv zu einer pädagogisch und fachdidaktisch motivierten, individuellen Förderung beitragen können. Das Projekt wird gefördert aus Mitteln des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg und befindet sich derzeit in der ersten Konzeptionsphase. Im Folgenden stellen wir das Konzept für die ko-konstruktive Entwicklung der Lehrveranstaltungen sowie den Plan für die empirische Evaluation vor.

Schlüsselbegriffe: ● KI-Methoden ● Lehr-Lernformate ● Sprachdidaktik ● MINT-Didaktik ● Informatik ● Computerlinguistik

Einleitung

In den vergangenen Jahren hat sich die Domäne der Künstlichen Intelligenz (KI), besonders seit der Veröffentlichung des textgenerierenden KI-Tools „ChatGPT“ der Firma OpenAI im November 2022 sowohl im akademischen Diskurs als auch in der öffentlichen Wahrnehmung signifikant verankert. Dieser Trend führte zu einer verstärkten Fokussierung auf die Nutzung von KI in verschiedenen Lebensbereichen wie der Didaktik (vgl. Chen/Chen/Lin 2020), Medizin (vgl. Rajpurkar et al. 2022), Wirtschaft (vgl. Aghion/Jones/Jones 2017), in Umwelttechnologien (vgl. Nishant/Kennedy/Corbett 2020) oder in kreativen Feldern (vgl. Anantrasirichai/Bull 2022). KI hat sich somit zu einer festen Größe entwickelt, die über die Grenzen der Informatik und diverser Forschungsbereiche hinaus im öffentlichen Raum rege diskutiert wird und starken Einfluss auf unseren Alltag hat.

Vor diesem Hintergrund kommt Lehrkräften eine entscheidende Rolle zu, da sie in der Lage sein müssen, KI-Tools kritisch zu reflektieren und fachdidaktisch fundiert einzusetzen. Das Projekt „WoLKE“ adressiert diese Notwendigkeit, indem es passgenaue Lehrveranstaltungen zum reflektierten Einsatz von KI-Tools für Lehramtsstudierende entwickelt und empirisch evaluiert. Dazu vereint das interdisziplinäre Projekt computerlinguistische und fachdidaktische Perspektiven, um die angehenden Lehrkräfte systematisch auf die Herausforderungen und Chancen des didaktisch motivierten Einsatzes von KI-Tools vorzubereiten.

Dieser Beitrag stellt das Konzept für die ko-konstruktive Entwicklung der Lehrveranstaltungen sowie den aktuellen Stand der Planung für die empirische Evaluation vor. Dabei wird der Ansatz des Design-based Research verfolgt und die Qualität der Lehrveranstaltungen mit Methoden der empirischen Bildungsforschung in mehreren aufeinanderfolgenden und ineinandergreifenden Zyklen zunächst pilotiert und anschließend je dreimal vollständig durchgeführt, evaluiert und überarbeitet. Das Projekt adressiert damit bestehende Desiderate und trägt zur fachdidaktischen Theorieentwicklung bezüglich der effektiven und kritisch reflektierten Integration von KI in die pädagogische Praxis bei. Im Folgenden wird ein Überblick über die Abschnitte des Beitrags gegeben.

Abschnitt 1. *Stand der Forschung und theoretischer Hintergrund*, insbesondere Abschnitt 1.1 *Definitionen von Künstlicher Intelligenz*, diskutiert verschiedene Aspekte des Konzepts der Künstlichen Intelligenz, einschließlich Definitionen, Anwendungen wie Maschinelles Lernen und selbstlernende Systeme sowie deren Bedeutung für das Natural Language Processing (NLP). Die potenzielle Bedeutung dieser Technologien für Bildungsumgebungen wird aufgrund ihrer Fähigkeit, komplexe sprachliche Interaktionen zu analysieren und darauf zu reagieren, hervorgehoben.

In Abschnitt 1.2 *KI in der Bildung* werden die vielfältigen Anwendungsmöglichkeiten von KI-Technologien im Bildungskontext beleuchtet, insbesondere intelligente tutorielle Systeme (ITS) und Large Language Models (LLMs). Das Potenzial von LLMs für die Erstellung adaptiver Lernumgebungen wird diskutiert, aber auch die damit verbundenen Herausforderungen, wie Vorurteile in Trainingsdaten und fehlendes Weltwissen. Darüber hinaus werden traditionelle KI-Methoden wie ITS als effektive Unterstützung für individualisiertes Lernen vorgestellt.

Abschnitt 1.2.1 *KI in der Sprachdidaktik* zeigt die vielfältigen Anwendungsmöglichkeiten von KI-Technologien wie ITS und textgenerierender KI in der Sprachdidaktik auf. Darin betont wird die Bedeutung dieser Technologien für die individualisierte Förderung von Lernenden, die Bereitstellung adaptiver Lernumgebungen und die spielerische Förderung des Schriftspracherwerbs. Es wird darin aber auch auf mögliche Problembereiche hingewiesen, insbesondere im Hinblick auf die Anpassung an individuelle Lernkontexte und die Vermeidung sprachlicher Überforderung.

In Abschnitt 1.2.2 *KI in der MINT-Didaktik* werden verschiedene Ansätze zur Integration von KI in die MINT-Lehramtsausbildung skizziert und die Bedeutung fachspezifischer digitaler Kompetenzen betont. Dabei werden sowohl die vielfältigen Einsatzmöglichkeiten traditioneller digitaler Werkzeuge im naturwissenschaftlichen Unterricht als auch bestehende Desiderate für die Weiterentwicklung von KI-Tools zur Berücksichtigung naturwissenschaftlicher Gütekriterien thematisiert.

Abschnitt 2. *Das Projekt „WoLKE“* beleuchtet die Projektziele, die Zusammensetzung des Verbunds, die Inhalte und Evaluation der geplanten Lehrveranstaltungen sowie den zeitlichen Horizont des Projekts.

Das Projekt zielt darauf ab, angehende Lehrkräfte für einen reflektierten Umgang mit KI-basierten Werkzeugen im Unterricht vorzubereiten, insbesondere in den Bereichen Sprach- und MINT-Didaktik. Durch die Entwicklung und Evaluation von Lehrveranstaltungen sowie die Bereitstellung von Open Educational Resources (OER) sollen methodische Kompetenzen vermittelt werden, um KI-Methoden zur Bewältigung schulischer Herausforderungen einzusetzen.

1. Stand der Forschung und theoretischer Hintergrund

1.1 Definitionen von Künstlicher Intelligenz

Der Begriff ‚Künstliche Intelligenz‘ besitzt bislang keine universell anerkannte Definition und wird in diversen Kontexten unterschiedlich ausgelegt. Wir verstehen unter dem Begriff KI ein weites, in mehrere Bereiche aufgeteiltes Forschungs- und Aufgabenfeld. Der übergeordnete Begriff der KI, der auch als funktionale Interpretation von KI bezeichnet werden kann, beschreibt den Versuch, menschliches Lernen und Denken auf den Computer zu übertragen und somit die Imitation menschlicher Intelligenz durch Computerprogramme. Kurz zusammengefasst zielt KI darauf ab, den Computer dazu zu befähigen, Aufgaben zu lösen, die – wenn sie vom Menschen gelöst werden würden – Intelligenz erfordern. Aus technischer Perspektive werden KI-Systeme häufig als (selbst)lernende Systeme charakterisiert, die entweder auf der Basis fester Regeln oder auf Grundlage von Datenpunkten lernen, um Prognosen für neue Datensätze zu erstellen (vgl. Goodfellow/Bengio/Courville 2016). Letzteres fällt unter den Begriff des Maschinellen Lernens und beschreibt die Teilmenge von KI, in der Algorithmen und Computer trainiert werden, selbstständig aus Daten und Erfahrungen zu lernen und sich zu verbessern, um Entscheidungen zu treffen und Prognosen zu stellen, anstatt explizit und regelbasiert dafür programmiert zu werden. In der Praxis sind hybride Ansätze, die sowohl feste Regeln als auch Trainingsdaten integrieren, weit verbreitet. Seit 2018 gewinnen selbstüberwachte Lernsysteme an Bedeutung, die durch den Einsatz umfangreicher Datensätze und einer Vielzahl von Parametern ihr eigenes Lernen kontinuierlich optimieren. Solche Systeme finden

Anwendung in aktuellen Sprachgenerierungsmodellen, die Texte erzeugen, indem sie fortlaufend das wahrscheinlichste Folgewort in einer Wortsequenz prognostizieren. Statt Prognosen über neue Daten zu treffen, generieren diese Modelle eigenständig neue Datenpunkte, in diesem Fall neue Texte.

Diese Systeme basieren auf Natural Language Processing (NLP), einem Zweig der Informatik und der KI-Forschung, bei welchem das regelbasierte Modellieren von Sprache mit statistischen Modellen aus dem Maschinellen Lernen kombiniert wird, um automatisiert Sprachdaten zu verarbeiten. So entstehen computerlinguistische Methoden zur Analyse, Verarbeitung und Generierung von Sprache. Diese Fortschritte in NLP und KI ermöglichen es, komplexe sprachliche Interaktionen zu analysieren und auf sie zu reagieren, was sie besonders nützlich in Bildungsumgebungen macht.

