

KI Tutoren: Eine Untersuchung zur Akzeptanz von KI Chatbots in digitalen Lernumgebungen

Bachelorarbeit von

Luca Hüttner

An der KIT-Fakultät für Informatik
KASTEL – Institut für Informationssicherheit und Verlässlichkeit

- | | |
|-------------------------|-----------------------------|
| 1. Prüfer/Prüferin: | Prof. Dr.-Ing. Anne Koziolk |
| 2. Prüfer/Prüferin: | Prof. Dr. Ralf Reussner |
| 1. Betreuer/Betreuerin: | M.Ed. Kai Marquardt |
| 2. Betreuer/Betreuerin: | Dr.-Ing. Lucia Happe |

27. Mai 2024 – 27. September 2024

Karlsruher Institut für Technologie
Fakultät für Informatik
Postfach 6980
76128 Karlsruhe

Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Quellen und Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, 27.09.2024

.....
(Luca Hüttner)

Zusammenfassung

Die individuelle Förderung von Schülerinnen und Schülern stellt in modernen Bildungssystemen aufgrund des oft hohen Lehrenden-Schüler-Verhältnisses eine große Herausforderung dar. Besonders nach Unterrichtschluss fehlt oft die notwendige Unterstützung, was den Lernerfolg beeinträchtigen kann. KI-basierte Chatbots bieten eine vielversprechende Lösung, da sie rund um die Uhr personalisierte Lernunterstützung ermöglichen. In dieser Arbeit wird die Akzeptanz von Lehrenden und Lernenden gegenüber solchen KI Chatbots anhand des Technology Acceptance Model (TAM) untersucht. Dazu wurde auf Basis von OpenAIs GPT-4 ein KI-Tutor exemplarisch für einen Online-Kurs von RockStartIT entwickelt. Hierfür wurden spezifische Instruktionen definiert, sowie eine Wissensdatenbank und ein Chatinterface in Form einer Web-App erstellt. Der KI-Tutor wurde von Lehrkräften und Lernenden getestet, die anschließend einen TAM-basierten Fragebogen ausfüllten. Ergänzend wurden qualitative Fragen gestellt, um Stärken, Verbesserungspotenzial sowie mögliche Entwicklungsrichtungen zu identifizieren. Die TAM-Komponenten Perceived Usefulness, Perceived Ease of Use und Intention to Use wurden in beiden Gruppen mehrheitlich positiv bewertet, was das Potenzial von KI-Tutoren in der Bildung unterstreicht.

Inhaltsverzeichnis

Zusammenfassung	i
1 Einleitung	1
1.1 Motivation	1
1.2 Forschungsfrage	2
2 Grundlagen	3
2.1 Chatbots	3
2.2 Natural Language Processing (NLP)	4
2.3 Large Language Models (LLMs)	5
2.4 Open AI	5
2.5 Technology Acceptance Model (TAM)	6
3 Verwandte Arbeiten	9
3.1 Chatbots aus der Sicht von Lehrkräften	9
3.2 Chatbots für Lernende	11
4 Anforderungen und Konzeption	15
4.1 Anforderungen	15
4.1.1 Pädagogische Anforderungen (PA)	15
4.1.2 Technische Anforderungen (TA)	16
4.2 Konzeption	16
5 Implementierung	19
5.1 Web-App	19
5.1.1 Architektur	19
5.1.2 Frontend	20
5.1.3 Backend	22
5.1.4 Assistant API Schnittstelle	22
5.2 KI-Tutor (OpenAI Assistant)	23
5.2.1 Wissensdatenbank und Formatwahl	23
5.2.2 Instruktionen	25
5.3 Abläufe	26
5.3.1 Thread-Erstellung	27
5.3.2 Nutzerinteraktion	28
5.4 Ergebnisse	30

5.5	Code Review	31
5.5.1	Markdown Unterstützung in KI-Tutor Antworten	32
5.5.2	Quellen Angaben	32
5.5.3	Magic Strings	32
5.5.4	Umgebungsvariablen	32
5.5.5	Testen der Web-App	33
6	Evaluation	35
6.1	Design	35
6.2	Fragebogen	35
6.2.1	Demografische Daten	36
6.2.2	Quantitative Fragen	36
6.2.3	Qualitative Fragen	37
6.3	Ergebnisse	38
6.3.1	Perceived Usefulness	38
6.3.2	Perceived Ease of Use	40
6.3.3	Intention to Use	42
6.4	Diskussion	43
7	Limitationen und zukünftige Arbeiten	49
7.1	Limitationen	49
7.2	Zukünftige Arbeiten	50
8	Fazit	51
	Literatur	55

Abbildungsverzeichnis

2.1	Die drei Rollen von Chatbots in der Bildung nach Deng u. Yu [12]	4
3.1	Übersicht über die Verwandten Arbeiten	10
5.1	Der Aufbau der KI-Tutor Web-App	19
5.2	Beispiel für den Aufbau der Markdown-Dateien in der Wissensdatenbank .	24
5.3	Ablauf Thread-Erstellung	27
5.4	Ablauf Nachricht Antwort Fluss	29
5.5	Beispiel Interaktion mit dem KI-Tutor	31
5.6	Benutzeroberfläche während der KI-Tutor antwortet	31
6.1	Perceived Usefulness (PU) von Lehrenden und Lernenden	39
6.2	PU der Lehrenden	40
6.3	PU der Lernenden	40
6.4	Perceived Ease of Use (PEOU) von Lehrenden und Lernenden	40
6.5	PEOU der Lehrenden	42
6.6	PEOU der Lernenden	42
6.7	Intention to Use (IU) von Lehrenden und Lernenden	42
6.8	IU der Lehrenden	43
6.9	IU der Lernenden	43

Tabellenverzeichnis

6.1	Übersicht über die Probanden	38
6.2	Deskriptive Statistik der Perceived Usefulness Aussagen	41

1 Einleitung

1.1 Motivation

In der heutigen Bildungsstruktur, bei der oft ein Lehrer dreißig Schüler betreut, stellt die Gewährleistung einer individuellen Unterstützung eine große Herausforderung dar. Dieses Ungleichgewicht begrenzt die Möglichkeit für personalisiertes Lernen, welches essenziell für die effektive Entwicklung und das Verständnis der Lerninhalte jedes einzelnen Schülers ist. Verschiedene Studien haben gezeigt, dass ungenügende persönliche Unterstützung zu schlechterer Leistung führen kann, während angepasste Unterstützung zu positiven Lerneffekten führen kann [7]. Ein weiteres Problem besteht darin, dass die Betreuung oft nach Unterrichtsschluss endet. Das bedeutet, dass bei Hausaufgaben oder generell auch in digitalen Lernsettings eine Unterstützungslücke besteht. Um Bildung skalierbarer zu machen und jedem Schüler die Aufmerksamkeit zu bieten, die er benötigt, müssen daher innovative Lösungen gefunden werden, die das Lehrer-Schüler-Verhältnis effektiv ergänzen und unterstützen können. Ein möglicher Ansatz hierbei könnten KI-Tutoren in Form von KI-Chatbots sein. In diesem Bereich hat es in den letzten Jahren eine große Entwicklung gegeben, die auch der Bildung zugute kommen kann. Sollten sich diese KI-Tutoren als gute Lösung erweisen, stünde die Unterstützung durch diese auch rund um die Uhr zur Verfügung und würde nicht nach Schulschluss enden.

Der Einsatz von Chatbots in der Bildung erfährt zunehmende Beachtung und wird auch zunehmend wissenschaftlich untersucht. Die Ergebnisse werden in verschiedenen systematischen Übersichtsarbeiten, wie den Reviews von Chee et al. [7] und Deng und Yu [12] zusammengefasst. Diese Arbeiten verdeutlichen das große Potenzial, das Chatbots für die Bildung bieten können. Ein Beispiel hierfür liefert die systematische Übersicht von Deng und Yu [12], die zu dem Schluss kommt, dass das Lernen mit Chatbots effektiver ist als traditionelle Lehrmethoden, insbesondere in den Bereichen explizites Argumentieren, Lernerfolg, Wissensbewahrung und Lerninteresse [12, S. 19]. Diese Erkenntnisse weisen auf signifikante Verbesserungen in vielen für die Bildung relevanten Metriken hin.

