Bundesministerium für Bildung und Forschung

Freie Universität Berlin

Friedrich-Alexander-Universität Erlangen-Nürnberg

ABSCHLUSSBERICHT: PetraKIP-Projekt

Zuwendungsempfänger: Freie Universität Berlin, Didaktik der Informatik sowie Dahlem Center for Machine Learning and Robotics (Teilprojekt FU Berlin) FAU Erlangen-Nürnberg, Lehrstuhl für Schulpädagogik mit dem Schwerpunkt empirische Unterrichtsforschung (Teilprojekt FAU Erlangen-Nürnberg)	**Förderkennzeichen:** 16DHB4019 (Teilprojekt FAU)
Vorhabenbezeichnung: PetraKIP: Persönliches transparentes KI-basiertes Portfolio für die Lehrerbildung	
Laufzeit des Vorhabens: 01.03.2021 bis 31.08.2024	
Berichtszeitraum: 01.03.2021 – 31.08.2024	

Prof. Dr. Michaela Gläser-Zikuda
Lehrstuhl für Schulpädagogik m. d. Schwerpunkt empirische Unterrichtsforschung Regensburger Str. 160, 90478 Nürnberg
Friedrich-Alexander-Universität Erlangen Nürnberg
E-Mail: michaela.glaeser-zikuda@fau.de; Telefon: +49 911 5302 95586

Projektkoordination:
Prof. Dr. Ralf Romeike
Didaktik der Informatik
Königin-Luise-Str. 24-26,14195 Berlin
Freie Universität Berlin
E-Mail: ralf.romeike@fu-berlin.de; Telefon: +49 30 838 65994

Weitere Beteiligte:
Prof. Dr. Christoph Benzmüller und Prof. Dr. Tim Landgraf
Dahlem Center for Machine Learning and Robotics
Arnimallee Str. 7, 14195 Berlin
Freie Universität Berlin
E-Mail: c.benzmueller@fu-berlin.de; Telefon: +49 30 838 52485
E-Mail: tim.landgraf@fu-berlin.de; Telefon: +49 30 838 75114

Inhaltsverzeichnis

I. KURZBESCHREIBUNG 3

II. EINGEHENDE DARSTELLUNG 4

1. Ziele und Ergebnisse 4

AP3: Didaktische Konzeptionierung 4

AP4: Evaluation 6

AP5: Projektkoordination / Kommunikation 14

2. Positionen des zahlenmäßigen Nachweises 15

3. Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten 15

4. Verwertbarkeit der Ergebnisse 15

5. Erfolgte und geplante Veröffentlichungen 15

III. ERFOLGSKONTROLLBERICHT 19

1. Wirtschaftliche und wissenschaftliche Ergebnisse des Vorhabens, erreichte Nebenergebnisse und die gesammelten wesentlichen Erfahrungen 19

2. Fortschreibung des Verwertungsplans 19

3. Angaben zu Arbeiten, die zu keiner Lösung geführt haben 19

4. Angaben über die Einhaltung der Ausgaben- und der Zeitplanung 19

I. Kurzbeschreibung

Trotz wachsender Potenziale von Anwendungen Künstlicher Intelligenz zeichnet sich für ihren Einsatz in der Hochschullehre ein Dilemma ab: Einerseits werden den Technologien umfangreiche Chancen, bspw. im Hinblick auf eine individualisierte Gestaltung von Studienangeboten zugeschrieben, um u.a. die Studienzufriedenheit zu fördern und Studienabbruchneigungen entgegenzutreten; andererseits lassen sich zunehmend studentische Abwehrhaltungen beobachten, die mit technologiebezogenen Unsicherheiten einhergehen. Eine besondere Herausforderung stellt dies für die Lehrkräftebildung dar: Während KI-Technologien durch die Verstärkung der sich bereits als erfolgreich erwiesenen Maßnahmen individualisierter Studienunterstützung einen signifikanten Beitrag zur Lösung des Lehrkräftemangels leisten könnten, lässt sich gerade bei Lehramtsstudierenden eine größere Skepsis gegenüber neuen Technologien wahrnehmen – was sich durch ihre zukünftige Multiplikatorenrolle nochmal auf weitere Generationen manifestieren könnte. Im Projekt PetraKIP werden Technologieentwicklung und Bildung als miteinander in Beziehung stehende Aufgaben verstanden. Ziel des Projekts ist die interdisziplinäre Erforschung von KI-Methoden (insbes. Textanalyse, Maschinelles Lernen, Benutzermodellierung, kontrollierter Dialog) für die Lehrkräftebildung und deren Realisierung innerhalb eines “KI-Portfolios”. Die Implementierung und Evaluation erfolgen innerhalb eines E-Portfolios, um die Verwertbarkeit der Ergebnisse sicherzustellen. Entwicklungsbegleitend wird untersucht, inwiefern Transparenz und Bildungsangebote hinsichtlich der eingesetzten KI-Verfahren zur Bildung und Mündigkeit der Lehramtsstudierenden und somit dem erfolgreichen Einsatz von KI Technologien beitragen können.

II. Eingehende Darstellung

Im Folgenden wird das Teilprojekt der FAU des Vorhabens PetraKIP näher beschrieben. Hierzu werden die Ziele und die daraus resultierenden Ergebnisse der einzelnen Arbeitspakete (AP) beschrieben.

1. Ziele und Ergebnisse

Die nachstehenden Ausführungen zu den Projektarbeiten und deren Ergebnissen betreffen die Arbeitspakete (AP) 3, 4 und 5 des Projekts. Andere Aspekte, die hier nicht erwähnt werden, wurden vom Teilprojekt der FU Berlin durchgeführt.

AP3: Didaktische Konzeptionierung

- UAP 3.1: Ermittlung der KI-bezogenen Lehr-Lernvoraussetzungen
- UAP 3.5: Schulung der kooperierenden Dozent:innen

AP4: Evaluation

- UAP 4.1. Zusammenstellung von stand. Skalen, Ermittlung der portfoliobezogenen Lehr-Lern-Voraussetzungen (Stud., Doz., Stakeh.)
- UAP 4.2: Formative und summative Evaluation
- UAP 4.3: Delphi-Studie mit Expert:innen verschiedener Disziplinen, Ableitung von Überarbeitungsempfehlungen für das KI-Portfolio
- UAP 4.4 Formative Evaluation (Entwicklungsbegleitende formative Evaluation bei größeren Studierendengruppen (u.a. KI-Akzeptanz, Reflexionsqualität, Kompetenz, Selbstwirksamkeit, Studienzufriedenheit) und summative Evaluation

AP5: Projektkoordination / Kommunikation

- UAP 5.2: Einbindung der assoziierten Partner und weiterer Akteur:innen in der Lehrkräftebildung (Workshops etc.)
- UAP 5.3: Auftakt-, Zwischen- und Abschlussworkshops
- UAP 5.4: Dokumentation, Aufbereitung, Vorbereitung von Publikationen und Transfers

AP3: Didaktische Konzeptionierung

Die FU Berlin trägt die Verantwortung für das Arbeitspaket drei (AP3), welches sich auf die didaktische Entwicklung der vorgesehenen App konzentriert. Dieses Arbeitspaket umfasst Unterarbeitspakete, die einerseits die Erfassung der KI-relevanten Lehr- und Lernbedingungen aus Sicht von Studierenden, Lehrenden und Interessensgruppen beinhalten und andererseits die Fortbildung von Lehrkräften umfassen. An der Umsetzung dieser Unterarbeitspakete ist auch die FAU Erlangen-Nürnberg beteiligt.

UAP 3.1: Ermittlung der KI-bezogenen Lehr-Lernvoraussetzungen

Um die KI-bezogenen Lehr-Lernvoraussetzungen zu ermitteln, wurden Studierende per Online-Fragebogen befragt. Auf der Basis dieser Ergebnisse erfolgt die Adaption, Weiterentwicklung und Implementierung der KI-Lehr-Lernmethoden, die Auswertung, Materialaufbereitung sowie die Transfervorbereitung mit anschließender Schulung der kooperierenden Dozierenden.