1.2 KI in der Bildung

Im Bildungskontext ermöglicht der Einsatz von KI-Technologien eine Optimierung etablierter Lehr- und Lernmethoden sowie eine effektivere Gestaltung des Unterrichts. KI kann als Instrument eingesetzt werden, um die Adaptivität im Lehrkontext zu erhöhen, eine individualisierte Förderung in heterogenen Lerngruppen zu unterstützen und Lehrkräfte in ihrer Arbeit zu entlasten (vgl. Colling/Heck/Meurers 2023). Intelligente Tutorssysteme (ITS), die auf KI- und NLP-Technologien basieren, können das Verhalten und die Bedürfnisse der Lernenden modellieren (vgl. Rudzewitz et al. 2019), Lernangebote individualisieren (vgl. Heck/Meurers/Nuxoll 2022) und zeitnahes Feedback auf Lernaktivitäten bereitstellen (vgl. Parrisius et al. 2022). Dies trägt zur Förderung des Lernfortschritts unabhängig von den individuellen Voraussetzungen der Lernenden bei und ermöglicht Lehrkräften eine effizientere Nutzung ihrer Unterrichtszeit. Solche ITS setzen oft hybride Systeme ein, die sowohl regelbasierte als auch datengetriebene Ansätze integrieren, was ihnen eine hohe Flexibilität und Anpassungsfähigkeit an unterschiedliche Lernsituationen verleiht. Im Folgenden wird ein Überblick über den Forschungsstand sowohl zu traditionellen als auch neueren KI-Tools und den Möglichkeiten sowie Grenzen ihres Einsatzes in den Sprach- und MINT-Didaktiken präsentiert.

Potenzial und Grenzen von Large Language Models (LLMs)

Textgenerierende KI-Tools wie ChatGPT basieren auf Large Language Models (LLMs), sogenannten großen Sprachmodellen, auf deren Basis sie sprachlich attraktive, aber inhaltlich nicht weiter in Welt- oder Fachwissen verankerte, letztlich fiktionale Texte erzeugen können. Damit können zum Beispiel schnell Texte zu unterschiedlichen Themen und Antworten auf Fragen generiert oder Dialoge unterstützt werden. Da die von LLMs generierten Texte allerdings immer eine Aneinanderreihung von wahrscheinlichen, aber nicht notwendigerweise inhaltlich korrekten Wortfolgen sind, benötigt ihr Einsatz grundsätzlich einen reflektierten Umgang und entsprechende Sachkenntnis, die Lehramtsstudierende für einen professionellen Umgang mit (text-)generierender KI entwickeln sollten.

LLMs bieten das Potenzial, Lehr- und Lerninhalte zu erstellen (vgl. Leiker et al. 2023), Lehrpläne zu entwerfen (vgl. Van Brummelen/Lin 2020) und durch individuell zugeschnittenes Feedback (vgl. Kasneci et al. 2023) adaptive Lernumgebungen zu schaffen, wodurch sie eine flexible Anpassung an die vielfältigen Bedürfnisse im Bildungswesen ermöglichen. So können

KI-generierte Texte auf das Kompetenzniveau der Lernenden und ihre persönlichen Interessen zugeschnitten werden, um die individuelle Motivation zu steigern. Dies kann sowohl für die Sprach- als auch für MINT-Didaktik genutzt werden: Im Mathematikunterricht können Aufgaben in der Erst- und Zweitsprache generiert werden, während im Sprachunterricht Aufgaben zum kreativen Schreiben erstellt werden können. Textgenerierende KIs können auch im Sinne eines Lerntandems genutzt werden, welche die Lernenden dabei unterstützen, komplexe Konzepte zu verstehen, oder gezielt Feedback zu einzelnen Aufgaben geben. In einem nächsten Schritt ermöglichen Large Multimodal Foundation Models (LMFMs), Input und Output nicht nur durch Eingabe textbasierter Prompts zu geben bzw. zu generieren, sondern auch via gesprochenem Text, Bildern oder Videos. Vor- und Nachteile verhalten sich ähnlich zu denen von LLMs, können aber jeweils unterschiedlich stark ausgeprägt sein (vgl. Küchemann et al. 2024).

Die Integration dieser Modelle in den Bildungsbereich bringt jedoch nicht nur Chancen, sondern auch Herausforderungen mit sich. Ein zentrales Problem, das in diesem Zusammenhang adressiert werden muss, sind die in den Trainingsdaten verankerten Vorurteile und Biases (vgl. Ferrara 2023). Da LLMs auf Daten trainiert werden, die existierende Vorurteile widerspiegeln, können diese Ungleichheiten auch in den generierten Inhalten reproduziert werden. Dies kann besonders problematisch sein, wenn Lernende die KI nicht im Rahmen einer Unterrichtseinheit nutzen, da die generierten Texte nicht von den Lehrkräften überprüft werden können. Da LLMs außerdem nicht über Weltwissen verfügen, sondern Texte nur basierend auf probabilistischen und nicht deterministischen Methoden generieren, kann es vorkommen, dass generierte Texte fachliche Fehlvorstellungen enthalten (vgl. Kasneci et al. 2023). Es ist daher wichtig, dass sowohl Lehrende als auch Lernende Urteilsvermögen im Umgang mit Fehlinformationen sowie kritisches Denken entwickeln – insbesondere die kritische Reflexion der generierten Texte.

Angesichts der zunehmenden Verbreitung und den potenziellen Vor- und Nachteilen von LLMs im Bildungsbereich ist ein verantwortungsbewusster Umgang mit diesen Technologien essenziell. Die Einsatzmöglichkeiten von LLMs müssen systematisch untersucht, geprüft und weiterentwickelt werden, wobei didaktische, ethische und rechtliche Aspekte zu berücksichtigen sind (vgl. Köller et al. 2024).

Ob Studierende generative KI künftig in ihrem eigenen Unterricht nutzen, hängt jedoch hochgradig mit deren Verständnis von generativer KI, der bisherigen Nutzung von generativer KI sowie dem wahrgenommenen Nutzen generativer KI zusammen (vgl. Chan/Zhou 2023).

Eine repräsentative Befragung aus dem Jahr 2023 zeigt: Etwa die Hälfte der Schüler*innen haben bereits generative KI genutzt und 58 % der befragten Schüler*innen wollen lernen, wie sie sinnvoll generative KI verwenden können (vgl. Bitkom 2023).

Einsatz traditioneller KI-Methoden in der Bildung

Einige dieser Herausforderungen können durch die Anwendung traditioneller KI-Methoden adressiert werden, da diese in geschlossenen Systemen operieren, in denen Daten und Funktionsweisen kontrolliert werden und vorhersehbar sind, und somit deterministische Ergebnisse produzieren.

So existieren schon länger etablierte KI-Methoden, die das Lehren und Lernen nachweislich effektiv unterstützen können. Hier sind insbesondere ITS zu nennen, die explizit den Wissens-

stand der Lernenden modellieren, um eine individuelle, passgenaue Lernunterstützung zu bieten. Diese Modellierung können ITS dazu nutzen, um z. B. in Form von interaktiven Übungsheften für die oder den Lernende*n individualisiertes und adaptives Feedback (vgl. Meurers et al. 2019) zu präsentieren oder basierend auf den Stärken und Schwächen der Lernenden adaptiv Aufgaben zu generieren (vgl. Heck/Meurers/Nuxoll 2022). Es ist wichtig hervorzuheben, dass ITS in der Regel für bestimmte Anwendungsfälle konzipiert werden und daher primär in spezifischen Lehr-Lernszenarien oder für ausgewählte Themengebiete Anwendung finden. Im Folgenden wird dargelegt, wie pädagogische Konzepte mithilfe von modernen, KI-basierten ITS unterstützt werden können, wobei die Anwendbarkeit einzelner Tools im Unterricht sorgfältig evaluiert werden muss. Die kritische Reflektion von ITS und deren Anwendbarkeit wird daher Teil der in WoLKE zu konzipierenden Lehrveranstaltung sein.