Weiterhin betonen die Autoren in ihrer Studie, dass zukünftige Forschungen die Einstellungen der Nutzer gegenüber der Chatbot-Technologie sowie die Lernhaltungen der Studierenden eingehender untersuchen sollten. Dies würde es ermöglichen, ein tieferes Verständnis dafür zu entwickeln, wie diese Technologien am effektivsten in Lernumgebungen integriert werden können und welche Auswirkungen sie auf die Motivation und das Engagement der Lernenden haben.

1.2 Forschungsfrage

Existierende Studien zeigen bereits, dass Chatbots in der Bildung ein großes Potenzial bieten. In dieser Arbeit liegt der Fokus auf der praktischen Umsetzung und der Akzeptanz dieser Technologien im Schulalltag sowie in der digitalen Lehre. Die übergeordnete Frage lautet dabei: Wie können KI-Chatbot-Technologien den Lehr- und Lernprozess effektiv unterstützen?

Im Rahmen dieser Arbeit wird die Einstellung von Lehrkräften und Lernenden gegenüber Chatbots untersucht, um daraus abzuleiten, wie ein solcher gestaltet werden muss, damit er als sinnvolle Ergänzung zu traditionellen Lehrmethoden angesehen wird und praktische Anwendung findet. Der Fokus liegt dabei vor allem auf der Akzeptanz und Wahrnehmung der Lehrenden und Lernenden gegenüber dieser Technologie. Daraus ergibt sich die zentrale Forschungsfrage:

- RQ: Wie ist die Akzeptanz von Lehrenden und Lernenden gegenüber KI-Chatbot-Technologie in digitalen Lernumgebungen?

Diese Forschungsfrage soll mit dem in der Literatur weit verbreiteten und genutzten *Technology Acceptance Model (TAM)* [11] (siehe Abschnitt 2.5) untersucht werden. Dazu wird in dieser Arbeit ein Chatbot am Beispiel eines Onlinekurses implementiert und mithilfe des TAM die Einstellung von Lernenden und Lehrenden zu diesem untersucht.

2 Grundlagen

In diesem Kapitel werden zentrale Begriffe und Konzepte erläutert, die für das Verständnis der weiteren Arbeit von Bedeutung sind.

2.1 Chatbots

Ein *Chatbot*, auch als *Conversational Agent (CA)* bezeichnet, ist ein Programm, das mit dem Nutzer in natürlicher Sprache interagiert, mit dem Ziel, eine menschliche Konversation zu simulieren [26]. Dafür können Text- oder Sprachform oder auch beides zusammen verwendet werden [24]. Ursprünglich entwickelten Programmierer Chatbots zur Unterhaltung und verwendeten einfache Keyword-Matching-Techniken, um Benutzereingaben zu beantworten, wie etwa beim ersten Chatbot ELIZA [27] [26, S. 30].

Seitdem gab es jedoch bedeutende Entwicklungen im Bereich der Verarbeitung natürlicher Sprache (*Natural Language Processing, NLP*). Insbesondere die Fortschritte in der Künstlichen Intelligenz (KI) und die Einführung von *Large Language Models (LLMs)* haben die Qualität und das Potenzial von Chatbots erheblich gesteigert. Modelle wie GPT-3 und seine Nachfolger nutzen Milliarden von Parametern sowie umfangreiche Trainingsdaten, um menschenähnliche Konversationen zu ermöglichen und vielfältige Aufgaben zu bewältigen [16, S. 156]. Diese technologischen Fortschritte haben die Anwendungsfelder von Chatbots erheblich erweitert und heute finden Chatbots breite Anwendung in verschiedenen Bereichen darunter Gesundheit, Dienstleistungen, als persönliche Assistenten und auch in der Bildung wird ihre Anwendung erforscht [20, S. 200].

Nach Gupta et al. lassen sich Chatbots basierend auf der Komplexität der Benutzeroberfläche, den Algorithmen und den verwendeten Technologien in drei Kategorien klassifizieren [17, S. 255]:

Menü-/Schaltflächenbasierte Chatbots (Menu/Button-Based Chatbots): Dies sind die am häufigsten verwendeten und einfachsten Chatbots. Sie arbeiten mit Schaltflächen und Dropdown-Menüs und folgen den Prinzipien von Entscheidungsbäumen, bei denen der Benutzer durch Auswählen von Optionen zu einer endgültigen Antwort gelangt. Diese Chatbots sind jedoch oft langsamer und nicht immer zuverlässig in der Bereitstellung der gewünschten Antworten [17, S. 255].

Chatbots auf Basis von Schlüsselwörterkennung (Keyword Recognition-Based Chatbots): Diese Chatbots erkennen spezifische Schlüsselwörter, um die gewünschte Antwort zu liefern. Sie hören auf die Eingaben der Benutzer und reagieren entsprechend. Mithilfe

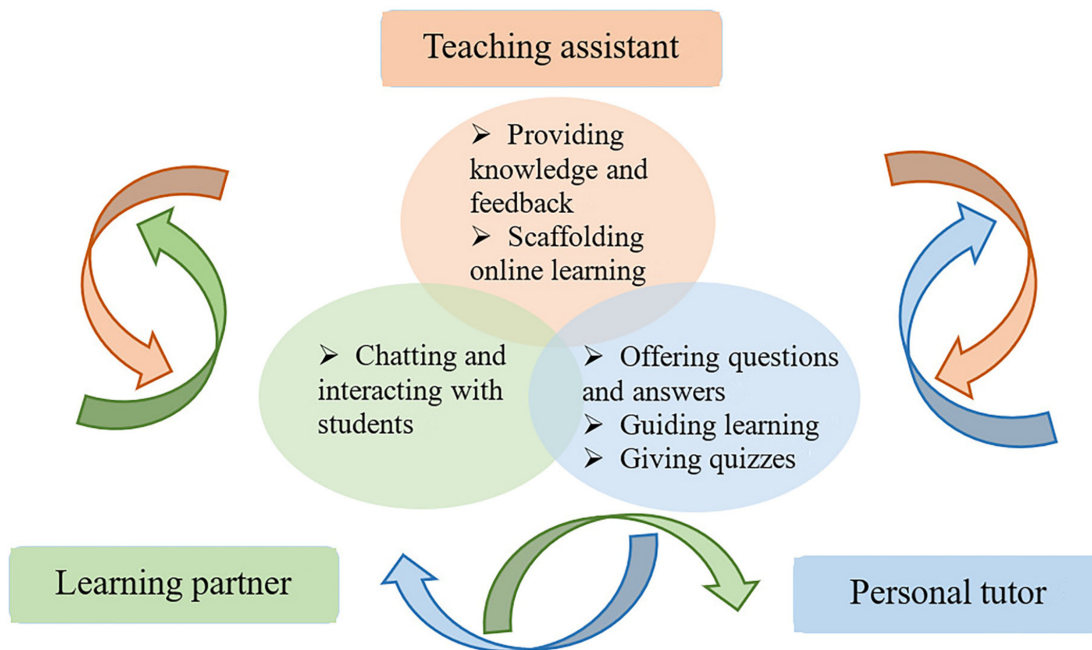


Abbildung 2.1: Die drei Rollen von Chatbots in der Bildung nach Deng u. Yu [12]

von KI-Technologie und angepassten Schlüsselwortlisten bestimmen sie die passende Antwort. Allerdings können sie bei Schlüsselwortüberschneidungen zwischen verschiedenen ähnlichen Fragen versagen [17, S. 255].

Kontextuelle Chatbots (Contextual Chatbots): Diese fortschrittlichen Chatbots nutzen maschinelles Lernen und KI-Technologien wie Spracherkennung und Sprach-zu-Text-Algorithmen, um die Absichten der Benutzer zu verstehen und entsprechend durch das Erkennen von Mustern in der Datenbank eine durchdachte Antwort zu geben [17, S. 255].