Die erste Befragung erfolgte im Sommersemester 2021. Eingebettet in die erste Version des auf selbstreguliertes Lernen und professionsbezogenes Reflektieren ausgerichteten Portfoliokonzepts wurden zu Beginn der Vorlesungszeit unter anderem KI-bezogenen Lehr-Lernvoraussetzungen erhoben. Das E-Portfolio wurde in die auf ILIAS basierende Lernplattform StudON der FAU implementiert und anschließend evaluiert. Kernbestandteil des Portfolios sind zwei Themenbereiche einer schulpädagogischen Vorlesung, die so aufbereitet wurden, dass die Studierenden diese nach einer kurzen Einweisung im Rahmen der Lehrveranstaltung selbstständig bearbeiten konnten. Die Studierenden beschäftigten sich im Rahmen der Portfolioarbeit mit einem Fallbeispiel, wobei ihnen hierbei diverse, hierarchisch angelegte Prompts zur Verfügung standen. Um die erarbeitete Lösung zum Fallbeispiel einschätzen zu können, erhielten die Studierenden nach der Einreichung als Rückmeldung einen Erwartungshorizont in unterschiedlichen Formaten (Stichpunkte mit und ohne Erläuterungen, Fließtext mit und ohne Erläuterungen sowie didaktisch konzipiertes Lernmaterial mit und ohne Erläuterungen). Abschließend reflektierten die Studierenden ihren Arbeitsprozess sowie ihre erstellte Lösung zum Fallbeispiel.

Im Anschluss erfolgte in Form einer zweiten Fragebogenerhebung eine Evaluation der eingesetzten Lehr-Lernmethode. Dieser Fragebogen enthielt neben weitere soziodemographischer Daten unter anderem auch standardisierte Skalen zum Lehr-Lernprozess, zum Nutzungsverhalten sowie zur Zufriedenheit mit der Umsetzung des E-Portfolio-Konzepts. Darüber hinaus wurden ebenfalls standardisierte Skalen eingesetzt, um die Akzeptanz und prospektive Nutzung von Künstlicher Intelligenz (KI) aus Sicht der Studierenden zu erfassen.

Zur Validierung der Instrumente wurde im Wintersemester 2021/2022 die Erhebung in identischer Form wiederholt; anschließend erfolgte für weitere Durchläufe eine geringfügige Adaption der Fragebögen und eingesetzten Portfoliomaterialien. Die ermittelten Daten der Testkonstruktion lieferten ausnahmslos gute bis sehr gute Werte.

Der Milestone, die KI-bezogenen Lehr-Lernvoraussetzungen der Studierenden zu erfassen, wurde somit vollumfänglich erreicht.

Im Sommersemester 2024 führte das FAU-Team eine Fragebogenerhebung von Hochschullehrenden an der FU Berlin sowie der FAU Erlangen-Nürnberg durch. Der Online-Fragebogen umfasste standardisierte Skalen in englischer Sprache: KI-Akzeptanz mit den Subdimensionen berufsbezogene Relevanz von KI, subjektive Norm, wahrgenommene Nützlichkeit, wahrgenommene Benutzerfreundlichkeit, wahrgenommener Freude, KI-Angst sowie auf KI bezogenes intelligentes TPACK (Technological pedagogical content knowledge) mit den Dimensionen intelligentes TPK (technological pedagogical knowledge), intelligentes TCK (technological content knowledge), intelligentes TK (technological knowledge) sowie KI-bezogene Ethik.

An der Erhebung nahmen insgesamt 113 Lehrende aus verschiedenen Fachbereichen der FAU teil, wobei 52 gültige Fragebögen letztlich für die Analysen verwendet werden konnten. Die Ergebnisse zeigen, dass die berufsbezogene Relevanz von KI eine entscheidende Rolle für die KI-Akzeptanz von Lehrenden spielt. Die befragten Lehrenden gaben insgesamt an über ein relativ hohes Maß an technologischem, pädagogischem Wissen und ethischem Bewusstsein zu verfügen.

UAP 3.5: Schulung der kooperierenden Dozent:innen

PapagAI wurde im Sommersemester 2023 und im Wintersemester 2023/24 den Lehrpersonen des Fachbereichs Schulpädagogik und den Mitgliedern des Zentrums für Lehrerinnen- und Lehrerbildung der FAU vorgestellt.

Im März 2023 präsentierten Projektmitarbeitende Forschenden und Lehrenden verschiedener europäischer Universitäten im Rahmen der Fachtagung „European Teacher Education Network“ in Nürnberg das Projekt PetraKIP und das KI-gestützte Feedbacksystem für reflexives Schreiben.

Im Juli 2023 beteiligten sich die Mitarbeitenden des PetraKIP-Projekts am „DIGILLAB Inspiration Day“ der FAU in Nürnberg, der sich eingehend mit der Rolle von KI-Technologien in der Bildungslandschaft auseinandersetzte. Besondere Aufmerksamkeit wurde der möglichen Rolle von KI-Technologien bei der Förderung digitaler Kompetenzen für Hochschullehrende und Lehrkräfte an Schulen gewidmet. Im Rahmen der Veranstaltung fanden Diskussionen über den Nutzen von KI für digitale Lehr- und Lernressourcen im Hochschulsektor statt. Ein zentraler Fokus lag auf den Einflüssen der KI auf die Lehrkräftebildung , wobei unterschiedliche Anwendungsszenarien und Forschungsansätze intensiv diskutiert wurden.

Im Juni 2024 wurde auf der FAU-Tagung „Let´s Talk – KI in der Lehre“ das im Rahmen des PetraKIP-Projekts erarbeitete E-Portfolio einer großen Gruppe von FAU-Dozierenden nähergebracht. Neben Fragen der Implementation wurden die bisherigen Erfahrungen sowie mögliche Chancen durch KI-Systeme im Zusammenhang mit systematischen Rückmeldungen an Studierende vorgestellt.

AP4: Evaluation

Im Rahmen des 4. Arbeitspakets (AP4) wurde im Sommersemester 2023 und im Wintersemester 2023/2024 jeweils eine quasi-experimentelle Studie durchgeführt, um die unterschiedlichen Wirkungen von menschlichem Feedback im Vergleich zu KI-generierten-Feedback zu untersuchen(Vergleich von Feedback durch Hochschuldozent:innen mit Feedback durch die im Projekt entwickelte KI-basierte App). Des Weiteren wurde eine Expert:innenbefragung im Rahmen einer Delphi-Studie durchgeführt. In dieser wurden Expert:innen aus den Fachbereichen Bildungs- und Erziehungswissenschaft (Schul- und Medienpädagogik, Päd. Psychologie) sowie der Informatik befragt. Die Ergebnisse wurden im Projektteam besprochen und für weitere Entwicklungen berücksichtigt. Darüber hinaus wurde eine Online-Umfrage zur Einstellung von Hochschuldozent:innen zu KI-basierten Anwendungen geplant.

UAP 4.1. Zusammenstellung von stand. Skalen, Ermittlung der Portfolio-bezogenen Lehr-Lern-Voraussetzungen (Stud., Doz., Stakeh.)