1.2.1 KI in der Sprachdidaktik

Im Bereich der (Fremd-)Sprachdidaktik bieten KI und insbesondere ITS (z. B. [FeedBook](#), vgl. Meurers et al. 2019) Unterstützungsmöglichkeiten für eine Vielzahl fachdidaktischer Aspekte, von denen wir einige im Folgenden ausführen. So können z. B. Tools für automatisches Input Enhancement (Sharwood-Smith 1993) – also die automatische (visuelle) Aufbereitung von Sprachlernmaterial – zur Förderung von Noticing (vgl. Schmidt 1995) und Sprachbewusstheit (vgl. Wildemann/Bien-Miller 2023) eingesetzt werden. Mit webbasierten Tools können so authentische Lesertexte nach lernförderlichen Mustern visuell und individualisiert auf Grundlage der Bedürfnisse der Lernenden aufbereitet ([COAST](#), vgl. Holz et al. 2018) oder bestimmte sprachliche Mittel hervorgehoben werden (vgl. Chinkina/Meurers 2016; Meurers et al. 2010). Im Sinne von Input Enrichment (vgl. Chinkina/Meurers 2016) – der Anreicherung bzw. Sicherstellung der Häufigkeit bestimmter sprachlicher Mittel in einem Sprachlernertext – können mittels linguistisch informierter Sprach-Suchmaschinen automatisch kompetenzadaptive Lehr-Lernmaterialien ausfindig gemacht und aufbereitet werden ([FLAIR](#); vgl. Chinkina/Meurers 2016; [KANSAS](#); vgl. Mayer/Jambor-Fahlen/Kholin 2023; Weiss/Dittrich/Meurers 2018).

Einen wesentlichen Beitrag können solche KI-basierten Suchmaschinen in hybriden Lehrer*innen-KI-Teams leisten, um entwicklungsproximale und motivierenden Texte, welche auf die Bedürfnisse der Schüler*innen zugeschnitten sind, bereitzustellen, um die (Schrift-)Sprachentwicklung zu fördern (vgl. Gass/Mackey 2014). Durch automatisierte Analyse sprachlicher Komplexität (vgl. Chen X./Meurers 2016) können Lehrkräfte sicherstellen, dass sich Lehr-Lernmaterialien im Rahmen der *Zone of Proximal Development* (vgl. Vygotskii/Kozulin 1986) befinden und nicht sprachlich überfordernd sind. Textgenerierende KI kann die adaptive Suche ergänzen und Texte generieren, umformulieren oder übersetzen, um für heterogene Sprachniveaus entsprechende Binnendifferenzierungsmöglichkeiten zur Verfügung zu stellen.

Außerdem kann die automatisierte Analyse sprachlicher Komplexität auch von Lehrkräften genutzt werden, um das sprachliche Niveau der Lernenden einzuschätzen, um so wiederum die sprachliche Heterogenität einer Klasse festzustellen und die Lehr-Lernmaterialien entsprechend anzupassen.

Im Sinne der Skill Acquisition Theory (vgl. DeKeyser 2020) werden KI-Systeme konzipiert, die Aufgaben adaptiv generieren und den Lernenden vorschlagen (vgl. Heck/Meurers/Nuxoll

2022) und so zum systematischen Üben komplexer Konzepte fördern. Diese Systeme können auch, basierend auf der individuellen Leistung der Lernenden, automatisch Feedback generieren und Lernende durch Scaffolding (vgl. Wood/Bruner/Ross 1976) zur korrekten Lösung führen.

Im Bereich der Schriftsprachförderung werden auch spielbasierte ITS erfolgreich eingesetzt, um z. B. im oder außerhalb des Unterrichts die Lesefähigkeiten (vgl. Görge et al. 2020) und Rechtschreibfähigkeiten ([Prosodiya](#); vgl. Holz et al. 2023; Witzel et al. 2024) von Grundschulkindern zu fördern. Solche spielbasierten Lese- und Rechtschreibfördermaßnahmen werden auch gezielt zur Unterstützung von Kindern mit schwachen Lese-Rechtschreibleistungen eingesetzt und tragen so dazu bei, dass Bildungsungleichheiten verringert werden.

Die Adaptivität konventioneller digitaler Tools konzentriert sich jedoch typischerweise auf sehr wenige Diversitätsmerkmale, nämlich das Ausgangsniveau und die Lernfortschritte im jeweiligen Lernbereich. Aus pädagogischer Perspektive sollte Sprachlernmaterial in der Alphabetisierung aber an den individuellen Kontext und die Interessen der Lernenden angepasst sein und angemessen Lerninput geben (vgl. Löffler/Weis 2016). Besonders im Grundbildungsbe-
reich ist es vor dem Hintergrund möglicher negativer Erfahrungen mit schulischem Lernen zentral, direkt an die Lebenswelt und Interessen der Lernenden anzuknüpfen, ohne sie sprachlich zu überfordern (vgl. Mayer/Jambor-Fahlen/Kholin 2023). Hier kann das Synergiepotenzial von textgenerierenden und etablierten KI-Methoden ausgeschöpft werden. Während textgenerierende KI Sprachlernertexte passend zu den individuellen Interessen der Lernenden automatisch erstellen kann, stellt eine ergänzende Sprachkomplexitätsanalyse sicher, dass sich die Texte in der entwicklungsproximalen Zone der Lernenden befinden.

1.2.2 KI in der MINT-Didaktik

Auch im MINT-Unterricht haben sich KI-Tools, insbesondere ITS wie Bettermarks ([bettermarks.com](#)), für das Fach Mathematik in Ergänzung zum traditionellen Unterricht als effektiv erwiesen (vgl. Blanc/Reichow/Paaßen 2023). ITS ermöglichen einerseits eine individualisierte und adaptive Unterstützung für Lernende, indem sie personalisierte Lernpfade und -materialien bereitstellen, um die Lernenden z. B. bei der Lösung mathematischer Probleme zu unterstützen oder komplexe naturwissenschaftliche Konzepte interaktiv zu vermitteln. Die Wirksamkeit des Einsatzes solcher ITS im Unterricht hängt jedoch davon ab, wie kompetent und erfahren die jeweilige Lehrkraft im Umgang mit ITS ist (vgl. Blanc/Reichow/Paaßen 2023).

Darüber hinaus eröffnen auch generative KI-Tools neue Horizonte und zugleich neue Spannungsfelder für das Lehren und Lernen in den MINT-Fächern. Ein fachspezifisches Spannungsfeld besteht darin, dass die rein probabilistische Funktionsweise von LLMs und generativen KI-Tools, die nicht im Welt- und Fachwissen verankert ist, in starkem Kontrast zur ‚naturwissenschaftlichen Methode‘ und zum logisch stringenten, wissenschaftlichen Argumentieren steht. Um generative KI-Tools gewinnbringend im Unterricht einsetzen zu können, ist es unerlässlich, dass Lehrkräfte die Funktionsweise und den Output der jeweiligen Tools vor dem Hintergrund dieses Spannungsfeldes reflektieren können.

Daraus ergibt sich die Notwendigkeit, (angehende) MINT-Lehrkräfte durch Bildungsangebote und Materialsammlungen zu kompetenten Multiplikator*innen für den reflektierten und didaktisch sinnvollen Einsatz KI-gestützter digitaler Hilfsmittel zu befähigen.

Allgemein wird in der MINT-Didaktik die Nutzung digitaler Hilfsmittel seit langer Zeit durch Unterrichtsvorschläge und Materialsammlungen unterstützt. Ein aktuelles Kompendium bietet beispielsweise die *Toolbox für den Unterricht* (vgl. Meßinger-Koppelt/Maxton-Küchenmeister 2018). Hier wird die Breite der Anwendungsmöglichkeiten digitaler Medien deutlich. Sie reicht von interaktiven Simulationen wie z. B. PhET (phet.colorado.edu) und der digitalen Messwerterfassung (z. B. mit dem Smartphone) über den Einsatz von Wärmebildkameras und 3D-Druckern bis hin zu AR- und VR-Anwendungen.

Ein weiteres digitales Hilfsmittel, das bereits seit etwa 15 Jahren in sämtlichen MINT-Bereichen zum Einsatz kommt, ist die *Wissensberechnungsmaschine* (vgl. Gottwald 2009) Wolfram Alpha (wolframalpha.com), die natürlichsprachliche Eingaben verarbeitet und eine Vielzahl von Informationen und Berechnungen liefert. Wolfram Alpha kombiniert umfangreiche Wissensdatenbanken mit leistungsfähigen Algorithmen, um komplexe mathematische und naturwissenschaftliche Probleme zu lösen sowie Zusammenhänge und Konzepte zu erklären. Im MINT-Unterricht bietet Wolfram Alpha einen Mehrwert durch den sofortigen Zugriff auf präzise Berechnungen, Grafiken und Erklärungen, die es den Lernenden ermöglichen, ihr Verständnis zu vertiefen, komplexe Probleme zu lösen und fachliche Zusammenhänge besser zu verstehen. Die seit kurzem in Form eines Plugins verfügbare Integration von Wolfram Alpha in GPT-4 erscheint daher vielversprechend für die Überbrückung der Diskrepanz zwischen der probabilistischen, fiktionalisierenden Funktionsweise von textgenerierender KI und den Anforderungen an die Darstellung naturwissenschaftlich korrekter Zusammenhänge und mathematisch korrekter Berechnungen (vgl. Spannagel 2023).

Für KI-Werkzeuge gilt, was gleichermaßen auch für digitale Werkzeuge im Allgemeinen gilt: sie können im naturwissenschaftlichen Unterricht die Unterrichtsqualität steigern, wenn folgende Aspekte berücksichtigt werden: a) digitale Hilfsmittel sollten als Ergänzung zu traditionellen Lehr-Lern-Materialien genutzt werden, b) digitale Hilfsmittel sollten für kurze Zeit und nicht kontinuierlich eingesetzt werden, c) Lernende sollten kooperativ mit digitalen Werkzeugen arbeiten und d) Lehrkräfte sollten im Umgang mit digitalen Hilfsmitteln geschult sein (vgl. Hillmayr et al. 2017).