Nach Deng und Yu [12] übernehmen Chatbots in der Bildung drei Rollen (siehe Abbildung 2.1): Lehrassistenten, Lernpartner und persönliche Tutoren. Chatbots in der Rolle als Lehrassistenten bieten professionelles Wissen und kontinuierliches Feedback. Als Lernpartner unterhalten sich Chatbots mit den Schülern und interagieren entweder durch Text oder mit Stimmen. Die Tutoren-Rolle erfordert, dass Chatbots Fragen und Antworten anbieten, die Schüler beim Beginn ihres Lernens anleiten und Quizfragen erstellen. Diese drei Bildungsrollen sind miteinander verflochten, da sie sich gegenseitig ergänzen und gemeinsam zu einem effektiven Lehr- und Lernprozess beitragen [12, S.2].

2.2 Natural Language Processing (NLP)

Computerlinguistik, auch bekannt als *Natural Language Processing (NLP)*, ist ein Teilbereich der Informatik, der sich mit der Anwendung computergestützter Techniken befasst, um menschliche Sprache zu verstehen, zu verarbeiten und zu erzeugen [18, S. 261]. Systeme

der Computerlinguistik können verschiedene Zwecke erfüllen: Das Ziel kann die Unterstützung der menschlichen Kommunikation sein, wie etwa bei der maschinellen Übersetzung (*Machine Translation, MT*). Ebenso kann es die Mensch-Maschine-Kommunikation fördern, wie bei Chatbots. Zudem profitieren sowohl Menschen als auch Maschinen von der Analyse und dem Lernen aus der riesigen Menge an verfügbaren menschlichen Sprachinhalten im Internet [18, S. 261].

2.3 Large Language Models (LLMs)

Große Sprachmodelle (*Large Language Models, LLMs*) wie GPT-3 haben in den letzten Jahren erhebliche Fortschritte bei der Verarbeitung natürlicher Sprache (NLP) gemacht. Diese Modelle werden auf riesigen Mengen von Textdaten trainiert und sind in der Lage, menschenähnlichen Text zu erzeugen, Fragen zu beantworten und andere sprachbezogene Aufgaben mit hoher Genauigkeit zu erledigen [19, S. 1].

LLMs arbeiten, indem sie den Eingabetext (*Input-Prompt*) in einen hochdimensionalen Vektorraum kodieren, in dem semantische Beziehungen zwischen Wörtern und Sätzen erhalten bleiben. Das Modell dekodiert dann diese Darstellung, um eine Antwort zu generieren, die sich an den erlernten statistischen Mustern orientiert [8, S. 2].

Um möglichst qualitativ hochwertige Antworten zu generieren, wird oft das Verfahren des *Prompt-Engineerings* genutzt. Prompt-Engineering bezeichnet den systematischen Entwurf und die Optimierung von Input-Prompts, um die Antworten von LLMs zu beeinflussen und Genauigkeit, Relevanz und Kohärenz der generierten Ausgabe zu verbessern. Dieser Prozess ist von entscheidender Bedeutung, um das volle Potenzial dieser Modelle auszuschöpfen und sie in verschiedenen Bereichen zugänglicher und anwendbarer zu machen [8, S. 2].

Eine der in diesem Zusammenhang verwendeten Techniken ist das sogenannte *Role-Prompting*. Dabei wird dem KI-Modell eine bestimmte Rolle oder Identität zugewiesen, um die Qualität und Relevanz der generierten Antworten zu verbessern. Durch die klare Vorgabe einer Rolle, beispielsweise als Lehrer, Experte in einem Bereich oder Tutor, kann das Modell kontextsensitivere und zielgerichtetere Antworten liefern, da es seine Antworten aus der Perspektive der zugewiesenen Rolle generiert. Diese Technik hilft, das Verhalten des Modells besser zu steuern und es auf den spezifischen Anwendungsfall anzupassen [8, S. 5f].

2.4 Open AI

OpenAI ist ein Forschungslabor für künstliche Intelligenz, das mit dem Ziel gegründet wurde, eine ‚freundliche künstliche Intelligenz‘ zum Nutzen der Menschheit zu entwickeln. Im Jahr 2022 veröffentlichten sie ChatGPT, basierend auf ihrem selbst entwickelten LLM GPT-3, welches zu dieser Zeit das fortschrittlichste und leistungsfähigste Sprachmodell

war, das je produziert wurde [21, S. 389]. Auch heute ist OpenAI führend in der Entwicklung von Künstlicher Intelligenz und bietet Entwicklern zahlreiche Möglichkeiten, ihre Sprachmodelle für eigene Projekte anzupassen und zu nutzen. Dafür stellen sie verschiedene Programmierschnittstellen (APIs) bereit.

Einer dieser APIs ist die **Assistant API**. Die Assistant API von OpenAI ermöglicht es, angepasste Chatbots (Assistenten) auf Basis von OpenAIs GPT-Modellen zu erstellen, die spezifische Aufgaben erfüllen und auf bestimmte Bedürfnisse zugeschnitten sind. Ein zentrales Merkmal dieser API ist die Möglichkeit, dem Modell zusätzliche Instruktionen zu geben, die das Verhalten und die Reaktionsweise des Chatbots steuern. Dadurch können Entwickler sicherstellen, dass der Chatbot genau die Informationen liefert und die Ausdrucksweise verwendet, die für den jeweiligen Anwendungsfall erforderlich ist. Neben den angepassten Instruktionen bietet die Assistant API auch die Möglichkeit, den Chatbot mit zusätzlichen Datenquellen zu erweitern. Standardmäßig ist das Modell auf eine große Menge an öffentlich zugänglichen Daten trainiert, hat jedoch keinen Zugriff auf Informationen, die über den Trainingszeitpunkt hinausgehen oder auf spezielle anwendungsbezogene Daten. Um diese Lücke zu schließen, können Entwickler dem Chatbot den Zugriff auf externe Wissensquellen wie z.B. Dokumente ermöglichen. Diese müssen dafür in die Wissensdatenbank (Knowledgebase) hochgeladen werden.

Wenn man einen eigenen Chatbot mit der Assistant API erstellt hat, kann man mit ihm interagieren. Um einen neuen Chat zu starten, muss man zunächst einen **Thread** erstellen. Dieser Thread stellt eine eigenständige Konversation zwischen dem Chatbot und dem Nutzer dar. Alle vom Nutzer gesendeten Nachrichten sowie die Antworten des Chatbots werden in diesem Thread hinzugefügt und gespeichert, wodurch der Chatbot in der Lage ist, sich auf frühere Nachrichten zu beziehen.

Wenn der Nutzer eine Nachricht an den Chatbot sendet, wird diese dem Thread hinzugefügt, jedoch erfolgt die Antwort des Chatbots nicht sofort. Zunächst muss ein **Run** auf dem Thread ausgeführt werden. Ein Run ist ein Aufruf des Chatbots auf einen Thread. Dabei nutzt der Chatbot seine Konfiguration (einschließlich Instruktionen und Wissensdatenbank) sowie die vorhandenen Nachrichten im Thread, um seine in den Instruktionen festgelegte Aufgabe zu erfüllen. Während dieses Prozesses fügt der Chatbot eine Antwort auf die Nutzeranfrage dem Thread hinzu [3].

2.5 Technology Acceptance Model (TAM)

Das Technology Acceptance Model (TAM) von Davis [11] wurde entwickelt, um den Prozess der Akzeptanz von Informationstechnologie zu erklären. Es wurde in verschiedenen Kontexten weit verbreitet angewendet, um das Verhalten der Nutzer hinsichtlich der Akzeptanz neuer Technologien zu verstehen. Das TAM ist speziell auf Kontexte von Informationssystemen zugeschnitten und ist das bisher am meisten anerkannte Modell in Bezug auf die Forschung zum Benutzerakzeptanzverhalten. Es ist hoch generalisierbar und ermöglicht eine einfache Übertragung und Anwendung in verschiedenen Kontexten [10, S. 4]. Das

TAM untersucht hierbei die Komponenten: Perceived Usefulness und Perceived Ease of Use welche beide Einfluss auf die Nutzungsabsicht (Intention to Use, IU) haben.