In einem ersten Arbeitsschritt wurde ein Skalenhandbuch erstellt, das folgende Konstrukte beinhaltet: Soziodemographischen Daten, standardisierten Skalen zur Erfassung der Persönlichkeit (Big Five), der Einstellung zum Studium (studienbezogene Selbstwirksamkeit, Lern- und Leistungsmotivation, Studienzufriedenheit, Relevanz des Studiums, Berufswahlmotivation und Prokrastination), zu selbstreguliertem Lernen (metakognitive Strategien, kognitive Lernstrategien und ressourcenbezogene Strategien), zu Emotionen (Perceived Stress), zu digitaler Kompetenz (Suchen und Verarbeiten, Kommunizieren und Kollaborieren, Produzieren und Präsentieren, Schützen und sicher Agieren, Problemlösen und Handeln, Analysieren und Reflektieren, Unterrichten und Implementieren) sowie zu Einstellungen und Akzeptanz der Studierenden von Künstlicher Intelligenz (Perceived Usefulness, Perceived Ease of Use, AI Self-Efficacy, Perceptions of External Control, AI Anxiety, Perceived Enjoyment, Subjective Norm, Job Relevance und Behavioral Intention). Ergänzt wurden Skalen zur Evaluation der App-Nutzung (Einstellung zur Portfolioarbeit, Evaluation des E-Portfolios und Kurzskala zur Lehrevaluation und Wissenstest).

Um Lehr-Lernvoraussetzungen der Studierenden zu ermitteln, wurde das unter UAP 3.1: Ermittlung der KI-bezogenen Lehr –Lernvoraussetzungen beschriebene Mini-E-Portfolio (ohne KI) als Basis herangezogen, um die Studierenden zur Auseinandersetzung und zur Reflexion mit professionsbezogenen Themen anzuregen. Entsprechende Prompts wurden entwickelt und über die Lernplattform schriftlich angeboten.

Die im Sommersemester erstmals ermittelten Evaluationsergebnisse flossen in die Optimierung des E-Portfolios für die Nutzung im Wintersemester 2021/2022 ein.

Der Milestone der Zusammenstellung von standardisierten Skalen und der Ermittlung der portfoliobezogenen Lehr-Lernvoraussetzungen war für März bis August 2021 eingeplant und wurde somit vollständig erreicht.

UAP 4.2 / UAP 4.4: Formative und summative Evaluation

Während der Projektlaufzeit fanden mehrere formative und summative Erhebungen statt. Eingebettet waren diese in ein E-Portfolio, das Teil der Lehrveranstaltung "Einführung in die Schulpädagogik" an der FAU war. Die Studierenden verschiedener Lehramtsstudiengänge hatten die Möglichkeit, lehrveranstaltungsbegleitend an den Erhebungen teilzunehmen. Diese Erhebungen beinhalteten diverse Fragebögen, einen Wissenstest, Interviews, Reflexionsaufgaben sowie fachspezifische Aufgabenstellungen. Nachfolgend werden die Erhebungen während der Semester chronologisch und detailliert dargestellt.

Erhebung im Sommersemester 2021

Fragebogen:

Der Fragebogen im Sommersemesters 2021 konzentrierte sich auf die Validierung der vorab festgelegten Skalen (siehe UAP 4.1). Die Ergebnisse zeigen, dass die verwendeten Skalen umfassende empirische Untersuchungen ermöglichen. Um den Wissenserwerb der Studierenden im Verlauf des Lernprozesses mit dem Mini-E-Portfolios untersuchen zu können, wurde ein speziell konzipierter fachbezogener Wissenstest in den Fragebogen integriert. Es nahmen 24 Studierende an dieser ersten Erhebung teil. Der Fragebogen wurde einmalig eingesetzt.

Interview:

Um die Benutzerfreundlichkeit und Akzeptanz dieser ersten E-Portfolio-Version zu erfassen, fanden ergänzend zum Fragebogen Thinking-Aloud-Interviews mit Studierenden statt. Ziel dieser Interviews war, Hinweise zur Implementierung des E-Portfolios zu ermitteln und von den Studierenden Verbesserungsvorschläge zu erhalten.

Erhebung Wintersemester 2021/2022

Fragebogen:

Auf Grundlage der Fragebogenergebnisse des Sommersemesters 2021 sowie der ausgewerteten Interviews wurde das E-Portfolio-Konzept überarbeitet und im Wintersemester 2021/22 in modifizierter Form eingesetzt. Anders als im Sommersemester 2021 fand nun vor und nach der Anwendung des E-Portfolios eine Fragenbogenerfassung statt.

Im Rahmen der ersten Befragung wurden zur Erfassung von individuellen Einstellungen und Kompetenzen sowie der Wahrnehmung von Studienbedingungen folgende Skalen eingesetzt:

Digitale Kompetenz, Akzeptanz von KI, Selbstregulation und Lernstrategien, studienbezogene Selbstwirksamkeit, Lern- und Leistungsmotivation, Prokrastination, Emotionen, Persönlichkeit, Berufswahlmotivation, Studienzufriedenheit sowie Relevanz des Studiums. Teilgenommen haben an dieser Befragung insgesamt 708 Lehramtsstudierende, davon waren 454 Fragebögen vollständig.

Beim zweiten Befragungszeitpunkt am Ende des Wintersemesters 2021/2022 wurden Skalen genutzt, um die Einstellung zur Portfolioarbeit zu erfassen, aber auch um das Mini-E-Portfolio in Hinblick auf die Motivation, den Lernfortschritt, die Umsetzung der Aufgaben, die Zufriedenheit mit sowie die Nutzung und Nützlichkeit der Prompts zu evaluieren. Ebenso wurde der zur Verfügung gestellte Erwartungshorizont begutachtet. Auch Fragen zur Lehrveranstaltung sowie zum bereits genannten Wissenstest zu den beiden Themengebieten waren Bestandteil der Befragung. An dieser nahmen 529 Studierende teil; 245 Fälle waren vollständig.

Die Daten wurden mithilfe eines Strukturgleichungsmodells analysiert, das mit der Open-Source-Statistiksoftware R durchgeführt wurde. Es wurden die wahrgenommene Benutzerfreundlichkeit ($\beta = 0{,}297^{***}$) und die wahrgenommene Nützlichkeit ($\beta = 0{,}501^{***}$) als primäre Faktoren identifiziert, welche die Absicht von Lehramtsstudierenden KI zu nutzen, vorhersagen. Darüber hinaus zeigten die Ergebnisse der latenten Mittelwertdifferenzen, dass sich zwei Konstrukte, KI-Angst ($z = -3{,}217^{**}$) und wahrgenommene Freude ($z = 2{,}556^{*}$), signifikant hinsichtlich des Geschlechts voneinander unterscheiden. Zudem wurden Pfadanalysen berechnet und es zeigte sich, dass die Pfade von der KI-Angst zur wahrgenommenen Benutzerfreundlichkeit ($p = 0{,}018^{*}$) und von der wahrgenommenen Benutzerfreundlichkeit zur wahrgenommenen Nützlichkeit ($p = 0{,}002^{**}$) durch das Geschlecht moderiert werden.

Interviews:

Im Wintersemester 2021/22 wurden ergänzend zur Fragebogenerhebung auch Leitfadeninterviews mit 22 Studierenden durchgeführt. Hierbei lag der Fokus vor allem auf der Nutzung digitaler Medien für den Lernprozess der Studierenden und der Akzeptanz von Künstlicher Intelligenz.

Die inhaltsanalytische Auswertung der Interviews liefert aufschlussreiche Einblicke in die Nutzung und Wahrnehmung digitaler Medien sowie den Umgang mit KI-Technologien im Lernkontext. Von den befragten Studierenden nutzen 21 Personen regelmäßig digitale Medien, wobei die Einstellung ihnen gegenüber mehrheitlich positiv ist. Allerdings äußern fünf Studierende auch kritische Perspektiven. Im Umgang mit digitalen Medien berichten die meisten Befragten sowohl von unterstützenden als auch von herausfordernden Erfahrungen. Während einige Plattformen wie StudyDrive und YouTube als hilfreich wahrgenommen werden, geben drei Studierende an, dass digitale Medien ihnen im Lernprozess keine Unterstützung bieten.