Zum Erwerb solcher Kompetenzen im Umgang mit digitalen Werkzeugen wird im MINT-Bereich für die Orientierung an fachspezifischen Aspekten und Anforderungen argumentiert (vgl. Thyssen et al. 2020) und häufig auf den Orientierungsrahmen *Digitale Kompetenzen für das Lehramt in den Naturwissenschaften* (DiKoLAN) rekurriert (vgl. Becker et al. 2020; Thyssen et al. 2023).

Daran anknüpfend gibt es derzeit an vielen deutschen Hochschulen Bemühungen, auch KI in die MINT-Lehrkräftebildung zu integrieren, z. B. mittels disziplinübergreifender sogenannter Micro-Degrees an der Universität Hannover (vgl. Schanze/Kühne 2024) oder mittels auf die Lernenden zugeschnittenen formativen Feedbacks, um selbstreguliertes Lernen zu fördern, an der LMU München (vgl. Steinert et al. 2023). Auch das Digitale Lehr-Lern-Labor der Technischen Universität München (TUM-DigiLLab) konzentriert sich auf die Entwicklung innovativer Lehr-Lern-Konzepte im Themenfeld der KI und hat zum Ziel, die Lücke zwischen theoretischer Wissensvermittlung und alltagsnaher Unterrichtspraxis zu schließen (vgl. Bewersdorff/Nerdel 2023).

Trotz der zunehmenden Bemühungen um die Integration von Kompetenzförderung zum Einsatz von KI in der (MINT-)Lehrkräfteausbildung bleibt ersichtlich, dass diese Bemühungen derzeit fachlich wie räumlich weit gestreut sind und noch nicht die systematische und flächendeckende Lehrkräfteausbildung im Bereich KI bewirken können, die angesichts der fortschreitenden Durchdringung aller möglichen Lebensbereiche durch KI-Technologien erforderlich wäre. Daher wollen wir im Rahmen des WoLKE-Projekts zur Systematisierung der KI-Bildung von angehenden Lehrkräften beitragen.

2. Das Projekt „WoLKE“

Im Folgenden stellen wir das vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg geförderte Projekt „WoLKE“ vor, in dem für die zuvor skizzierten fachdidaktischen Desiderate Lehrveranstaltungen für angehende Lehrkräfte entwickelt, evaluiert und als Open Educational Resources (OER) bereitgestellt werden.

2.1 Ziele

Das übergeordnete Ziel des Projektes ist es, angehende Lehrkräfte zu einem reflektierten Umgang mit KI-basierten Werkzeugen im Unterricht zu befähigen, die sie hinsichtlich konkreter schulischer Herausforderungen in fachdidaktisch wohlmotivierter Weise einsetzen können. Das Akronym „WoLKE“ steht dabei für den Kern der im Projekt bearbeiteten Fragestellung: „*Wo* bieten KI-Methoden *Lösungen* für fachdidaktische Herausforderungen? Computerlinguistisch fundierte *Konzeption* und *Evaluation* curricular verankerter Lehrveranstaltungen für die Sprach- und MINT-Didaktik“.

Dafür entwickeln wir in WoLKE Lehrveranstaltungen für angehende Lehrkräfte für den didaktisch sinnvollen Einsatz von KI-Methoden und -Werkzeugen im Unterricht. In den Lehrveranstaltungen können angehende Lehrkräfte methodische Kompetenzen erwerben, wie KI-basierte Werkzeuge unterstützend im Unterricht eingesetzt werden können, um konkreten schulischen Herausforderungen zu begegnen, insbesondere in den Sprachen und MINT-Fächern. Dabei stehen der fachdidaktische Praxisbezug im Sprach- und MINT-Unterricht sowie ein methodisch und ethisch fundiertes Verständnis des Potenzials und der Grenzen der KI-Methoden im Mittelpunkt. Die angehenden Lehrkräfte sollen dadurch zu einem reflektierten Umgang mit KI-basierten Werkzeugen im Unterricht befähigt werden.

Als Grundlage hierfür wird eine Übersicht über etablierte und aktuelle, textgenerierende KI-Methoden erstellt, deren Eignung und Reifegrad mit Blick auf die unterschiedlichen fachdidaktischen Desiderate systematisch analysiert werden. Der fachdidaktische Einsatz etablierter KI-Methoden wurde bereits in den Kapiteln *KI in der Sprachdidaktik* und *KI in der MINT-Didaktik* genauer dargestellt. Diese KI-Methoden, wie z. B. ITS, bieten sich insbesondere mit ihrer expliziten Lernermodellierung und der individuellen Unterstützung zur Ergänzung des lehrkraftorchestrierten Unterrichts an, während textgenerierende KI-Methoden u. a. viel Potenzial bei der Unterstützung der Lehrenden in Form eines hybriden Teams im Sinne einer Mensch-Maschine-Kooperation statt nur einer Mensch-Maschine-Interaktion bieten. Eine Aufarbeitung des aktuellen Forschungsstandes zu Grenzen des Einsatzes von KI im Unterricht (vgl. u. a. Küchemann et al. 2024) soll den Studierenden darüber hinaus helfen, ein Bewusstsein dafür zu entwickeln, für welche Lehr-Lernsituationen und für welche Lernenden KI sich

didaktisch zielführend und für welche weniger gewinnbringend zur Gestaltung von fachlich wohlmotiviertem Unterricht einsetzen lässt.

Das Projekt lotet auch die Kombination unterschiedlicher KI-Methoden aus, etwa um aufzuzeigen, wie sich Schwächen von textgenerierender KI bei der fachlichen Modellierung adressieren lassen.

Auf dieser Basis werden die Lehrveranstaltungen zur Vermittlung der grundlegenden methodischen Kompetenzen und didaktischen Möglichkeiten für Lehramtsstudierende der Sprachen und MINT-Fächer evidenzbasiert entwickelt und evidenzorientiert empirisch evaluiert.

Die Lehrveranstaltungen werden als im Lehramtsstudium curricular verankerte Seminare an den Hochschulstandorten des Verbunds (siehe unten) angeboten und besitzen sowohl eine gemeinsame fachübergreifende Basis als auch fachspezifische Schwerpunkte.

Begleitend zu den beiden Lehrveranstaltungen bieten wir zu den für das Projekt entwickelten OER-Materialien der Lehrveranstaltung jeweils einen Transferworkshop für Hochschullehrende an, die diese Lehrveranstaltungen an ihrer jeweiligen Hochschule anbieten möchten.

2.2 Verbund

Das Projekt zeichnet sich durch einen interdisziplinären und standortübergreifenden Verbund aus (Fremd-)Sprachdidaktik, Informatik und Informatikdidaktik, MINT-Didaktiken, Computerlinguistik und empirischer Bildungsforschung aus, der in Abbildung 1 dargestellt ist. Dieser ermöglicht eine effektive Ko-Konstruktion der Lehrveranstaltungen und Transferworkshops sowie eine qualitätssichernde Wirksamkeitsüberprüfung. Die im Verbund angelegte parallele Einbindung von Sprach- und MINT-Didaktiken bietet die Möglichkeit, die spezifischen Charakteristiken und Limitationen der unterschiedlichen KI-Methoden hinsichtlich sprachlicher und fachlicher Anforderungen explizit zu machen.

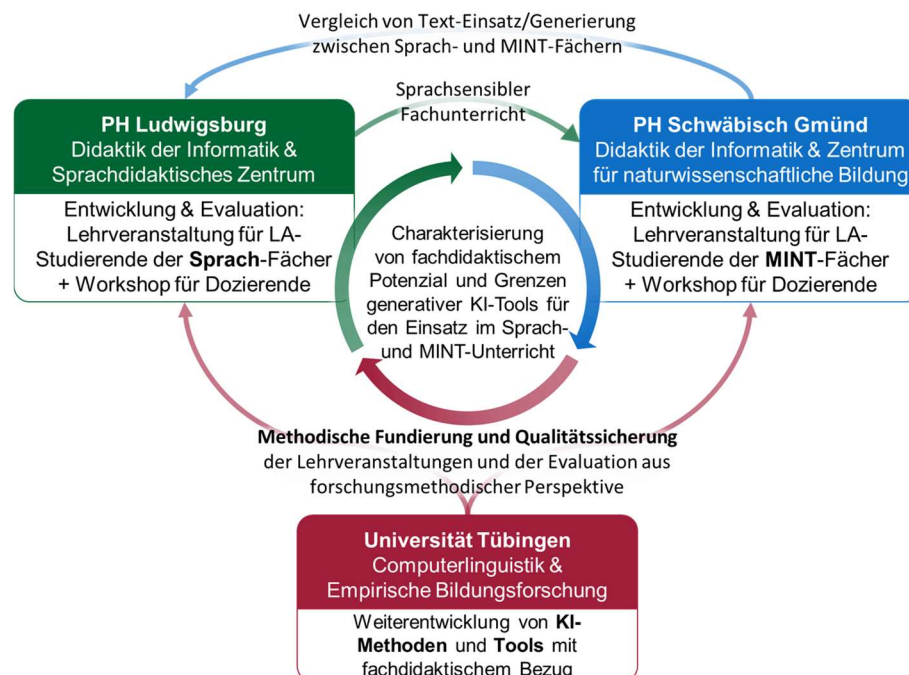


Abbildung 1: Der WoLKE-Projektverbund

Die Entwicklung und Evaluation der Seminare für Lehramtsstudierende in den sprachlichen Fächern erfolgen an der Pädagogischen Hochschule Ludwigsburg sowie in den MINT-Fächern an der Pädagogischen Hochschule Schwäbisch Gmünd und basieren auf einer gemeinsam erarbeiteten Konzeption.