Perceived Usefulness (PU) ist die wahrgenommene Nützlichkeit. Sie gibt an, in welchem Ausmaß eine Person das Gefühl hat, dass der Einsatz einer bestimmten Technologie ihre Fähigkeit, ihre Arbeit zu erledigen, verbessert. Es beschreibt die subjektive Einschätzung des Nutzers, dass die Nutzung des Systems für ihn von Vorteil ist und ist somit ein entscheidender Faktor für den Wunsch des Einzelnen, die Technologie zu verwenden [11].

Perceived Ease of Use (PEOU) ist die wahrgenommene Benutzerfreundlichkeit. Diese bezieht sich auf den Grad, zu dem eine Person glaubt, dass die Nutzung eines bestimmten Systems ohne Anstrengung möglich ist. Bei gleichen Bedingungen wird angenommen, dass eine Anwendung, die als benutzerfreundlicher wahrgenommen wird, eher von den Nutzern akzeptiert wird [11, S. 320].

3 Verwandte Arbeiten

In diesem Kapitel werden die verwandten Arbeiten untersucht. Abbildung 3.1 bietet eine Übersicht über die betrachteten Studien. Zunächst werden Arbeiten, die die Perspektive der Lehrenden beleuchten, näher betrachtet, gefolgt von Studien, welche die Sicht der Lernenden untersuchen.

3.1 Chatbots aus der Sicht von Lehrkräften

Chocarro et al. [10] und Belda-Medina und Calvo-Ferrer [4] untersuchten die Akzeptanz, Einstellung und Wahrnehmung von Lehrkräften sowie angehenden Lehrkräften gegenüber Chatbots im Bildungsbereich. Beide Studien verwendeten unter anderem das TAM-Modell, jedoch verfolgten sie unterschiedliche Ansätze.

In der Untersuchung von Chocarro et al. [10] wurde kein Chatbot implementiert, und die Lehrkräfte konnten auch nicht selbst mit einem Chatbot interagieren. Stattdessen wurden verschiedene Szenarien anhand von Screenshots von potenziellen Interaktionen dargestellt. Diese Szenarien zeigten mögliche Interaktionen von Lernenden mit einem Chatbot, etwa um organisatorische Aufgaben wie die Abfrage des Stundenplans für den Tag oder das Erfragen der Telefonnummer des Sekretariats durchzuführen. Die Studie von Chocarro et al. [10] zeigte, dass die Faktoren der wahrgenommenen Nützlichkeit (Perceived Usefulness, PU) und der wahrgenommenen Benutzerfreundlichkeit (Perceived Ease of Use, PEOU) im TAM-Modell entscheidende Einflussfaktoren für die Akzeptanz von Chatbots sind. Lehrkräfte, die diese beiden Aspekte positiv bewerteten, zeigten eine höhere Bereitschaft, Chatbots im Unterricht einzusetzen. Dennoch war in dieser Studie die Verhaltensabsicht, Chatbots im Unterricht zu nutzen, moderat. Obwohl Lehrkräfte die potenziellen Vorteile erkannten, waren sie zögerlich, diese Technologie sofort in ihren Unterricht zu integrieren. Weitere Erkenntnisse waren, dass weder das Alter der Lehrkräfte noch deren digitale Kompetenzen einen signifikanten Einfluss auf die Akzeptanz von Chatbots in der Studie hatten.

Belda-Medina und Calvo-Ferrer [4] untersuchten hingegen die Akzeptanz angehender Lehrkräfte gegenüber Chatbots im Sprachenlernen, wobei tatsächlich nutzbare Chatbots eingesetzt wurden. Diese Chatbots waren jedoch gewöhnliche „soziale Chatbots“, die als virtuelle Freunde oder Gesprächspartner mit den Nutzern interagieren sollten und nicht speziell für den Zweck des Sprachenlernens entwickelt worden waren. Die angehenden Lehrkräfte hatten vier Wochen Zeit, diese Chatbots zum Sprachenlernen auszuprobieren, und füllten anschließend einen Fragebogen basierend auf dem TAM-Modell aus. Die Ergebnisse des TAM zeigten positive Bewertungen hinsichtlich der wahrgenommenen Nützlichkeit,

Autor	Jahr	TAM	Lehrende	Lernende	Lehrende	Lernende	Chatbot	Nutzbar	Nutz GPT	Organisatorisch	Kursbezogen	Lower Education	Quantitative Studie	Methodik	Ziel der Untersuchungen
Chocarro	2023	Ja	Ja	Nein	Nein	Nein	Nein	Nein	Nein	Ja	Nein	Nein	Ja, 225 Lehrer	TAM	Akzeptanz bei Lehrer
Belda-Medina	2022	Ja	Ja	Nein	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Nein	Ja, 176 Lehrer	TAM	Akzeptanz bei Lehrer für Sprachunterricht
Bliquise	2024	Ja	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Ja	Nein	Nein	Ja, 207 Studierende	extended TAM	Akzeptanz bei Studenten
Al-Abdulatif	2023	Ja	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Ja, 432 Studierende	TAM, VAM	Akzeptanz bei Studenten
Malik	2021	Ja	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Ja, 372 Studierende	extended TAM	Akzeptanz bei Studenten
Atzu Deveci Topal	2021	Nein	Nein	Ja, Schüler	Ja	Nein	Ja	Nein	Nein	Nein	Ja	Ja	Ja, 41 Schüler	pre-postest	Leistung, Meinung
Essel	2022	Nein	Nein	Ja, Studenten	Ja	Nein	Nein	Nein	Nein	Nein	Ja	Nein	Ja, 68 Studierende	Test, Fokusgruppen	Akademisch Leistung, indirekt Wahrnehmung
Chen	2023	Nein	Nein	Ja, Studenten	Ja	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Ja, 195 Studierende	Quiz-Fragen	Meinung von Lernenden, Effektivität
Neo	2022	Ja	Nein	Ja, Studenten	Ja	Nein	Nein	Nein	Nein	Nein	Ja	Nein	Ja	TAM	Akzeptanz und Wahrgenommener nutzen
Tarek Ait Baha et al.	2023	Nein	Nein	Ja, Schüler	Ja	Nein	Nein	Nein	Nein	Nein	Ja	Ja	Ja	pre-postest	Leistung, Zufriedenheit durch Fragebogen

Abbildung 3.1: Übersicht über die Verwandten Arbeiten

Benutzerfreundlichkeit und Einstellung zur Integration von Chatbots, jedoch nur ein moderates Interesse an deren Nutzung in naher Zukunft. Zusätzlich wurden qualitative Daten in Form von offenen Fragen zu den Vor- und Nachteilen der Chatbots erhoben und analysiert. Besonders positiv wurden die lexikalische Vielfalt und die natürliche Interaktion der Chatbots bewertet. Es gab jedoch auch einige Schwächen in der Grammatik und Fehlerkorrektur. Die Chatbots erkannten Fehler häufig nicht oder korrigierten diese nicht, was insbesondere für Lernende mit geringerem Sprachniveau als Nachteil angesehen wurde.

Im Gegensatz zu den oben erwähnten Arbeiten bezieht sich diese Arbeit zwar ebenfalls auf eine Akzeptanzuntersuchung mithilfe des TAM-Modells, jedoch wird in dieser Untersuchung ein real nutzbarer Chatbot eingesetzt, der speziell für den Lehrkontext sowie für einen spezifischen Onlinekurs entwickelt wurde.