Neun Studierende haben bereits Erfahrungen mit Online-Schulungsprogrammen gesammelt, insbesondere durch Angebote wie Kurse der Virtuellen Hochschule Bayern oder Anwendungen, die sie in früheren Ausbildungsphasen genutzt haben. Ein Großteil der Befragten (18 von 22) vermutet, dass KI-Technologien in den von ihnen genutzten digitalen Medien integriert sind, und gibt an, solche Technologien regelmäßig zu verwenden. Dabei werden Instagram, Netflix, Alexa/Siri und Google Maps häufig genannt. Die Integration von KI-Technologien ruft gemischte oder neutrale Gefühle hervor. Während Praktikabilität und Mehrwert oft betont werden, äußern einige Studierende Unsicherheiten und Sorgen über mögliche zukünftige Entwicklungen. Vier Studierende bewerten KI-Technologien ausschließlich positiv, fünf hingegen ausschließlich negativ.

Für die Lehrer:innenbildung wird der Einsatz von KI-Technologien von einer deutlichen Mehrheit (17 von 22 Studierende) positiv gesehen. Zudem würden 18 Studierende KI-Technologien auch für ihren eigenen Lernprozess nutzen. Hinsichtlich der gewünschten Art von KI-Feedback wird vor allem der Wunsch nach Verbesserungsvorschlägen und Rückmeldungen zu inhaltlichen Fehlern geäußert. Zusätzlich wünschen sich die Studierenden Lob für richtige Erarbeitungen sowie eine Orientierungshilfe in Form von Noten oder Lernstandseinschätzungen. Viele Studierende betonen außerdem den Wunsch, dass KI langfristige Fortschritte im Lernprozess erkennt und diese rückmeldet. Eine abschließende Bewertung rein durch die KI wird jedoch abgelehnt.

Das Mini-Portfolio hatte nach Meinung der Interviewten ebenfalls Auswirkungen auf das Lernverhalten der Studierenden. Zehn Befragte berichten von einer Veränderung in ihrem selbstregulierten Lernen und 16 Studierende geben an, dass das Portfolio ihre Motivation positiv beeinflusst hat.

Zusammenfassend verdeutlichen die Ergebnisse der Interviews die Relevanz digitaler Medien und KI-Technologien aus Sicht der Studierenden. Die Mehrheit der Befragten zeigt eine grundsätzlich positive Haltung gegenüber der Integration von KI in Bildungskontexten, formuliert jedoch differenzierte Bedürfnisse und Erwartungen an deren Einsatz. Das Mini-Portfolio scheint aus Sicht einiger Studierender ein geeignetes Instrument zu sein, um selbstreguliertes Lernen und Motivation zu fördern.

Reflexionsschreiben:

Im Rahmen des E-Portfolios wurden die Studierenden aufgefordert zwei Reflexionstexte zu den zuvor bearbeiteten Themenbereichen anzufertigen. Auf diese Weise entstanden 516 Reflexionen zum Bereich Klassenmanagement und 506 zum Bereich Pädagogische Diagnostik. Aus den Texten wurden randomisiert 100 der 506 vorliegenden Reflexionstexte zum Thema Pädagogische Diagnostik analysiert. Diese Reflexionstexte wurden einerseits quantitativ mit der Software LIWC (Linguistic Inquiry and Word Count ist der Goldstandard bei Software zur Analyse der Wortverwendung), andererseits inhaltsanalytisch qualitativ mit einem deduktiv orientierten Kategoriensystem händisch von zwei Mitarbeiter:innen analysiert. Für die Computerlinguistik wurde die LIWC-Software verwendet, um die linguistischen Merkmale des reflexiven Textes in den morphologischen, syntaktischen und semantischen Ebenen zu extrahieren. Für die Qualitative Inhaltsanalyse wurden die Reflexionstexte theoriegeleitet vier modifizierten Reflexionsniveaus (vgl. Hatton & Smith, 1995; Fütterer, 2019) zugeordnet. Die Einstufung der Qualität der Reflexionstexte erfolgt basierend auf der Zusammenführung der Ergebnisse aus beiden Verfahren. Die Ergebnisse zeigten, dass die Qualität der Reflexionen der angehenden Lehrkräfte eher niedrig war.

Erhebung Sommersemester 2022

Fragebogen:

Der erste Fragebogen für das Sommersemester 2022 gleicht dem des Wintersemesters 2021/22. In der ersten Befragung wurden verschiedene Skalen verwendet, um individuelle Einstellungen und Kompetenzen sowie die Wahrnehmung von Studienbedingungen zu erfassen. Zu diesen standardisierten Skalen gehörten digitale Kompetenz, Akzeptanz von KI, Selbstregulation und Lernstrategien, studienbezogene Selbstwirksamkeit, Lern- und Leis-

tungsmotivation, Prokrastination, Emotionen, Persönlichkeitsmerkmale, Berufswahlmotivation, Studienzufriedenheit sowie Relevanz des Studiums. Insgesamt nahmen 295 Lehramtsstudierende an dieser Befragung teil, 183 Fragebögen wurden vollständig ausgefüllt.

Am zweiten Befragungszeitpunkt zum Ende des Sommersemesters 2022 wurden verschiedene Skalen genutzt, um die Nutzererfahrung, die Einstellung zur Portfolio-Arbeit und das Mini-E-Portfolio in Bezug auf verschiedene Aspekte, wie Motivation, Lernfortschritt, Umsetzung der Aufgaben, Zufriedenheit, Nutzung und Nützlichkeit der Prompts zu evaluieren. Außerdem wurden die im digitalen Mini-E-Portfolio eingesetzten Materialien, wie z.B. ein Erwartungshorizont für die bearbeiteten Aufgaben, von den Studierenden bewertet; es wurden Angaben zur Lehrevaluation gemacht sowie ein Wissenstest zu den beiden Themengebieten der Lehrveranstaltung eingesetzt. An der Befragung nahmen insgesamt 226 Studierende teil, 140 Fragebögen waren gültig. Diese Daten wurden aufgrund der Ähnlichkeit zum Vorsemester bislang nicht ausgewertet.

Reflexionsschreiben:

Im Sommersemester 2022 wurden zu den Themen Klassenmanagement und Pädagogische Diagnostik jeweils 74 bzw. 97 Reflexionsschreiben erfasst.

Erhebungen Wintersemester 2022/2023

Fragebogen:

Im Wintersemester 2022/2023 wurde der Fragebogen auf die KI-Thematik fokussiert. Er wurde von den Kooperationspartnern der FU Berlin konzipiert und bei Lehramtsstudierenden der FAU Erlangen-Nürnberg eingesetzt. Der Fragebogen bestand aus drei Teilen, wobei in den ersten beiden Abschnitten die persönliche Einstellung der Studierenden zur KI sowie ihre Einschätzung zur Entwicklung dieser Technologie abgefragt wurde. Der letzte Teil bestand aus offenen Fragen, die den Studierenden die Möglichkeit gaben, ihre Meinungen und Ansichten zum Thema KI frei zu äußern. Insgesamt nahmen 822 Personen an dieser Befragung teil, wobei 411 Fragebögen vollständig waren. Für die Erhebung im Wintersemester 2022/2023 wurde der Fragebogen des Sommersemesters 2022 ohne Änderungen übernommen. Bei dieser Befragung waren 348 Fragebögen gültig. Die Auswertung dieser Erhebungen oblag dem Zuständigkeitsbereich der FU Berlin.

Reflexionsschreiben:

Im Wintersemester 2022/2023 wurden im Themenbereich Klassenmanagement 383 und im Bereich Pädagogische Diagnostik 384 Reflexionsschreiben erhoben.