2.3 Inhalte der Lehrveranstaltungen

Beide Lehrveranstaltungen setzen sich mit drei zentralen Dimensionen des Potenzials und der Grenzen von KI-Tools für den reflektierten, computerlinguistisch fundierten und fachdidaktisch passgenauen Einsatz im Unterricht auseinander.

Zunächst wird die Dimension der computerlinguistischen Hintergründe und fachdidaktischen Aspekte im Kontext des Einsatzes von KI im Unterricht behandelt. Der Fokus liegt dabei auf der Vermittlung grundlegender Kenntnisse über die Funktionsweise von LLMs und textgenerierender KI, insbesondere im Hinblick auf deren probabilistische Grundlagen und die Fiktionalität der generierten Inhalte. Dabei wird der traditionelle Prozess des Verfassens wissenschaftlicher Texte, der quellenbasiert, argumentativ und an den Bedürfnissen der Adressaten orientiert ist, mit der probabilistischen Generierung durch KI kontrastiert. Darüber hinaus werden die Möglichkeiten und Grenzen des Einsatzes von KI-Werkzeugen im (sprachsensiblen) Unterricht diskutiert, insbesondere im Hinblick auf Prompt Engineering, Quellenrecherche und die reflektierte Überarbeitung von KI-generierten Texten.

In der Lehrveranstaltung für Lehramtsstudierende der sprachlichen Fächer werden aus sprachdidaktischer Perspektive Qualitätskriterien für KI-generierte Texte erarbeitet, die Aspekte wie Form, Inhalt und Kohärenz berücksichtigen.

In der Lehrveranstaltung für Lehramtsstudierende der MINT-Fächer werden aus MINT-didaktischer Perspektive Qualitätskriterien für KI-generierte Texte betrachtet, wobei neben der Form auch der fachliche Inhalt und die naturwissenschaftliche Argumentation im Fokus stehen.

Die zweite zentrale Dimension der Lehrveranstaltung sind *(fach-)didaktische und unterrichtspraktische Überlegungen*, die die *Rolle*, den *Mehrwert* sowie die *Grenzen* von KI-Werkzeugen im Unterricht beleuchten. Dabei wird untersucht, welchen spezifischen Mehrwert textgenerierende KI-Werkzeuge im Vergleich zu traditionellen Methoden und anderen digitalen (KI-)Tools bieten. Die Möglichkeiten und Grenzen von KI-Ansätzen sowohl mit als auch ohne explizite Modellierung der curricularen Lernziele und des Lernprozesses werden diskutiert. Besonderes Augenmerk liegt auf der synergetischen Nutzung expliziter Modellierung in traditionellen KI-Tools und aktueller textgenerierender KI-Methoden.

Das Seminar für Lehramtsstudierende sprachlicher Fächer widmet sich spezifischen Anwendungen von KI im sprachdidaktischen Kontext. Dabei werden Themen wie ITS, automatische Analyse, adaptive Auswahl und Anreicherung sprachdidaktischer Lehr-Lernmaterialien behandelt. Es wird diskutiert, welche Vor- und Nachteile eine explizite Modellierung von Sprache im Vergleich zu einer impliziten probabilistischen Sprachverarbeitung bietet, auch im Hinblick auf curriculare Vorgaben entsprechend der Bildungspläne.

Die Lehrveranstaltung für MINT-Lehramtsstudierende fokussiert den Einsatz von KI im MINT-Unterricht. Dabei werden sowohl klassische algorithmische Lösungen in Übungsplattformen wie Bettermarks oder in Simulationen als auch Ansätze wie Augmented Reality (AR)

betrachtet. Zudem wird die empirisch-naturwissenschaftliche Methode mit der Big-Data-basierten, probabilistischen Arbeitsweise der textgenerierenden KI kontrastiert, um die Unterschiede und Potenziale dieser Ansätze zu verdeutlichen.

Die dritte Dimension, die in beiden Kursen behandelt wird, betrachtet die *(wissenschafts-)ethischen und gesellschaftlichen Aspekte* im Zusammenhang mit textgenerierenden KI-Werkzeugen. Die wissenschaftlichen, ethischen und gesellschaftlichen Implikationen, die mit dem Einsatz solcher Werkzeuge verbunden sind, insbesondere im Hinblick auf die Reproduktion von Biases, werden eingehend reflektiert.

Die Lehrveranstaltung für Lehramtsstudierende sprachlicher Fächer legt einen Schwerpunkt auf die Implikationen des Sprachgebrauchs durch textgenerierende KI für das menschliche Selbstverständnis. Dabei werden auch Fragen der Zuschreibung von Emotionen und deren Bedeutung für die Kommunikation und Interaktion zwischen Mensch und Maschine diskutiert.

Die Lehrveranstaltung für Lehramtsstudierende der MINT-Fächer befasst sich mit den Implikationen generativer KI für das Wissenschaftsverständnis und das Vertrauen in naturwissenschaftliche Erkenntnisse. Es wird diskutiert, wie der Einsatz von KI-Werkzeugen das Verständnis wissenschaftlicher Methodik und die Wahrnehmung naturwissenschaftlicher Erkenntnisse beeinflussen kann und welche potenziell negativen Auswirkungen dies auf das Wissenschaftsverständnis haben und das Vertrauen in die Wissenschaft schwächen könnte.

Durch die ganzheitliche Betrachtung dieser ethischen und gesellschaftlichen Aspekte können angehende Lehrkräfte zu einem umfassenden Verständnis der Herausforderungen und Möglichkeiten des Einsatzes von KI im Unterricht gelangen.

2.4 Evaluation der Lehrveranstaltungen

Für die Evaluation der Lehrveranstaltungen und Lehr-Lernmaterialien soll ein Mixed-Methods-Ansatz verwendet werden, der quantitative und qualitative Methoden kombiniert. Nach dem Prinzip des Design-based Research (vgl. Anderson/Shattuck 2012) werden die Evaluationsmaßnahmen eng mit der Entwicklung und Umsetzung des didaktischen Konzepts sowie mit der Ausarbeitung fachdidaktischer Theorien bezüglich der effektiven und kritisch reflektierten Integration von KI in die pädagogische Praxis verzahnt. Dies unterstützt eine explizite Verknüpfung von theoretischen Annahmen, praktischer Umsetzung und Evaluation, um ein umfassendes Verständnis der Wirksamkeit der Lehrveranstaltungen und ihres didaktischen Konzepts zu erhalten. Um die Lehrveranstaltungen qualitativ zu beforschen, sollen aufgezeichnete Interviews genutzt werden, die ein tieferes Verständnis der individuellen Eindrücke der Studierenden ermöglichen. Diese qualitativen Daten sollen dann mit quantitativen Daten aus standardisierten Tests und Fragebögen kombiniert werden, um einen umfassenden Einblick in den Lernfortschritt der Studierenden zu erhalten. Die quantitative Wirksamkeitsüberprüfung erstreckt sich über mehrere Messzeitpunkte, während die qualitativen Daten mit den Studierenden als retrospektive Interviews hinsichtlich ihrer Erfahrungen und Einschätzungen geplant sind. Durch die Kombination dieser unterschiedlichen empirischen Evaluationsansätze kann ein ganzheitliches Bild über die Wirksamkeit des didaktischen Konzepts gewonnen werden. Die spezifischen Evaluationsmethoden und Messinstrumente werden derzeit auf Basis ei-

ner ausführlichen Exploration der verfügbaren Literatur unter Beratung ausgewiesener Forschender aus der empirischen Bildungsforschung zusammengestellt.

2.5 Zeitlicher Horizont

Die dreijährige Laufzeit des Projekts erstreckt sich von Dezember 2023 bis November 2026 und verläuft in insgesamt drei Phasen (vgl. Abbildung 2). In Phase 1 wird die Lehrveranstaltung zu Beginn ko-konstruktiv im Verbund entwickelt, basierend auf der methodisch fundierten Analyse etablierter und aktueller KI-Methoden. Dabei werden Lehrkräfte im Rahmen von Teilabordnungen speziell für das Projekt WoLKE als Praxispartner*innen eng in die Konzeption, Erprobung und Evaluation der zu entwickelnden Inhalte einbezogen. Darüber hinaus werden Hochschuldozierende aus den Sprach- und MINT-Didaktiken in allen Phasen des Projekts als beratende Expert*innen für die Entwicklung und Umsetzung der Lehrveranstaltungen herangezogen.