3.2 Chatbots für Lernende

Bei der Suche zu Chatbots in der Bildung konnten mehr Studien im Bezug auf die Akzeptanz und Auswirkungen auf Lernende gefunden werden als auf Lehrende. Verschiedene Studien untersuchten die Akzeptanz von Chatbots aus der Sicht von Lernenden [5, 1, 20, 22]. Unter anderem hat Bilquise et al. [5] bei einer Online Umfrage mit 207 Hochschulstudenten über einen fürs Studium beratenden Chatbot herausgefunden, dass vor allem die wahrgenommene Benutzerfreundlichkeit (PEOU) einen Einfluss auf die Nutzungsabsicht der Studierenden hatte. Außerdem gab es Hinweise darauf, dass Studierende Chatbots eher nutzen würden, wenn diese menschliche Züge wie soziale und emotionale Intelligenz zeigen [5, S. 6374]. In der Studie von Al-Abdullatif [1] mit 432 Studierenden in Saudi-Arabien hatte die PEOU hingegen keine signifikanten Auswirkungen auf die Einstellung der Studierenden gegenüber dem Einsatz von Chatbot-Technologie. In dieser Untersuchung war die wahrgenommene Nützlichkeit (PU) der größte Faktor, der die Akzeptanz beeinflusste. In der Studie von Malik et al. [20], bei der 372 Studierende weltweit befragt wurden, hatten beide Faktoren, PEOU und PU, einen signifikanten positiven Einfluss auf die Einstellung der Studierenden, Chatbots zum Lernen zu nutzen.

Eine Limitation dieser Studien ist, dass keiner der Chatbots tatsächlich nutzbar war. Sie basierten entweder auf Szenarien, die in der Umfrage mithilfe von Screenshots dargestellt wurden [5], oder auf den bisherigen persönlichen Erfahrungen der Studierenden [1, 20]. Nur bei der TAM-Untersuchung von Neo [22] wurde die Akzeptanz von Studierenden anhand eines für einen Online-Kurs nutzbaren Chatbots evaluiert.

Dafür wurden 102 Studierende in Malaysia mithilfe eines TAM-basierten Fragebogens befragt. Die Ergebnisse waren überwiegend positiv und weisen auf das Potenzial von Chatbots in digitalen Lernumgebungen hin. Die Faktoren PEOU und PU wurden beide sehr positiv bewertet und auch die Nutzungsabsicht (Intention to Use, IU) für Chatbots wie diesen fiel positiv aus. Das zusätzlich abgefragte Feedback und die Kommentare der Studierenden bestätigten diese Ergebnisse und zeigten, dass der Chatbot insbesondere in vier Bereichen im Lernprozess effektiv war: 1) Unterstützung des Lernprozesses durch

Feedback beim asynchronen Lernen, 2) Förderung einer aktiveren Beteiligung und eines tieferen Verständnisses, 3) Steigerung des Lerninteresses durch sofortiges Feedback und 4) Erhöhung der Teilnahme sowie Schaffung einer immersiven Lernerfahrung [22, S. 45].

Chen et al. [9] untersuchten die Meinung von Studierenden zu Chatbots, um herauszufinden, wie solche sie am besten beim Lernen unterstützen können. Dafür wurden 215 Studierende befragt. Genannt wurden unter anderem die Beantwortung von Fragen zum Kursmaterial und das Bereitstellen weiterführender Beispiele. Als besondere Stärken wurden die schnelle Reaktionsfähigkeit, die Möglichkeit, rund um die Uhr Antworten zu erhalten, sowie die erhöhte Interaktivität im Vergleich zu Lernvideos hervorgehoben [9, S. 168]. In der Studie wurde auch die Effektivität der Vermittlung von Lerninhalten durch Chatbots untersucht. Dazu wurden Grundlagen zu KI mithilfe eines exemplarischen Chatbots vermittelt. Nach der Vermittlung wurden die Inhalte ebenfalls vom Chatbot abgefragt, wobei die Studierenden gute Ergebnisse erzielten, was die Autoren als vielversprechend bewerteten. Allerdings war die Untersuchung limitiert, da es keine Kontrollgruppe oder Vergleichsmöglichkeiten gab.

Die Leistung von Lernenden nach der Inhaltsvermittlung durch Chatbots wurde auch in weiteren Studien untersucht [14, 2, 13]. Ait Baha et al. [2] und Deveci Topal et al. [13] konnten keine signifikanten Änderungen oder Verbesserungen feststellen. Ait Baha et al. [2] untersuchten die Auswirkungen von Bildungs-Chatbots bei Sekundarschülern in Marokko. Zu diesem Zweck wurde ein kursbegleitender Chatbot entwickelt, der als virtueller Tutor die Programmiersprache Logo vermitteln sollte. In der Studie wurde die Leistung einer Experimentalgruppe mit Chatbot-Zugriff mit einer Kontrollgruppe verglichen, die keinen Chatbot zur Verfügung hatte und ausschließlich traditionelle Lernmethoden nutzte. Es wurden ein Pretest und ein Posttest durchgeführt. Zusätzlich wurden in der Chatbot-Gruppe die Zufriedenheit sowie die Lernerfahrung mit dem Chatbot erfasst. In Bezug auf die akademische Leistung lagen beide Gruppen auf einem ähnlichen Niveau, jedoch deuten die Ergebnisse darauf hin, dass die Schüler der Chatbot-Gruppe eine deutlich bessere Lernerfahrung hatten, mit weniger Stress, mehr Motivation und der Möglichkeit, in ihrem eigenen Tempo zu lernen. Deveci Topal et al. [13] kamen zu einem ähnlichen Ergebnis. In dieser Studie wurde ein kursbegleitender Telegram-Chatbot für Schüler der 5. Klasse entwickelt. An der Studie nahmen 41 Schüler teil, die in eine Experimental- und eine Kontrollgruppe aufgeteilt wurden. Die Experimentalgruppe nutzte zusätzlich zum regulären Unterricht den Chatbot. Es wurde ein Pretest und ein Posttest durchgeführt, und die Erfahrungen der Experimentalgruppe wurden in einer Fokusgruppe erfasst. Das Ergebnis der Studie zeigte, dass es zwar keinen signifikanten Unterschied in der akademischen Leistung zwischen der Experimental- und der Kontrollgruppe gab, jedoch die Chatbot-Anwendung die Online-Lernerfahrung der Schüler in der Experimentalgruppe positiv beeinflusste.

Die Ergebnisse von Essel et al. [14] stützen ebenfalls die verbesserte Lernerfahrung. Die Autoren untersuchten den Einfluss eines kursangepassten Chatbots auf das Lernen von Studierenden in Ghana. Der Chatbot wurde auf der WhatsApp-Plattform mithilfe von FloXO implementiert und diente dazu, Studierende in einem Multimedia-Programmierkurs zu unterstützen. In der Studie wurden 68 Studierende in eine Experimentalgruppe mit Chatbot-Interaktion und eine Kontrollgruppe, die ausschließlich mit Dozenten interagierte,

aufgeteilt. Untersucht wurden primär die akademische Leistung nach dem Kurs sowie indirekt die Wahrnehmung des Chatbots. Die Ergebnisse zeigten, dass Studierende, die mit dem Chatbot interagierten, signifikant bessere akademische Leistungen erzielten als jene, die nur mit dem Dozenten kommunizierten. Besonders positiv hervorgehoben wurden die sofortige Verfügbarkeit des Chatbots rund um die Uhr und die Fähigkeit, schnelle Antworten auf Fragen zu geben. Gleichzeitig bemängelten die Studierenden jedoch, dass die Antworten des Chatbots teilweise nicht tiefgehend genug und in einigen Fällen veraltet waren. Essel et al. schlussfolgern, dass Chatbots besonders in Kontexten mit hohen Studierenden-Lehrenden-Verhältnissen hilfreich sein können, da sie die Interaktion verbessern und den Lehrenden entlasten. Gleichzeitig betonen sie, dass die Technologie weiterentwickelt werden muss, um tiefere, kontextbezogene Antworten zu ermöglichen und so eine nachhaltigere Unterstützung des Lernprozesses zu bieten.

Diese Studien zeigen das Potenzial, das Chatbots in der Bildung haben können. Insgesamt deuten die Ergebnisse auf eine verbesserte Lernerfahrung hin [22, 2, 13] und teilweise auch auf bessere akademische Leistungen [14]. Allerdings wurden in mehreren dieser Studien die Umfragen ohne tatsächlich nutzbare Chatbots durchgeführt [5, 1] und in den Fällen, in denen nutzbare Chatbots getestet wurden, handelte es sich meist um regelbasierte Systeme, die nicht von den neuesten Fortschritten im Bereich der großen Sprachmodelle (LLMs) profitierten [14, 22]. Diese Fortschritte könnten einige der in den Studien erwähnten Nachteile, wie nicht tiefgehend genug oder veraltete Antworten der Chatbots, beheben.