Zu Beginn der Analyse wurde eine Zufallsauswahl von 200 Reflexionstexten (jeweils 100 aus beiden Themenbereichen) vorgenommen. Die so gewonnenen 198 Texte (2 missing values) wurden quantitativ mittels der LIWC-Software (Linguistic Inquiry and Word Count) und qualitativ mittels einer deduktiv orientierten Qualitativen Inhaltsanalyse (nach Mayring & Gläser-Zikuda, 2008) von zwei Mitarbeiterinnen ausgewertet. Die Kodierungen, welche in Anlehnung an die qualitative Inhaltsanalyse vorgenommen wurden, orientierten sich an einem theoriegeleiteten Kategoriensystem mit vier Reflexionsebenen (vgl. Hatton & Smith, 1995; Fütterer, 2019). Für die computerlinguistische Analyse wurden die linguistischen Merkmale der Reflexionstexte auf morphologischer, syntaktischer und semantischer Ebene mit der LIWC-Software

extrahiert. Die Bewertung der Reflexionstexte erfolgt durch die Zusammenführung der Ergebnisse beider Analyseverfahren. Zusätzlich wurde BERTopic auch zur Untersuchung des Reflexionsthemas verwendet. Die Ergebnisse der Qualitativen Inhaltsanalyse deuten darauf hin, dass die Teilnehmer:innen in erster Linie deskriptiv und auf niedrigem Niveau schreiben. Computergestützte linguistische Analysen ergaben weiter, dass die Verwendung affektiver und kognitiver Begriffe je nach Reflexionsebene unterschiedlich ausfiel, wobei ein häufigerer Einsatz derartiger Begriffe mit einer tieferen Reflexionsebene korrelierte. Die Ergebnisse der BERTopic-Analyse zeigen, dass sich die Reflexionsinhalte hauptsächlich auf das Lernmaterial konzentrierten und affektive Themen eher in Texten zu finden waren, die eine höheres Reflexionsebene aufwiesen.

Erstes Quasi-Experiment: Sommersemester 2023:
Das erste Quasi-Experiment wurde im Rahmen des E-Portfolio-Einsatzes im Sommersemester 2023 durchgeführt. Die Studierenden waren wie in den Semestern zuvor angehalten, innerhalb der beiden Themenbereiche sowohl eine fachliche Ausarbeitung zu einem Fallbeispiel, sowie ein Reflexionsschreiben anzufertigen. Zu den Reflexionsschreiben erhielten die Studierenden in diesem Semester unterschiedliches Feedback. Hierfür kam ein randomisiertes Design zur Anwendung. Die Studierenden wurden zunächst zwei Gruppen zugeordnet: Die erste Gruppe erhielt das Feedback zu den schriftlichen Reflexionen von dem nun einsatzfähigen KI-System PapagAI, während die zweite Gruppe das Feedback von einer/m Dozierenden bekam. Beide Gruppen setzten sich aus jeweils 30 Studierenden zusammen. Von diesen 60 Teilnehmenden bearbeiteten 45 alle Kursaufgaben, was einer effektiven Teilnahmequote von 75 Prozent entspricht.

Im Sommersemester 2023 wurden insgesamt drei Fragebögen an drei verschiedenen Messzeitpunkten eingesetzt (T0, T1, T2). In der ersten Befragung wurden verschiedene Skalen verwendet, um individuelle Einstellungen und Kompetenzen sowie die Wahrnehmung von Studienbedingungen zu erfassen. Zu diesen standardisierten Skalen gehörten Geschlecht, Alter, Fachsemester, Studium, Notendurchschnitt Studium, Notendurchschnitt im Abitur, Studienort, studienbezogene Selbstwirksamkeit, Berufswahlmotivation, Akzeptanz von KI, Notwendigkeit zur Selbstreflexion im Schulkontext, Notwendigkeit zur Selbstreflexion allgemein, Engagement in Selbstreflexion, Einsichten aus der Selbstreflexion. Insgesamt nahmen 121 Lehramtsstudierende an dieser Befragung teil, 79 Fragebögen wurden vollständig ausgefüllt. Der zweite Fragebogen wurde eingesetzt, nachdem die Studierenden ihr erstes Feedback erhalten hatten. Erhoben wurden Glaubwürdigkeit, Nützlichkeit, Unterstützungsgrad, ergebnisorientierte Dimension über das Feedback sowie offene Fragen zum Feedback. Es gab 101 Teilnehmer und 68 gültige Fragebögen. Zur abschließenden Befragung am Ende des Sommersemesters 2023 wurden verschiedene Skalen herangezogen, um die Nutzererfahrung sowie die Einstellungen der Studierenden zur Portfolio-Arbeit und zum Mini-E-Portfolio zu erfassen. Dabei wurden unterschiedliche Aspekte wie Motivation, Lernfortschritt, Umsetzung der Aufgaben, Zufriedenheit, Nutzung und Nützlichkeit der Prompts betrachtet. Zusätzlich bewerteten die Studierenden die im Mini-E-Portfolio eingesetzten Materialien, beispielsweise den Erwartungshorizont für die bearbeiteten Aufgaben. Es wurden auch Angaben zur Lehrveranstaltung erfasst sowie ein Wissenstest zu den beiden zentralen Themengebieten der Lehrveranstaltung durchgeführt. Darüber hinaus beinhaltete die Erhebung Skalen zur Reflexion (Notwendigkeit zur Selbstreflexion im Schulkontext; Notwendigkeit zur Selbstreflexion allgemein; Engagement in Selbstreflexion; Einsichten aus der Selbstreflexion) und zum Feedback (Glaubwürdigkeit; Nützlichkeit; Unterstützungsgrad; Ergebnisorientierte Dimension). An der Befragung nahmen

insgesamt 98 Studierende teil und 78 Fragebögen waren gültig. Im Rahmen der Erhebung wurden 72 schriftliche Reflexionen im Bereich Pädagogische Diagnostik und 67 schriftliche Reflexionen im Themengebiet Klassenmanagement angefertigt

Die Ergebnisse zeigen, dass beide Formen des Feedbacks einen minimalen Effekt auf die Tiefe der Reflexionen der Studierenden hatten. Die anschließende Analyse mittels *natural language processing* (NLP) zeigte keine signifikanten Unterschiede in den Metriken für Lesbarkeit, Wortschatzdichte, Satzstruktur, kognitive Metriken, zeitliche Metriken und affektive Metriken in den Aufsätzen der Studierenden. Die statistischen Analysen zeigten keine signifikanten Unterschiede in der Wahrnehmung der KI-generierten und der von den Dozierenden gegebenen Rückmeldungen durch die Studierenden.

Zweites Quasi-Experiment: Wintersemester 2023/2024

Das zweite Quasi-Experiment fand im Wintersemester 2023/2024 statt, wobei Aufbau und Kontext dem des ersten Quasi-Experiments im Sommersemester 2023 entsprachen. Der wesentliche Unterschied bestand in der Integration einer Gruppe, die Feedback von ChatGPT erhielt. Die Ergebnisse dieser Studie wurden sorgfältig ausgewertet und sollen im Rahmen der Arbeitsgruppe für empirische pädagogische Forschung (AEPF24) sowie bei der Gesellschaft für Empirische Bildungsforschung (GEBF25) präsentiert werden. In diesem Experiment verteilten sich die Teilnehmenden wie folgt: 25 befanden sich in der PapagAI-Feedback-Gruppe, 25 in der Dozent:innen-Feedback-Gruppe und 30 in der ChatGPT-Feedback-Gruppe.

Es fanden ebenfalls wieder drei Fragebogenerhebungen statt: Vor dem Start der Intervention, nach dem Feedback zur ersten Reflexion und nach dem Feedback zur zweiten Reflexion. Der erste Fragebogen sammelte grundlegende Informationen über die Studierenden, ihre medienbezogenen Lehrkompetenzen, die Akzeptanz von Künstlicher Intelligenz sowie ihre digitalen Kompetenzen. Aus dieser Erhebung wurden insgesamt 427 gültige Fragebögen gewonnen. Der zweite Fragebogen fokussierte auf die Einschätzung von Feedback, die Selbstregulation der Lernfähigkeit, Motivation, Prokrastination und studienbezogene Selbstwirksamkeit, woraus 326 gültige Fragebögen resultierten. Der dritte Fragebogen zielte erneut auf die Bewertung von Feedback, studienbezogene Selbstwirksamkeit und Überzeugungen bezüglich digitaler Kompetenzen ab. Letztlich konnten 317 gültige Fragebögen herangezogen werden. Zusätzlich konnten 341 schriftliche Reflexionen im Themenbereich Klassenmanagement und 345 Reflexionen im Themenbereich der Pädagogischen Diagnostik gesammelt.