Als Grundlage für die Lehrveranstaltungen werden traditionelle, wissensbasierte und aktuelle datengetriebene, textgenerierende KI-Methoden hinsichtlich Eignung, Reifegrad und Synergiepotenzial für den fachdidaktisch motivierten Einsatz analysiert, zum Teil weiterentwickelt, und die Grenzen und Möglichkeiten der jeweiligen Methoden für unterschiedliche Lernsituationen und -gruppen explizit gemacht. Zur Realisierung des Potenzials werden dabei für konkrete Herausforderungen der Schulpraxis – wie etwa Unterrichtsvorbereitung oder Automatisierung von Aufgabenstellungen und Feedback – auf Basis der wissenschaftlichen Erkenntnisse der Fachdidaktik gezielt KI-basierte Methoden exploriert, die eine Lösung unterstützen können.

In Phase 2 ab dem Wintersemester 2024/2025 sind erste Pilotierungen der Lehrveranstaltungsmodule an den jeweiligen Hochschulstandorten geplant. Basierend auf den in der Pilotierung gewonnenen Erkenntnissen werden die Lehrveranstaltungen vollständig ausgearbeitet und in den darauffolgenden drei Semestern entsprechend den Prinzipien des Design-based Research in iterativen Überarbeitungszyklen ausgebracht und wiederholt evaluiert. Hierfür werden die Veranstaltungen an den Verbundhochschulen curricular verankert angeboten.

Damit einhergehend erfolgt auch die empirische Begleitung und Untersuchung der Lehrveranstaltungen nach dem Mixed-Methods-Design (vgl. Abschnitt 2.4 *Evaluation der Lehrveranstaltungen*) mit quantitativer und qualitativer Datenerhebung mit mehreren Messzeitpunkten sowie der Datenerfassung und -aufbereitung. Aktuell befindet sich das Projekt in Recherche und Beratung zur genauen forschungsmethodischen Vorgehensweise und der Auswahl der zur didaktischen Zielsetzung passenden Erhebungsinstrumente.

In Phase 3, die durch ineinandergreifende Zyklen zeitlich mit den anderen beiden Phasen verzahnt ist, werden die aus den Daten gewonnenen Befunde mit den Zielen der Lehrveranstaltungen abgeglichen und gegebenenfalls Änderungen an den Lehrveranstaltungen vorgenommen. Parallel dazu werden die Materialien der Lehrveranstaltungen für den Transfer an die Verbundhochschulen vorbereitet und für abschließende Transferworkshops und -tagungen als OER aufbereitet.

Insbesondere der Transfer der entwickelten Seminare innerhalb des Projektverbunds, aber auch durch die Transfertagungen und Workshops für Dozierende, soll die Erkenntnisse und die entwickelten Lehrveranstaltungen multiplikativ in die Breite tragen.

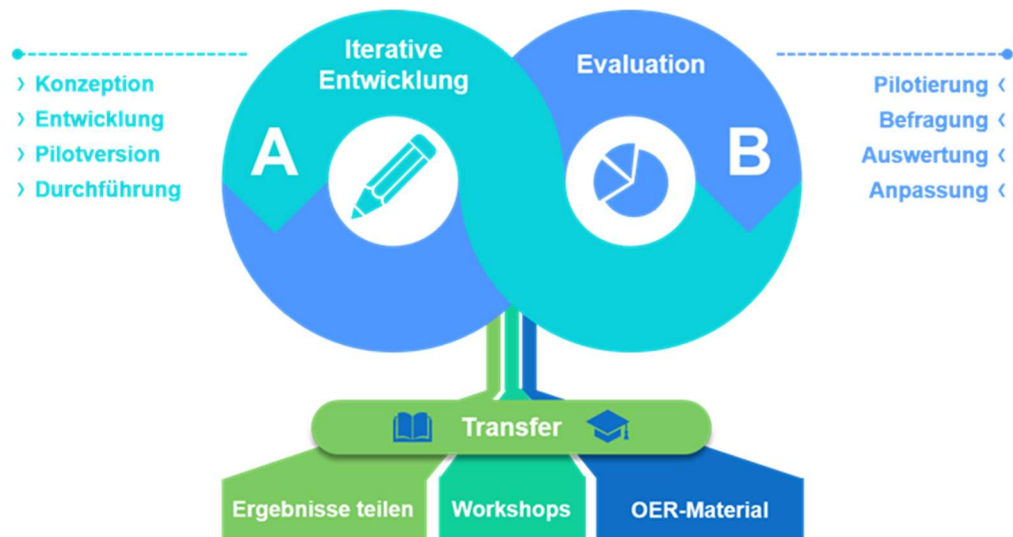


Abbildung 2: Forschungsdesign und Projektverlauf von WoLKE

3. Schluss

Zusammenfassend stellt das Projekt „WoLKE“ einen ganzheitlichen Ansatz zur Integration von KI-Tools in die Lehrkräftebildung dar, das die Entwicklung überfachlicher, fachdidaktischer und methodischer Kompetenzen befördern soll, um den fachdidaktisch passgenauen, reflektierten Einsatz von KI-Tools sowohl im Sprach- als auch im MINT-Unterricht zu fördern.

Durch die interdisziplinäre Zusammenarbeit und die evidenzbasierte Entwicklung der Lehrveranstaltungen, die sich auf aktuelle Befunde aus der computerlinguistischen, informatischen und fachdidaktischen Forschung zu KI in der Bildung stützt, soll das Projekt dazu beitragen, die Potenziale und Grenzen von KI-Tools im Unterricht zu kontextualisieren, zu analysieren und zu bewerten. Darüber hinaus zielt das Projekt darauf ab, einen bedeutenden Beitrag zur Lehrkräftebildung zu leisten, indem es angehenden Lehrkräften die Möglichkeit bietet, die notwendigen Kenntnisse und Kompetenzen für einen verantwortungsvollen Einsatz von KI-Tools zu erlangen. Ziel ist es, dass Lehrkräfte KI-Tools didaktisch fundiert und technisch sowie gesellschaftlich verantwortungsbewusst einsetzen können, um nachhaltig zur Bildung für ein souveränes Agieren in der digitalen Welt beizutragen.

Literatur

- Aghion, Philippe/Jones, Benjamin F./Jones, Charles I. (2017): Artificial intelligence and economic growth. National Bureau of Economic Research.
- Anantrasirichai, Nantheera/Bull, David (2022): Artificial intelligence in the creative industries: A review. In: *Artificial Intelligence Review*, 55 (1), pp. 589–656. doi.org/10.1007/s10462-021-10039-7.
- Anderson, Terry/Shattuck, Julie (2012): Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, 41 (1), pp. 16–25. doi.org/10.3102/0013189X11428813.
- Becker, Sebastian/Bruckermann, Till/Finger, Alexander/Huwer, Johannes/Kremser, Erik/Meier, Monique/Thoms, Lars-Jochen/Thyssen, Christoph/von Kotzebue, Lena (2020): Orientierungsrahmen Digitale Kompetenzen für das Lehramt in den Naturwissenschaften–DiKoLAN. In: Becker, Sebastian/Meßinger-Koppelt, Jenny/Thyssen, Christoph (Hrsg.): *Digitale Basiskompetenzen – Orientierungshilfe und Praxisbeispiele für die universitäre Lehramtsausbildung in den Naturwissenschaften*, S. 14–43. joachim-herz-stiftung.de/fileadmin/Redaktion/Projekte/Naturwissenschaften/2020_Nawi_Digitale_Basiskompetenzen_web.pdf.