Diese Arbeit zielt darauf ab, diese Lücken zu adressieren, indem ein tatsächlich nutzbarer KI-Tutor auf Basis von OpenAI's GPT-4o, einem der derzeit führenden LLMs, implementiert wird. Auf dieser Grundlage wird eine Untersuchung mithilfe des Technologie-Akzeptanz-Modells durchgeführt. Eine weitere Abgrenzung dieser Arbeit ist, dass die Akzeptanz sowohl von Lehrenden als auch von Lernenden mit demselben Chatbot untersucht wird, wodurch direkte Vergleiche zwischen den beiden Gruppen möglich sind. Dies erlaubt es, spezifische Unterschiede in den Wahrnehmungen und Erwartungen beider Zielgruppen herauszuarbeiten, was wertvolle Einblicke in die unterschiedlichen Anforderungen und Optimierungspotenziale für den Einsatz von KI-Tutoren im Bildungsbereich bietet.

4 Anforderungen und Konzeption

4.1 Anforderungen

Da der KI-Tutor im Lernkontext angewendet werden soll und sich an Lernende der 7. bis 10. Klasse richtet, ergeben sich verschiedene pädagogische, sowie technische Anforderungen. In diesem Kapitel werden die Anforderungen an den KI-Tutor aufgelistet und erklärt. Ziel ist es, ein umfassendes Verständnis der notwendigen Rahmenbedingungen zu schaffen, die für die erfolgreiche Implementierung eines KI-Tutors erforderlich sind.

4.1.1 Pädagogische Anforderungen (PA)

Pädagogische Anforderungen sind entscheidend, da sie sicherstellen, dass der KI-Tutor die Lernenden effektiv unterstützt. Ein guter KI-Tutor sollte in der Lage sein, sich an die individuellen Bedürfnisse der Schülerinnen und Schüler anzupassen und ihnen auf verständliche und motivierende Weise Lerninhalte zu vermitteln. Gerade für Lernende der 7. bis 10. Klasse ist es wichtig, dass der Tutor altersgerechte Inhalte und Methoden nutzt, um den Lernprozess zu fördern. In diesem Abschnitt werden die grundlegenden pädagogischen Anforderungen beschrieben, die der KI-Tutor für diese Arbeit erfüllen sollte.

PA1: Altersgerechte Sprache

PA1.1: Verwendung einer einfachen und klaren Sprache, die für die Zielgruppe leicht verständlich ist.

PA2: Unterstützung bei Lerninhalten

PA2.1: Bereitstellung von Erklärungen und Beispielen zu den im Kurs behandelten Themen.

PA2.2: Vermeidung der direkten Ausgabe von Lösungen; stattdessen soll der Nutzer durch gezielte Hinweise zur Lösung hingeführt werden.

PA3: Personalisierung

PA3.1: Anpassung der Antworten und Hilfestellungen an den individuellen Lernfortschritt und die Bedürfnisse der Schüler (Alter/Niveau).

PA3.2: Vertiefung in Themen, an denen die Lernenden besonderes Interesse zeigt.

PA4: Motivationsförderung

PA4.1: Nutzung von positiven Verstärkungen, wie Lob und Anreize, um die Schüler zu motivieren.

4.1.2 Technische Anforderungen (TA)

Die technischen Anforderungen dienen dazu, sicherzustellen, dass die Lernenden einen zuverlässigen KI-Tutor zur Verfügung haben, der eine stabile und unterstützende Lernumgebung bietet und ihre Bildungsziele optimal unterstützt. Sie setzen sich aus folgenden Punkten zusammen:

TA1: Benutzerfreundlichkeit

TA1.1: Einfache und intuitive Bedienung, sodass Lernende ohne großen Aufwand mit dem Chatbot interagieren können.

TA1.2: Kompatibilität mit verschiedenen Geräten (Smartphones, Tablets, Laptops, etc.).

TA2: Fehlertoleranz

TA2.1: Fähigkeit, Rechtschreib- und Grammatikfehler in den Eingaben der Schüler zu erkennen und trotzdem korrekt zu reagieren.

TA3: Zugänglichkeit

TA3.1: Rund-um-die-Uhr-Verfügbarkeit, um den Schülern jederzeit Unterstützung bieten zu können.

TA3.2: Nutzbarkeit des KI-Tutors ohne die Notwendigkeit externer Accounts, um den Zugang für alle Kursteilnehmer zu ermöglichen.

TA4: Konsistenz der Lerninhalte

TA4.1: Der Chatbot sollte stets auf die Kursinhalte zurückgreifen, um sicherzustellen, dass die Unterstützung immer relevant und im Kontext des aktuellen Lernstoffs bleibt.

4.2 Konzeption

In diesem Abschnitt wird die Konzeption des KI-Chatbots behandelt, wobei zunächst die Technologie-Wahl im Fokus steht, um die in Abschnitt 4.1 definierten Anforderungen bestmöglich zu erfüllen.

Der KI-Chatbot wurde in Form eines KI-Tutors für den ersten Teil des Kurses „Rettet die Bienen!“ von der Online-Kurs Webseite „RockStartIT“ umgesetzt.

Um der Anforderung *TA1.2: Kompatibilität mit verschiedenen Geräten* so gut wie möglich zu entsprechen, wurde sich für eine Umsetzung des KI-Tutors als Web-App entschieden. Diese Lösung bietet den Vorteil, dass sie plattformunabhängig ist und über jeden Webbrowser genutzt werden kann, unabhängig vom verwendeten Gerät oder Betriebssystem. Dadurch wird gewährleistet, dass Lernende auf Smartphones, Tablets und Laptops gleichermaßen Zugriff auf den KI-Tutor haben. Die Web-App ermöglicht es zudem, zentrale Updates und Verbesserungen problemlos durchzuführen, sodass alle Nutzer stets die neueste Version verwenden, was die Wartung und Verwaltung vereinfacht. So wird die Anforderung der Gerätekompatibilität vollständig erfüllt und den Schülern eine flexible und zugängliche Lernumgebung geboten. Die Umsetzung als Web-App bietet auch das Potenzial, den KI-Tutor später direkt in den Online-Kurs integrieren zu können.

Bei der Wahl für das zugrundeliegende Large Language Model (LLM) für den KI-Tutor wurde sich für OpenAI's GPT-4o Model entschieden. Nach einer Analyse von Fulgencio [16] zeichnen sich OpenAI's GPT-Modelle besonders durch ihre benutzerfreundliche Handhabung, die Fähigkeit, kohärente und kontextuell relevante Antworten zu erzeugen, sowie ihre vielseitigen Anpassungsmöglichkeiten aus. Diese Eigenschaften machen es zu einem wertvollen Werkzeug für die Erstellung eines Bildungs-Chatbots, der speziell auf einen Kurs zugeschnitten ist. Die Fulgencio [16] kommt zu dem Schluss, dass OpenAI's GPT-Modelle derzeit die optimale Wahl sind. Sie bieten eine einfache Implementierung für Lehrkräfte mit unterschiedlichen technologischen Kenntnissen und fesseln die Lernenden durch adaptive, sprachbasierte Interaktionen, die auf verschiedene Lernstile und Bedürfnisse eingehen [16, S. 159].

OpenAI bietet derzeit drei Möglichkeiten, um ChatGPT an eigene Anwendungszwecke und Daten anzupassen: Custom GPTs, die Assistant API und das Finetuning der OpenAI-Modelle (Nachtrainieren mit eigenen Daten). Finetuning ist dabei die teuerste und aufwändigste Methode, und OpenAI empfiehlt in ihrer Dokumentation, diese Funktionalität erst zu verwenden, wenn Custom GPTs oder die Assistant API nicht zu zufriedenstellenden Ergebnissen führen [15].