Aus den vollständigen Fällen wurden randomisiert insgesamt 72 Lehramtsstudierende ausgewählt (76.34 % weiblich, M = 2.33 Semester). Diese verfassten zwei schriftliche Reflexionen und erhielten basierend auf denselben Kriterien entweder ein schriftliches Feedback durch Dozierende, ChatGPT oder die im Projekt entwickelten KI-App. Die Reflexionstexte wurden theoriebasiert und mit einer Qualitativen Inhaltsanalyse ausgewertet. Es zeigte sich, dass die Reflexionstexte der ChatGPT-Feedback-Gruppe am positivsten bewertet wurden, gefolgt von der Dozierenden-Gruppe und schließlich der KI-App-Gruppe ($F(2,69) = 6.22^{**}$, $\eta^2 = 0.15$). Allerdings ergaben sich keine signifikanten Unterschiede zwischen den Rückmeldeformaten in Bezug auf die Reflexionsfähigkeit der Studierenden ($\beta = -1.50$, $t(144) = -1.42$).

Drittes Quasi-Experiment: Sommersemester 2024

Das dritte Quasi-Experiment fand im Sommersemester 2024 in der gleichen Lernumgebung wie das erste und das zweite statt. Hauptunterschied war die Streichung einer Rückmeldung

durch die PapagAI-App. Da die PapagAI-App keine weiteren inhaltlichen Anpassungen mehr erhielt sowie Probleme in der Usability bestanden, wurden keine weitere Erhebungen mit diesem Tool mehr angedacht. Ziel dieses Quasi-Experiments war die unterschiedliche Wirkung zwischen KI-generierten Feedback (durch ChatGPT) und Dozierenden-Feedback vertieft zu untersuchen. Es nahmen 78 Personen teil (71,80 % weiblich, *M* = 2.79 Semester): D KI-Feedback-Gruppe bestand aus 36, die Dozierenden-Feedback-Gruppe aus 42 Teilnehmer:innen. Die Ergebnisse zeigten eine Verbesserung der Leistung beim Schreiben der zweiten Reflexion im Vergleich zur ersten Reflexion, wobei ein signifikanter Unterschied zwischen der KI-Feedback-Gruppe und der Dozierenden-Feedback-Gruppe zu Gunsten des Dozierenden-Feedbacks vorlag (β = -0.58, *KI*[-.099; -.17], p = <.01).. Bezüglich eines Einflusses von Lernstrategien ließen sich signifikante Beiträge der metakognitiven Strategien erkennen (β = .63, *KI*[-.19; 1.06], p <.01). Im Gegensatz dazu hatten kognitive und ressourcenorientierte Strategien keinen Einfluss.

UAP 4.3: Delphi-Studie mit Expert:innen verschiedener Disziplinen, Ableitung von Überarbeitungsempfehlungen für das KI-Portfolio

Im Rahmen des UAP 4.3 wurde ein Interviewleitfaden erstellt, um Expert:innen zu befragen. Insgesamt 9 Expert:innen aus der Erziehungswissenschaft (Schul- und Medienpädagogik, Hochschuldidaktik) und der Pädagogischen Psychologie wurden im Zeitraum von Dezember 2022 bis Februar 2023 zur Anwendung eines KI-basierten Portfolios, zu KI-Anwendungen allgemein im Bildungsbereich, als auch speziell in der Lehrer:innenbildung befragt. Weitere 3 Expert:innen aus dem Bereich der Informatik wurden im Sommer 2023 gebeten, sich hinsichtlich der Nutzerfreundlichkeit sowie der Anwendung von KI-Technologien im Bildungsbereich und der Lehrer:innenbildung zu äußern. Die Gesamtstichprobe beträgt daher 12 Expert:innen verschiedenster Fachrichtungen, welche Input zur Entwicklung und Implementierung eines KI-basierten Portfolios geben konnten. Die Ergebnisse der Delphi-Studie zeigen, dass die Expert:innen die Projektidee grundsätzlich befürworten und einen Vorteil darin sehen, das Thema KI in das Lehramtsstudium zu integrieren, um Vorbehalte dieser Personengruppe zu minimieren und den Kompetenzerwerb in diesem Bereich zu fördern. Folgende zentrale Ergebnisse wurden ermittelt:

Einschätzung der Akzeptanz gegenüber KI-basierten Lerntools:

Die Expert:innen setzten die Akzeptanz KI-basierter Lerntools in eine klare Abhängigkeit von der Technikaffinität von Lehramtsstudierenden. Sie vermuteten, dass bei diesen Studierenden eine Polarisierung zwischen Technikbegeisterten und Skeptikern besteht. Allgemein befürchten die Expert:innen, dass die Gruppe der Lehramtsstudierenden, die in der Literatur erfahrungsgemäß geringere Akzeptanz zeige, KI-basierten Lerntools mit Ängsten, offenen Fragen und Vorbehalten entgegen trete. Ebenso beeinflusse die Skepsis von Dozierenden ohne Medienbezug gegenüber KI-Technologien die Akzeptanz der Studierenden.

Die Förderung der Akzeptanz für KI-basierte Anwendungen beruhe den Expert:innen nach auf verschiedenen Schlüsselfaktoren: Eine detaillierte Aufklärung darüber, welche Daten und Informationen preisgegeben würden und welche konkreten Chancen und Risiken damit verbunden seien, sei letztlich entscheidend. Transparenz wurde wiederholt als wichtiger Punkt genannt, ebenso wie die Erklärung der technischen Hintergründe und der damit einhergehenden rechtlichen Konsequenzen, insbesondere im Sinne der Verantwortung der Institution (hier der Hochschule). Zudem müsse der Nutzen und Mehrwert der Anwendungen klar erkennbar sein.

Um die Akzeptanz und Nutzung zu begünstigen, spiele die Usability eine zentrale Rolle. KI-basierte Lerntools sollten aus Sicht der Expert:innen einfach, schnell verständlich, selbsterklärend und ansprechend gestaltet sein. Zudem werde die Akzeptanz durch den pädagogischen Umgang mit KI von Seiten der Dozierenden, sowie der Integration von KI-Bildung gefördert. Dies könne Ängsten entgegenwirken und eine kritische Reflexion fördern.

Ethische Bedenken bezüglich KI-basierten Lerntools

Die Expert:innen äußerten, dass Informatiker:innen oft eine große Begeisterung für technische Möglichkeiten hätten, der pädagogische Ansatz aber häufig in den Hintergrund rücke und ethische Aspekte kaum berücksichtigt würden. Ähnlich würde es sich in vielen Diskussionen über den Einsatz und Mehrwert von KI verhalten, wo zunächst die Bereicherung durch KI im Vordergrund stünde und ethische Fragen erst im Nachgang erwähnt werden würden. Die Herausforderungen lägen oft darin, dass sich bei Entitäten - wie OpenAI - Dynamiken entwickelten, die schwer nachvollziehbar seien und sich bestimmte Entscheidungen und Prozesse der Kontrolle entzögen. Ein zentraler Punkt sei dabei die Erkenntnis, dass KI im Allgemeinen nicht frei von Diskriminierung sei und stark von den bereitgestellten Informationen und ihrer Verarbeitung abhänge. Eine weitere potentielle Gefahr sahen die Expert:innen in der Benachteiligung von Studierenden, wenn durch die KI-basierten Lerntools unterschiedliche Aufgaben zugewiesen würden und hierdurch ein Gefühl von Ungerechtigkeit erzeugt werden könne. Dies könne zu ethischen Problemen innerhalb des Lernumfelds führen. Weitere ethische Bedenken wurden vor allem bei der Erfassung und Verwendung von Daten, insbesondere im Hinblick auf Datenschutz geäußert. Zudem wurde als übergeordnete ethische Schwierigkeit die Frage nach dem Subjektstatus von KI thematisiert. Die Diskussion darüber, wie KI als entitätsähnliche Systeme behandelt werden sollte, werfe grundlegende ethische Fragen auf, die bisher nicht beantwortet werden könnten.