- Bewersdorff, Arne/Nerdel, Claudia (2023): Das TUM-DigiLLab: Lehr-Lernraum sowie Forschungs- und Entwicklungsumgebung zur Förderung digitaler Kompetenzen. In: Meier, Monique/Greefrath, Gilbert/Hammann, Marcus/Wodzinski, Rita/Ziepprecht, Kathrin (Hrsg.): Lehr-Lern-Labore und Digitalisierung Springer Fachmedien, S. 137–141. doi.org/10.1007/978-3-658-40109-2_10.
- Bitkom (2023, Mai 23): Hälfte der Schülerinnen und Schüler hat schon mal ChatGPT genutzt (Presseinformation) Bitkom e. V. bitkom.org/Presse/Presseinformation/ChatGPT-in-Schule-nutzen.
- Blanc, Berit/Reichow, Insa/Paaßen, Benjamin (2023): Was wirkt? Eine Literaturstudie zur Wirksamkeit von Systemeigenschaften in Mathematik-Lernumgebungen, S. 293–294. dl.gi.de/handle/20.500.12116/42224.
- Chan, Cecilia Ka Yuk/Zhou, Wenxin (2023): An expectancy value theory (EVT) based instrument for measuring student perceptions of generative AI. In: Smart Learning Environments, 10 (1), pp. 64–86. doi.org/10.1186/s40561-023-00284-4.
- Chen, Lijia/Chen, Pingping/Lin, Zhijian (2020): Artificial intelligence in education: A review. In: IEEE Access, 8, pp. 75264–75278. doi.org/10.1109/ACCESS.2020.2988510.
- Chen, Xiaobin/Meurers, Detmar (2016): CTAP: A Web-Based Tool Supporting Automatic Complexity Analysis. In: Brunato, Dominique/Dell’Orletta, Felice/Venturi, Giulia/François, Thomas/Blache, Philippe (Eds.): Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC). The COLING 2016 Organizing Committee. pp. 113–119. aclanthology.org/W16-4113.
- Chinkina, Maria/Meurers, Detmar (2016): Linguistically Aware Information Retrieval: Providing Input Enrichment for Second Language Learners. In: Tetreault, Joel/Burstein, Jill/Leacock, Claudia/Yannakoukakis, Helen (Eds.): Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications, pp. 188–198. doi.org/10.18653/v1/W16-0521.
- Colling, Leona/Heck, Tanja/Meurers, Detmar (2023): Reconciling Adaptivity and Task Orientation in the Student Dashboard of an Intelligent Language Tutoring System. In: Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023), pp. 288–299.
- DeKeyser, Robert (2020): Skill acquisition theory. In: Vanpatten, Bill/Keating, Gregory D./Wulff, Stefanie (Eds.): Theories in second language acquisition. New York: Routledge, pp. 83–104.
- Ferrara, Emilio (2023): Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models. In: First Monday, 28 (11). doi.org/10.5210/fm.v28i11.13346.
- Gass, Susan M./Mackey, Alison (2014): Input, interaction/output in second language acquisition. In: Vanpatten, Bill/Keating, Gregory D./Wulff, Stefanie (Eds.): Theories in second language acquisition. New York: Routledge, pp. 194–220.
- Goodfellow, Ian/Bengio, Yoshua/Courville, Aaron (2016): Deep learning. Cambridge: MIT Press.
- Görgen, Ruth/Huemer, Sini/Schulte-Körne, Gerd/Moll, Kristina (2020): Evaluation of a digital game-based reading training for German children with reading disorder. In: Computers & Education, 150, p. 103834. doi.org/10.1016/j.compedu.2020.103834.
- Gottwald, Susanne (2009): Die Wissensberechnungsmaschine Wolfram|Alpha. opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/1161.
- Heck, Tanja/Meurers, Detmar/Nuxoll, Florian (2022): Automatic exercise generation to support macro-adaptivity in intelligent language tutoring systems. Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022, p. 162.
- Hillmayr, Delia/Reinhold, Frank/Ziernwald, Lisa/Reiss, Kristina (2017): Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe: Einsatzmöglichkeiten, Umsetzung und Wirksamkeit. Hrsg. vom Zentrum für Internationale Vergleichsstudien. Münster: Waxmann.
- Holz, Heiko/Ninaus, Manuel/Schwerter, Jakob/Parrisius, Cora/Beuttler, Benedikt/Brandelik, Katharina/Meurers, Detmar (2023): A digital game-based training improves spelling in German primary school children – A randomized controlled field trial. In: Learning and Instruction, 87, p. 101771.

- doi.org/10.1016/j.learninstruc.2023.101771.
- Holz, Heiko/Weiss, Zarah/Brehm, Oliver/Meurers, Detmar (2018): COAST – Customizable Online Syllable Enhancement in Texts. A flexible framework for automatically enhancing reading materials. In: Tetreault, Joel/Burstein, Jill/Kochmar, Ekaterina/Leacock, Claudia/Yannakoudakis, Helen (Eds.): Proceedings of the thirteenth workshop on innovative use of NLP for building educational applications, pp. 89–100. doi.org/10.18653/v1/W18-0509.
- Kasneci, Enkelejda/Sessler, Kathrin/Küchemann, Stefan/Bannert, Maria/Dementieva, Daryna/Fischer, Frank/Gasser, Urs/Groh, Georg/Günnemann, Stephan/Hüllermeier, Eyke/Krusche, Stephan/Kutyniok, Gitta/Michaeli, Tilman/Nerdel, Claudia/Pfeffer, Jürgen/Poquet, Oleksandra/Sailer, Michael/Schmidt, Albrecht/Seidel, Tina/Kasneci, Gjergji (2023): ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, p. 102274. doi.org/10.1016/j.lindif.2023.102274.
- Köller, Olaf/Thiel, Felicitas/van Ackeren-Mindl, Isabell/Anders, Yvonne/Becker-Mrotzek, Michael/Cress, Ulrike/Diehl, Claudia/Kleickmann, Thilo/Lütje-Klose, Birgit/Prediger, Susanne/Seeber, Susan/Ziegler, Birgit/Lewalter, Doris/Maaz, Kai/Reintjes, Christian/Stanat, Petra (2024): Large Language Models und ihre Potenziale im Bildungssystem. Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz, S. 1–31. doi.org/10.25656/01:28303.
- Küchemann, Stefan/Avila, Karina E./Dinc, Yavuz/Hortmann, Chiara/Revenga, Natalia/Ruf, Verena/Stausberg, Niklas/Steinert, Steffen/Fischer, Frank/Fischer, Martin R./Kasneci, Enkelejda/Kasneci, Gjergji/Kuhr, Thomas/Kutyniok, Gitta/Malone, Dr Sarah/Sailer, Michael/Schmidt, Albrecht/Stadler, Matthias/Weller, Jochen/Kuhn, Jochen (2024): Are Large Multi-modal Foundation Models all we need? On Opportunities and Challenges of these Models in Education. doi.org/10.35542/osf.io/n7dvf.
- Leiker, Daniel/Finnigan, Sara/Gyllen, Ashley Ricker/Cukurova, Mutlu (2023): Prototyping the use of Large Language Models (LLMs) for adult learning content creation at scale (arXiv:2306.01815). arxiv.org/abs/2306.01815.
- Löffler, Cordula/Weis, Susanne (2016): Didaktik der Alphabetisierung. In: Löffler, Cordula/Korfkamp, Jens (Hrsg.): Handbuch zur Alphabetisierung und Grundbildung Erwachsener. Münster: Waxmann, S. 365–382.
- Mayer, Nadine/Jambor-Fahlen, Simone/Kholin, Mareike (2023): Einstellungen von Lehrpersonen zu digitalen Medien und die effektive Nutzung einer Suchmaschine (KANSAS). In: *Lernen und Lernstörungen*, 12 (2), S. 95–106. doi.org/10.1024/2235-0977/a000396.
- Meßinger-Koppelt, Jenny/Maxton-Küchenmeister, Jörg (2018): *Naturwissenschaften digital: Toolbox für den Unterricht*. Hamburg: Joachim Herz Stiftung Verlag.
- Meurers, Detmar/De Kuthy, Kordula/Nuxoll, Florian/Rudzewitz, Björn/Ziai, Ramon (2019): Scaling Up Intervention Studies to Investigate Real-Life Foreign Language Learning in School. In: *Annual Review of Applied Linguistics*, 39, pp. 161–188. doi.org/10.1017/S0267190519000126.
- Meurers, Detmar/Ziai, Ramon/Amaral, Luiz/Boyd, Adriane/Dimitrov, Aleksandar/Metcalf, Vanessa/Ott, Niels (2010): Enhancing Authentic Web Pages for Language Learners. In: Tetreault, Joel/Burstein, Jill/Leacock, Claudia (Eds.): Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications. *Association for Computational Linguistics*, pp. 10–18. aclanthology.org/W10-1002.
- Nishant, Rohit/Kennedy, Mike/Corbett, Jacqueline (2020): Artificial intelligence for sustainability: Challenges, opportunities/a research agenda. In: *International Journal of Information Management*, 53, p. 102104.
- Parrisius, Cora/Wendebourg, Katharina/Rieger, Sven/Loll, Ines/Pili-Moss, Diana/Colling, Leona/Blume, Carolyn/Pieronczyk, Ines/Holz, Heiko/Bodnar, Stephen/Schmidt, Torben/Trautwein, Ulrich/Meurers, Detmar/Nagengast, Benjamin (2022): Effective Features of Feedback in an Intelligent