Custom GPTs bieten die Möglichkeit, die Basis-GPT-Modelle von OpenAI mit zusätzlichen Datensätzen und spezifischen Instruktionen anzupassen. Dies ermöglicht die Erstellung maßgeschneiderter Chatbots, die genau auf bestimmte Anwendungsbereiche und Bedürfnisse zugeschnitten sind. Die Erstellung von Custom GPTs ist vollständig ohne Programmierkenntnisse möglich, da sie in OpenAI's ChatGPT integriert sind und kein Bereitstellen einer separaten Webseite erfordern. Custom GPTs können zudem öffentlich geteilt werden, sodass auch andere Nutzer auf sie zugreifen können. Ein entscheidender Nachteil von Custom GPTs ist jedoch, dass eine Integration in eine eigene Webseite nicht möglich ist und jeder Kursteilnehmer zum Stand dieser Arbeit einen eigenen OpenAI-Account zur Nutzung benötigt, was direkt *TA3.2: widerspricht*. Aus diesem Grund wurde die Entscheidung für die dritte Option getroffen: die OpenAI Assistant API.

Die OpenAI Assistant API stellt eine flexible Alternative zu den Custom GPTs dar und bietet fast identische Anpassungsmöglichkeiten. Mit der Assistant API können ebenfalls zusätzliche Datensätze bereitgestellt und spezifische Instruktionen definiert werden, um den Chatbot auf bestimmte Anwendungsbereiche zuzuschneiden. Ein wesentlicher Vorteil der

API ist die Möglichkeit, den Chatbot nahtlos in eigene Anwendungen integrieren zu können, was potentiell eine tiefere Einbettung in den Online-Kurs ermöglicht, beispielsweise durch ein Popup direkt im Kurs. Diese Flexibilität bringt jedoch auch einen höheren technischen Aufwand mit sich, da sowohl ein Backend als auch ein Chatbot-UI-Frontend erstellt und gehostet werden müssen.

5 Implementierung

In diesem Kapitel wird die Implementierung des KI-Tutors detailliert erläutert. Zunächst wird die Implementierung der Web-App behandelt, gefolgt von der Umsetzung des KI-Tutors auf Basis der Assistant API. Anschließend wird das Zusammenspiel der Web-App mit der Assistant API anhand der zwei wichtigsten Abläufe des KI-Tutors beschrieben. Abschließend werden die Ergebnisse der Implementierung vorgestellt.

5.1 Web-App

5.1.1 Architektur

Die Web-App besteht wie in Abbildung 5.1 zu sehen ist aus zwei Hauptkomponenten: dem Frontend und dem Backend. Das Frontend stellt die Benutzeroberfläche bereit und wird direkt im Browser des Nutzers ausgeführt. Es dient dazu, eine einfache und intuitive Interaktion mit dem KI-Tutor zu ermöglichen. Das Backend hingegen spielt eine wichtige Rolle im Schutz sensibler Informationen, insbesondere des OpenAI API-Schlüssels. Dieser Schlüssel muss verborgen bleiben, um zu verhindern, dass er von potenziellen Angreifern ausgelesen und für missbräuchliche Zwecke verwendet wird, was hohe Kosten verursachen könnte. Da das Backend hauptsächlich als Schnittstelle zur Weiterleitung von Anfragen

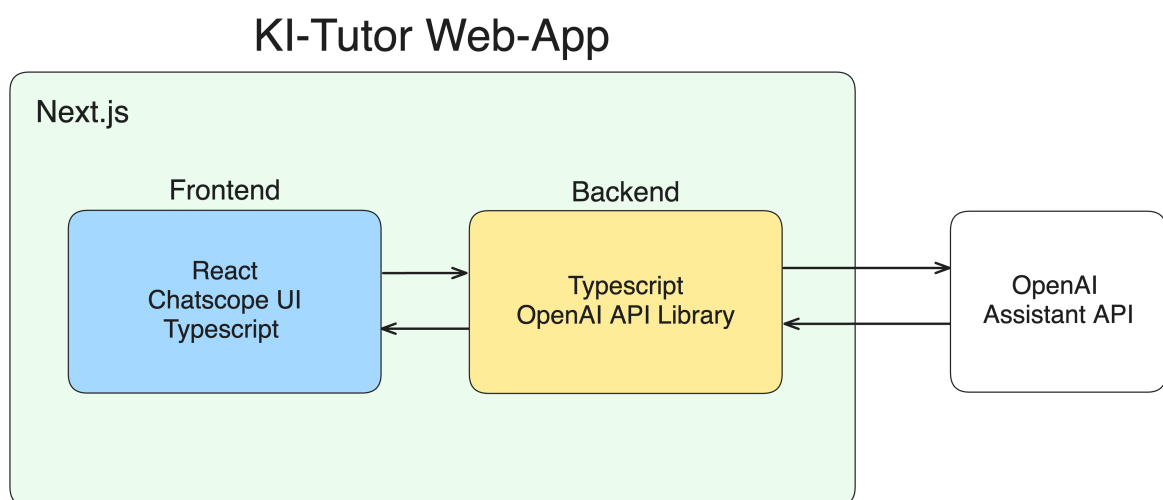


Abbildung 5.1: Der Aufbau der KI-Tutor Web-App

zwischen dem Frontend und der Assistant API fungiert, wurde es bewusst minimal gehalten, um Ressourcen zu schonen und die Wartung zu vereinfachen.

Zur Umsetzung der Web-App wurde das Next.js Fullstack Framework gewählt. Dies bietet den Vorteil, sowohl das Frontend als auch das Backend in einer einheitlichen Umgebung zu entwickeln, was die Integration und Wartung vereinfacht. Die zugrundeliegende Programmiersprache, für sowohl Frontend als auch Backend, ist hierbei Typescript.

Für die Implementierung der Benutzeroberfläche im Frontend wurde außerdem das weitverbreitete Framework React verwendet, welches bereits in Next.js integriert ist. Dies ermöglicht es, für die Chatoberfläche, mit der der Nutzer später interagiert, vorgefertigte Komponenten aus der Open-Source-Bibliothek Chat UI Kit React von Chatscope [6] zu nutzen. Dies beschleunigt und vereinfacht die Entwicklung der Chatoberfläche und sorgt gleichzeitig für eine bessere Wartbarkeit.

Im Backend wurde zusätzlich die von OpenAI bereitgestellte Bibliothek OpenAI Node API Library [23] verwendet, die die Interaktion mit der OpenAI API vereinfacht, da sie Typen und Hilfsfunktionen für den Zugriff auf die API bereitstellt.

5.1.2 Frontend

Das Frontend der Web-App verwendet React, ein weit verbreitetes Frontend-Framework. Es vereinfacht die Trennung der Benutzeroberfläche (UI) in einzelne Komponenten, die als wiederverwendbare UI-Bausteine fungieren. Jede Komponente repräsentiert dabei einen Teil der Benutzeroberfläche samt zugehöriger Logik und kann in andere Komponenten eingebettet werden, wodurch die Modularität und Wiederverwendbarkeit erhöht wird.

5.1.2.1 Benutzeroberfläche (UI)

Für die Implementierung der Chatoberfläche des KI-Tutors wurde die Library Chat UI Kit React von Chatscope [6] verwendet, die speziell entwickelte Komponenten für die Erstellung von Chat-Benutzeroberflächen bereitstellt. Dies vereinfacht die Implementierung erheblich und erhöht die Wartbarkeit der Anwendung. Die Gestaltung der Benutzeroberfläche des KI-Tutors besteht hauptsächlich aus der Kombination dieser vorgefertigten Komponenten und deren Verknüpfung mit der Logik des KI-Tutors. Die Zusammenstellung der UI-Komponenten für den KI-Tutor ist in Listing 5.1 zu sehen.

```
1 <div style={{ position:"relative", height: "100dvh" }}>
2   <MainContainer>
3     <ChatContainer>
4       <ConversationHeader>
5         <ConversationHeader.Content userName="KI-Tutor" />
6       </ConversationHeader>
7       <MessageList typingIndicator={
8         inputDisabled ? <TypingIndicator content="KI-Tutor schreibt..." /> : null
9       }>
10      {messages.map((msg, index) => (
```



```

11     <Message
12       key={index}
13       model={msg}
14     >
15       {msg.type === "custom" && (
16         <Message.CustomContent>
17           <Markdown>{msg.message}</Markdown>
18         </Message.CustomContent>
19       )}
20     </Message>
21   )}
22 </MessageList>
23 <MessageInput
24   placeholder="Stelle deine Frage..."
25   disabled={inputDisabled}
26   onSend={handleSend}
27   attachButton={false}
28 />
29 </ChatContainer>
30 </MainContainer>
31 </div>

```

Listing 5.1: JSX Code für die Chat-Benutzeroberfläche

5.1.2.2 Logik

Beim Laden der Webseite wird ein neuer Chat mit dem KI-Tutor gestartet. Dafür wird der *useEffect()*-Hook von React verwendet, der einmalig beim Seitenladen ausgeführt wird. Dabei wird eine für den Nutzer unsichtbare Nachricht an den KI-Tutor geschickt, mit der Aufforderung, sich vorzustellen und den Nutzer zu motivieren, Fragen zu stellen. Dies soll die Aufmerksamkeit des Nutzers wecken und den Start des Chats mit dem KI-Tutor erleichtern. Der genaue Ablauf wird mit einem Sequenzdiagramm in Unterabschnitt 5.3.1 beschrieben.