Die Überarbeitungsempfehlungen wurden innerhalb den Projektteams berücksichtigt.

AP5: Projektkoordination / Kommunikation

UAP 5.2: Einbindung der assoziierten Partner und weiteren Akteure in der Lehrer:innenbildung

Darüber hinaus wurden Kolleg:innen aus der Fakultät, die stark in die universitäre Lehrer:innenbildung eingebunden sind, im Rahmen eines Forschungstages im Juli 2023 informiert. Auf diese Weise konnten Lehrende in den Fachbereichen Chemie-, Deutsch-, Geographie- und Musikdidaktik sowie Grundschulpädagogik an der FAU Erlangen-Nürnberg erreicht werden.

UAP 5.3: Auftakt-, Zwischen- und Abschlussworkshops

Am 11. und 12. Oktober 2023 fand in Kooperation mit dem Interdisziplinären Zentrum für empirische Lehr-Lernforschung (IZeLL) an der FAU Erlangen-Nürnberg eine Autumn School zum Thema „KI & Co - Bedeutung für die Zukunft der Bildung“ mit Mitarbeitenden des PetraKIP-Teams (aus Nürnberg und Berlin) statt. Diese Veranstaltung diente als eine Plattform für den Austausch und die Diskussion über die Rolle Künstlicher Intelligenz und verwandter Technologien in der Bildungslandschaft. Expert:innen, Forschende und Interessierte kamen zusammen, um sich zu informieren, innovative Ansätze zu erkunden und die Implikationen dieser

Technologien für die Zukunft der Bildung zu erörtern. Die Autumn School bot zudem die Gelegenheit PetraKIP-Projektergebnisse zu präsentieren und um Forschende wie Lehrende zahlreicher Hochschulen in Deutschland, die an der Autumn School teilnahmen, zu erreichen.

UAP 5.4: Dokumentation, Aufbereitung, Vorbereitung von Publikationen und Transfer
Die erhobenen Daten werden nach Ende des Projekts veröffentlicht.
Alle Publikationen werden im Kapitel "5. Erfolgte und geplante Veröffentlichungen aufgelistet.

2. Positionen des zahlenmäßigen Nachweises

Die Positionen des zahlenmäßigen Nachweises sind dem tabellarischen Verwendungsnachweis zu entnehmen.

3. Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Alle durchgeführten Arbeiten des Teilprojekts der FAU, sowie die dafür aufgewandten Ressourcen waren notwendig und angemessen, da sie der im Projektantrag formulierten Planung weitestgehend entsprachen und alle wesentlichen im Arbeitsplan formulierten Aufgaben erfolgreich bearbeitet wurden.

4. Verwertbarkeit der Ergebnisse

Die Hauptergebnisse sind in erster Linie für die Weiterentwicklung von lernförderlichen digitalen Konzepten sowie KI-basierten Anwendungen in der Lehrkräftebildung von großer Bedeutung. Vor allem konnten erstmalig empirische Erkenntnisse generiert werden, ob Studierende KI akzeptieren, wie sie KI-gestützte Rückmeldungen erleben, wie diese ihre Kompetenzentwicklung und ihren Wissenserwerb positiv beeinflussen können und wie mit verschiedenen Forschungsmethoden Reflexionstexte analysiert werden können. Diese Erkenntnisse sind bereits in die Lehre an der FAU integriert worden und darüber hinaus wurden sie bereits in mehreren internationalen Journals mit entsprechend großer Sichtbarkeit publiziert und im Rahmen nationaler und internationaler wiss. Kongresse präsentiert. Damit fließen sie in die aktuelle Diskussion um den Einsatz von KI-basierten-Verfahren im Lehr-Lernkontext (an Hochschulen und speziell in der Lehrkräftebildung) ein. Darüber hinaus sind die Forschungsergebnisse auch für den schulischen Bereich relevant. In aktuellen Folgeprojekten der FAU (zur Fortbildung von Lehrkräften und zur Gestaltung von Lehrveranstaltungen unter Nutzung von KI-Anwendungen) wird auf diese Ergebnisse rekurriert.

5. Erfolgte und geplante Veröffentlichungen

Publikationen wurden in Zeitschriften mit Peer-Review-Verfahren eingereicht bzw. publiziert (FAU-Teilprojekt):

- Gläser-Zikuda, M., Zhang, C., Hofmann, F., Plößl, L., Pösse, L., & Artmann, M. (2024). Mixed Methods Research on Reflective Writing in Teacher Education. *Frontiers in Psychology*, *15*. https://doi.org/DOI: 10.3389/fpsyg.2024.1394641
- Zhang, C., Hofmann, F., Kamran, F., & Gläser-Zikuda, M. (2024). Self-Regulated Learning of Pre-Service Teachers in Digital Learning Environments. In S. Mistretta (Eds.), *Self-Regulated Learning – Insights and Innovations.* (pp. 1-19). London: Intech-Open.
- Zhang, C., Hofmann, F., Plößl, L., & Gläser-Zikuda, M. (2024). Classification of reflective writing: A comparative analysis with shallow machine learning and pre-trained language models. *Education and Information Technologies*, 1-27. https://doi.org/10.1007/s10639-024-12720-0.

- Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Evaluating Reflective Writing in Pre-Service Teachers: The Potential of a Mixed-Methods Approach. *Education Sciences*, 13(12), 1213. https://doi.org/10.3390/educsci13121213.
- Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: a multigroup analysis. *International Journal of Educational Technology in Higher Education*, 20(1), 49. https://doi.org/10.1186/s41239-023-00420-7.

Dissertation (FAU-Teilprojekt):

- Publikationsbasierte Promotion von Chengming Zhang im Fachbereich Pädagogik. Titel „*Fostering Pre-Service Teacher Reflection through AI-Based Feedback: From Understanding AI Acceptance to Developing Effective AI-Driven Feedback*“ (Abschluss 06/2024)

Konferenzbeiträge wurden eingereicht, angenommen und durch Projektmitarbeitende der FAU präsentiert:

- Hofmann, F. & Zhang, Ch. (2024, November). Symposium submitted for EARLI Conference (European Association of Research on Learning and Instruction) 2025 in Graz (Austria): AI in Teacher Training – Concepts, Potentials, and Challenges of Prompts and Feedback
- Hofmann, F., Zhang, C., Solopova, V., Plößl, L., & Gläser-Zikuda, M. (2024, September). Förderung der Reflexionsfähigkeit von Lehramtsstudierenden: Zum Potential von KI-basiertem Feedback. In: Proceedings of Sektionstagung der Arbeitsgruppe empirische pädagogische Forschung (AEPF) in der Deutschen Gesellschaft für Erziehungwissenschaft (DGfE). Osnabrück, Germany.
- Plößl, L.; Gläser-Zikuda, M. (2024, September). KI-basierte Portfolios aus Sicht von Expert:innen. In: Proceedings of the Tagung Digitale Transformation für Schule und Lehrkräftebildung gestalten. Potsdam, Germany.
- Ortwig,J., Zhang, C., Gläser-Zikuda, M. (2024, September). Reflexionsfähigkeit von Lehramtsstudierenden – eine Mixed-Methods-Studie. In: Proceedings of Sektionstagung der Arbeitsgruppe empirische pädagogische Forschung (AEPF) in der Deutschen Gesellschaft für Erziehungwissenschaft (DGfE). Osnabrück, Germany.
- Hofmann, F., Zhang, C., Artmann, M., & Gläser-Zikuda, M. (2024, March). Reflexives Schreiben in der Hochschulbildung - ein systematisches Review zu Konzepten, Forschungsthemen und -methoden. In: Proceedings of the Gesellschaft für Empirische Bildungsforschung 2024 (GEBF2024), Potsdam, Germany.
- Zhang, C., Solopova, V., Hofmann, F., & Gläser-Zikuda, M. (2024, March). Reflexives Schreiben von Lehramtsstudierenden: Eine Analyse mit Verfahren maschinellen Lernens und vortrainierten Sprachmodellen. In: Proceedings of the Gesellschaft für Empirische Bildungsforschung 2024 (GEBF2024), Potsdam, Germany.
- Kamran, F., Zhang, C, & Gläser-Zikuda, M. (2023, April). The Formative Assessment and Feedback Practices at Language and Science Classrooms at Public Higher Education Institutions in Pakistan. In: Proceedings of the Annual Conference of the European Teacher Education Network (ETEN) “Teacher Education – Connecting Glocal”, Nürnberg, Germany.