- Tutoring System-A Randomized Controlled Field Trial (Pre-Registration).
- Rajpurkar, Pranav/Chen, Emma/Banerjee, Oishi/Topol, Eric J. (2022): AI in health and medicine. In: *Nature medicine*, 28 (1), pp. 31–38.
- Rudzewitz, Björn/Ziai, Ramon/Nuxoll, Florian/De Kuthy, Kordula/Meurers, Detmar (2019): Enhancing a Web-based Language Tutoring System with Learning Analytics. EDM (Workshops), pp. 1–7.
- Schanze, Sascha/Kühne, Patricia (2024): AI in Teacher Education. In: Graulich, Nicole/Arnold, Julia/Sorge, Stefan/Kubsch, Marcus (Hrsg.): *Lehrkräftebildung von morgen*. Münster: Waxmann, S. 309–317. doi.org/10.31244/9783830997962.33.
- Schmidt, Richard (1995): Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In: *Attention and awareness in foreign language learning*, 9, pp. 1–63.
- Sharwood-Smith, Michael (1993): Input Enhancement in Instructed SLA: Theoretical Bases. *Studies in Second Language Acquisition*, 15 (2), pp. 165–179. doi.org/10.1017/S0272263100011943.
- Spannagel, Christian (2023): Hat ChatGPT eine Zukunft in der Mathematik? In: *Mitteilungen der Deutschen Mathematiker-Vereinigung*, 31 (3), S. 168–172. doi.org/10.1515/dmvm-2023-0055.
- Steinert, Steffen/Avila, Karina E./Ruzika, Stefan/Kuhn, Jochen/Küchemann, Stefan (2023): Harnessing Large Language Models to Enhance Self-Regulated Learning via Formative Feedback (arXiv:2311.13984). arxiv.org/abs/2311.13984.
- Thyssen, Christoph/Finger, Alexander/Becker, Sebastian/Huwer, Johannes/Bruckermann, Till/Kremser, Erik/Meier, Monique/Thoms, Lars-Jochen/von Kotzebue, Lena (2023): Die Lehrkräftebildung im Wandel: Über welche digitalen Kompetenzen sollten angehende Lehrkräfte (in den Naturwissenschaften) verfügen? In: Aufenanger, Julian/Bigos, Michael (Hrsg.): *Digitalisierung in der Lehrer:innenbildung*. Wiesbaden: Springer VS, S. 88–100.
- Thyssen, Christoph/Thoms, Lars-Jochen/Kremser, Erik/Finger, Alexander/Huwer, Johannes/Becker, Sebastian (2020): Digitale Basiskompetenzen in der Lehrerbildung unter besonderer Berücksichtigung der Naturwissenschaften. In: Beißwenger, Michael/Bulizek, Björn/Gryl, Inga et al. (Hrsg.): *Digitale Innovationen und Kompetenzen in der Lehramtsausbildung*. Duisburg: Universitätsverlag Rhein-Ruhr, S. 77–98. dx.doi.org/10.17185/dupublico/73330.
- Van Brummelen, Jessica/Lin, Phoebe (2020): Engaging Teachers to Co-Design Integrated AI Curriculum for K-12 Classrooms (arXiv:2009.11100). arxiv.org/abs/2009.11100.
- Vygotskiĭ, Lev Semjonowitsch/Kozulin, Alex (1986): *Thought and language* (Translation newly revised and edited). Cambridge: MIT Press.
- Weiss, Zarah/Dittrich, Sabrina/Meurers, Detmar (2018): A Linguistically-Informed Search Engine to Identify Reading Material for Functional Illiteracy Classes. In: Pilán, Ildikó/Volodina, Elena/Alfter, David/Borin, Lars (Eds.): *Proceedings of the 7th workshop on NLP for Computer Assisted Language Learning*. Stockholm: LiU Electronic Press, pp. 79–90. aclanthology.org/W18-7109.
- Wildemann, Anja/Bien-Miller, Lena (2023): *Sprachbewusstheit: Perspektiven aus Forschung und Didaktik*. Wiesbaden: Springer Fachmedien. doi.org/10.1007/978-3-658-39229-1.
- Witzel, Björn/Görgen-Rein, Ruth/Galuschka, Katharina/Huemer, Sini/Corvacho Del Toro, Irene/Schulte-Körne, Gerd/Moll, Kristina (2024): Digital game-based spelling intervention for children with spelling deficits: A randomized controlled trial. In: *Learning and Instruction*, 89, p. 101842. doi.org/10.1016/j.learninstruc.2023.101842.
- Wood, David/Bruner, Jerome S./Ross, Gail (1976): The role of tutoring in problem solving. In: *Journal of Child Psychology and Psychiatry*, 17 (2), pp. 89–100. doi.org/10.1111/j.1469-7610.1976.tb00381.x.

Informationen zu den Autor*innen



Luzia Leifheit* ist Tenure-Track-Professorin für Digitalisierung mit Schwerpunkt Didaktik der Algorithmik und Data Science an der Pädagogischen Hochschule Schwäbisch Gmünd. In ihrer Forschung setzt sie sich mit der Wirksamkeit unterschiedlicher, oft spielbasierter und körperlicher, Lernansätze für die Förderung von Motivation und Kompetenzen in der informatischen Bildung von der Grundschule bis zur Universität auseinander. * *Luzia Leifheit teilt sich die Erstantorinnenschaft mit Denise Loefflad*

luzia.leifheit@ph-gmuend.de

Denise Loefflad* ist wissenschaftliche Mitarbeiterin am Leibniz-Institut für Wissensmedien in Tübingen. In dem Projekt WoLKE fokussiert sie sich auf den Bereich der Computerlinguistik. Innerhalb des Projektes bewertet sie die Leistungsfähigkeit bestehender digitaler Tools und erforscht die Entwicklung neuer KI-basierter Tools, um Lehr-Lernformate zu unterstützen. * *Denise Loefflad teilt sich die Erstantorinnenschaft mit Luzia Leifheit*



d.loefflad@iwm-tuebingen.de



Sina Belschner promoviert als wissenschaftliche Mitarbeiterin im Projekt WoLKE (wolke.schule) an der Pädagogischen Hochschule Schwäbisch Gmünd. Hier fokussiert sich Frau Belschner auf den Bereich der MINT-Didaktik und wirkt mit an der Entwicklung einer Lehrveranstaltung zum reflektierten Einsatz von KI-Tools im Unterricht.

sina.belschner@ph-gmuend.de

Benedikt Beuttler ist Postdoc im Projekt WoLKE (wolke.schule). In dem Verbundprojekt zwischen der PH Ludwigsburg, der PH Schwäbisch-Gmünd und dem am Leibniz-Institut für Wissensmedien (IWM) Tübingen entwickelt das Team fachlich und didaktisch passgenaue Lehr-Lernformate für den reflektierten Einsatz von KI-Tools im Unterricht, um bereits angehende Lehrkräfte auf die neuen Herausforderungen und Chancen vorzubereiten. Herr Beuttler fokussiert sich in seiner Forschung auf die Digitalisierung im Bildungsbereich und auf Systeme mit Mensch-Computer Interaktionen, wie z.B. digitales spielbasiertes Lernen, intelligente Tutorssysteme und digitale Assessments.



benedikt.beuttler@ph-ludwigsburg.de



Jan Winkelmann ist Tenure-Track-Professor für naturwissenschaftliche Bildung an der Pädagogischen Hochschule Schwäbisch Gmünd und leitet dort das Zentrum für naturwissenschaftliche Bildung (www.zfnb.de). Den Rahmen seiner Arbeit bilden schwierigkeiterzeugende Merkmale im naturwissenschaftlichen Unterricht. Dabei fokussiert sein Forschungsinteresse auf die Bedeutung des Experiments und des Modells im naturwissenschaftlichen Erkenntnisprozess. Aktuell untersucht er den expliziten Umgang mit Idealisierungen in Erklärungen naturwissenschaftlicher Phänomene sowie Implikationen auf das Vertrauen in Wissenschaft.

jan.winkelmann@ph-gmuend.de



Prof. Dr. Detmar Meurers leitet die Arbeitsgruppe „Sprache und KI in der Bildung“ am Leibniz-Institut für Wissensmedien (IWM) Tübingen. Seine Forschungsinteressen liegen an der Schnittstelle der Computerlinguistik und der empirischen Bildungsforschung. Er erforscht und entwickelt KI-Methoden im Bildungs- und Zweitspracherwerb, von Intelligenten Tutorssystemen bis zur Analyse sprachlicher Komplexität und Lesbarkeit.

d.meurers@iwm-tuebingen.de

Heiko Holz ist Tenure-Track Professor für Informatik und Informatikdidaktik mit Schwerpunkt Digitalisierung im Bildungsbereich an der Pädagogischen Hochschule Ludwigsburg. Seine Forschung fokussiert sich darauf, Bedingungen der erfolgreichen Digitalisierung im Bildungsbereich zu identifizieren und in digitalen Lehr-Lernsystemen zu fördern, einschließlich Mensch-Computer Interaktion, digitales spielbasiertes Lernen, intelligente Tutorssysteme und digitale Interventionen und Assessments.



heiko.holz@ph-ludwigsburg.de

Zitationshinweis:

Leifheit, Luzia/Loefflad, Denise/Belschner, Sina/Beuttler, Benedikt/Winkelmann, Jan/Meurers, Detmar/Holz, Heiko (2024): KI im Unterricht – Entwicklung von Lehrveranstaltungen für Lehramtsstudierende der Sprach- und MINT-Fächer. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 24, S. 1–19. doi.org/10.21240/lbzm/24/08.