- Zhang, C, Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023, April). Portfolio-based reflective writings in teacher education - Results of a mixed methods study. In: Proceedings of the Annual Conference of the European Teacher Education Network (ETEN) “Teacher Education – Connecting Glocal”, Nürnberg, Germany.
- Zhang, C, & Solopova V. (2023, April). AI-powered automatic feedback on reflective writing. In: Proceedings of the Annual Conference of the European Teacher Education Network (ETEN) “Teacher Education – Connecting Glocal”, Nürnberg, Germany.
- Zhang, C, Romeike, R., Gläser-Zikuda, M., Benzmüller,C., Landgraf, T., Hofmann, F., Solopova V., Schießl, J., Witte, S, Plößl, L.(2023, April). PetraKIP: Personal transparency AI-based portfolio for teacher education. In: Proceedings of the Annual Conference of the European Teacher Education Network (ETEN) “Teacher Education – Connecting Glocal”, Nürnberg, Germany.
- Zhang, C, Schießl, J.,Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023, February). Portfolio-based reflective writings in teacher education - Results of a mixed methods study. In: Proceedings of the Gesellschaft für Empirische Bildungsforschung 2023 (GEBF2023), Essen, Germany.
- Zhang, C, Schießl, J.,Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2022, October, Online). Quality of pre-service teachers’ reflections – a mixed-method study. In: Proceedings of the Reflexionstagung 2022 - Reflexion in der Lehrkräftebildung: Empirisch – Phasenübergreifend – Interdisziplinär, Berlin, Germany.
- Zhang, C, Schießl, J.,Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2022, September). Reflexives Schreiben und Reflexionsqualität von Lehramtsstudierenden - eine Mixed-Methods-Studie. In: Proceedings of the Sektionstagung der Arbeitsgruppe empirische pädagogische Forschung (AEPF) in der Deutschen Gesellschaft füR Erziehungswissenschaft (DGfE), Stuttgart, Germany.
- Zhang, C, Schießl, J.,Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2022, September). Akzeptanz von künstlicher Intelligenz (KI) bei angehenden Lehrkräften – eine Frage des Geschlechts? In: Proceedings of the Sektionstagung der Arbeitsgruppe empirische pädagogische Forschung (AEPF) in der Deutschen Gesellschaft füR Erziehungswissenschaft (DGfE), Stuttgart, Germany.
- Zhang, C, & Gläser-Zikuda, M. (2022, June). Moderating effect of gender on the acceptance of artificial intelligence among pre-service teachers: a multi-group analysis. In: Proceedings of the EARLI SIG11 Conference 2022 Digital Transformation in Teaching and Teacher Education, Oldenburg, Germany.

Masterarbeiten im Masterstudiengang Erziehungswissenschaftlich-Empirische Bildungsforschung an der FAU, die aus dem Projekt hervorgegangen sind:

- Plößl, Lea (2024). KI-basierte Lerntools für Lehramtsstudierende. Ergebnisse einer Expert:innenbefragung (unveröffentlichte Masterarbeit, Erziehungswissenschaftliche-Empirische Bildungsforschung), Friedrich-Alexander-Universität Erlangen Nürnberg.
- Ortwig, Julia (2024). „Reflexionsfähigkeit von Lehramtsstudierenden - eine Mixed-Methods Studie“ (unveröffentlichte Masterarbeit, Erziehungswissenschaftliche-Empirische Bildungsforschung), Friedrich-Alexander-Universität Erlangen, Nürnberg.

- Doganay, Meltem (2024). „Untersuchung der Professionalität in der Lehrer:innenbildung durch die Reflexionsniveaus angehender Lehrkräfte“ (unveröffentlichte Masterarbeit, Erziehungswissenschaftliche-Empirische Bildungsforschung), Friedrich-Alexander-Universität Erlangen, Nürnberg.

Modul-Projektarbeiten im Masterstudiengang Erziehungswissenschaftlich-Empirische Bildungsforschung an der FAU, die aus dem Projekt hervorgegangen sind:

:

- Caglar, Tugba (2022). „Digitale Kompetenzen und selbstreguliertes Lernen von Lehramtsstudierenden im Kontext der Corona-Pandemie“ (unveröffentlichte Projektarbeit, Erziehungswissenschaftliche-Empirische Bildungsforschung), Friedrich-Alexander-Universität Erlangen, Nürnberg.
- Doganay, Meltem (2022). „Reliabilität und Validität von KI-Akzeptanz-Skalen“ (unveröffentlichte Projektarbeit, Erziehungswissenschaftliche-Empirische Bildungsforschung), Friedrich-Alexander-Universität Erlangen, Nürnberg.

III. Erfolgskontrollbericht

1. Wirtschaftliche und wissenschaftliche Ergebnisse des Vorhabens, erreichte Nebenergebnisse und die gesammelten wesentlichen Erfahrungen

Die wissenschaftlichen Ergebnisse tragen signifikant zur Weiterentwicklung von lernförderlichen dig. Konzepten mit KI-basierten Anwendungen in der Lehrkräftebildung und Hochschullehre bei. Sie verdeutlichen insbesondere das Potential KI-basierten Feedbacks für den Kompetenzerwerb von Studierenden. Die Erkenntnisse dieses interdisziplinären Projekts wurden kontinuierlich und international publiziert. Im Rahmen der Projektkooperation konnten sich mehrere Nachwuchswissenschaftler:innen (Master bis Post-Doc-Niveau) mit Blick auf dieses innovative und interdisziplinäre Forschungsthema erfolgreich (weiter) qualifizieren.

2. Fortschreibung des Verwertungsplans

Im Teilprojekt der FAU waren keine Schutzrechtanmeldungen oder ähnliches notwendig. Erworbene Lizenzen wurden im Rahmen der Projektlaufzeit genutzt und nach Projektende nicht verlängert. Wissenschaftliche Erfolgsaussichten sind aufgrund der Menge an erhobenen Daten zu erwarten. Ebenso besteht die wissenschaftliche Anschlussfähigkeit an bereits veröffentlichten Ergebnissen.

3. Angaben zu Arbeiten, die zu keiner Lösung geführt haben

Da die durch das Teilprojekt der FU Berlin entwickelte PapagAI-App nicht wie ursprünglich geplant entwickelt werden konnte, war es im Teilprojekt der FAU nicht möglich Lernarrangements zu konzipieren, um mit der App selbstreguliertes Lernen und Reflexionsfähigkeit zu fördern. KI-basierte Rückmeldungen waren mit der App nur eingeschränkt umsetzbar, sodass eine summative Evaluation der Wirksamkeit der PapagAI-App nicht realisiert werden konnte.

4. Angaben über die Einhaltung der Ausgaben- und der Zeitplanung

Insgesamt konnte die Projektlaufzeit nicht gänzlich eingehalten werden. Das Projekt wurde ab Februar 2024 kostenneutral um weitere sechs Monate verlängert. Aufgrund der nur zum Teil erreichten Entwicklung der PapagAI konnten keine Evaluation der App stattfinden. Die finanziellen Mittel im Teilprojekt der FAU wurden zielgerichtet und wie beantragt verwendet, aber nicht vollständig ausgeschöpft.