Um ein interaktiveres Nutzererlebnis zu ermöglichen, werden die Antworten des KI-Tutors gestreamt. Das bedeutet, dass neue Textteile der Antwort in Echtzeit an das Frontend gesendet werden, sobald sie generiert werden. Zur korrekten Handhabung dieses Streams wurde im Frontend eine Helferfunktion namens *handleReadableStream(stream)* implementiert. Der Frage-Antwort Fluss ist auch in Unterabschnitt 5.3.2 genauer ausgeführt.

In den Antworten des KI-Tutors erscheinen regelmäßig Annotationen, die für den Nutzer nicht lesbar sind und den Lesefluss stören. Diese werden von der Assistant API mitgesendet, wenn der KI-Tutor Informationen aus der Wissensdatenbank verwendet, um Quellen anzugeben. Da diese Funktion in der aktuellen Implementierung noch nicht genutzt wird und um zu verhindern, dass die Nutzer durch diese Zeichen verwirrt werden, werden sie mithilfe der *removeAnnotations*-Funktion aus dem Antworttext herausgefiltert.

Sobald der Nutzer eine Nachricht an den KI-Tutor abgeschickt hat und auf die Antwort wartet, wird das Eingabefeld deaktiviert und ein Hinweis eingeblendet, dass der KI-Tutor

antwortet. Nachdem die Antwort abgeschlossen ist, wird das Texteingabefeld für den Nutzer wieder aktiviert. Die Deaktivierung erfolgt, sobald der Nutzer auf „Senden“ klickt, und die Aktivierung, sobald die Funktion `handleReadableStream` das `thread.run.completed`-Event empfängt, das signalisiert, dass die Antwort abgeschlossen ist.

5.1.3 Backend

Wie bereits oben erwähnt, dient das Backend in der Implementierung dieser Bachelorarbeit hauptsächlich dazu, den API-Schlüssel vor potenziellen Angreifern zu verbergen und ist daher bewusst einfach und minimal gehalten. Für zukünftige Erweiterungen des KI-Tutors bietet das Vorhandensein eines Backends jedoch zusätzliche Vorteile. Da alle Anfragen der Nutzer über den Server laufen, könnten in Zukunft Funktionen wie Fehler-Logs, die die Erkennung von Problemen erleichtern, oder ähnliche Erweiterungen einfacher implementiert werden.

Das Frontende sendet alle Anfragen an die Assistant API über das Backend. Das Backend fügt den API-Schlüssel den Anfragen hinzu und leitet diese sicher an die Assistant API weiter. Dafür stellt das Backend folgende Endpunkte zur Verfügung:

POST /api/threads: Dieser Endpunkt ist dafür da, einen neuen Chat Thread mit dem KI-Tutor zu erstellen. Das Backend leitet diese Anfrage an OpenAI's API weiter, wartet auf die Antwort und gibt dann dem Frontend die ThreadId zurück.

POST /api/threads/[threadId]: Dieser Endpunkt sendet eine neue Nachricht an einen bestehenden Chat-Thread mit dem KI-Tutor. Die Anfrage enthält den Inhalt der Nachricht, die an den Thread gesendet wird. Das Backend leitet die Nachricht an die OpenAI-API weiter, startet den Assistant Run auf dem Thread und streamt die Antwort des KI-Tutors in Echtzeit zurück.

5.1.4 Assistant API Schnittstelle

In dieser Sektion werden die Endpunkte der Assistant API beschrieben, die von der Web-App genutzt werden. Der KI-Tutor selbst, einschließlich der Instruktionen und der Wissensdatenbank, wurde über die OpenAI Entwicklerplattform [3] definiert und erstellt. Dieser KI-Tutor verfügt über eine eindeutige *AssistantID*, die bei den API-Endpunkten referenziert und verwendet wird.

POST <https://api.openai.com/v1/threads>: Dieser Endpunkt erstellt einen neuen Thread. Nach der erfolgreichen Erstellung wird ein neuer Thread zusammen mit der zugehörigen *thread_id* zurückgegeben.

POST https://api.openai.com/v1/threads/thread_id/messages: Mit diesem Endpunkt wird eine neue Nachricht in dem durch die *thread_id* definierten Thread erstellt.

POST https://api.openai.com/v1/threads/thread_id/runs: Dieser Endpunkt führt einen Assistant Run auf dem durch die *thread_id* definierten Thread aus. Der auszuführende Assistant wird im Body der Anfrage durch die *AssistantID* spezifiziert.

5.2 KI-Tutor (OpenAI Assistant)

Bei der Implementierung des KI-Tutors wurde die OpenAI Assistant API verwendet. Diese ermöglicht es, wie in den Grundlagen (Sektion 2.4) beschrieben, spezialisierte Chatbots auf Basis der GPT-Modelle von OpenAI zu erstellen. Dazu können dem Chatbot sowohl zusätzliche Instruktionen als auch weitere Daten als Wissensbasis zur Verfügung gestellt werden. In dieser Sektion erläutere ich, welche Instruktionen für die Erstellung des KI-Tutors verwendet wurden, aus welchen Gründen diese ausgewählt wurden und wie der Kursinhalt strukturiert und dem KI-Tutor in der Wissensdatenbank zur Verfügung gestellt wird.

5.2.1 Wissensdatenbank und Formatwahl

Für die Implementierung der Wissensdatenbank des KI-Tutors wurde ein textbasiertes Format gewählt. Diese Entscheidung basiert auf der Stärke von ChatGPT, das besonders effizient bei der Verarbeitung und Erstellung von textbasierten Inhalten ist. Die Wahl fiel dabei auf das Markdown-Format (.md), da es dem reinen Textformat sehr nahe kommt und somit direkt und ohne Umwandlung von ChatGPT verarbeitet werden kann. Ein weiterer Vorteil von Markdown ist, dass es trotz seiner Einfachheit die Möglichkeit bietet, hierarchische Informationen wie Überschriften und Abschnitte zu vermitteln. Dadurch lassen sich Kursinhalte aus dem Kurs für Menschen leichter und besser strukturiert in das Wissensdatenbank-Format übertragen. Diese zusätzliche Strukturierung der Informationen kann auch von ChatGPT potenziell genutzt werden, um präzisere und kontextgerechtere Antworten zu generieren.

Der Inhalt des Kurses wurde kapitelweise in einzelne Markdown-Dateien unterteilt, die jeweils die Kapitelnummer sowie den Kapiteltitel im Dateinamen enthalten. Diese Trennung soll dem KI-Tutor dabei helfen, Informationen gezielt aus dem Kapitel bereitzustellen, das am besten zur jeweiligen Nutzeranfrage passt. Der KI-Tutor kann pro Anfrage nur eine begrenzte Zeichenmenge verarbeiten. Die kapitelweise Aufteilung stellt sicher, dass keine Informationen abgeschnitten werden. Da die Kosten der Assistant API nach der Anzahl der verarbeiteten Zeichen berechnet werden, reduziert diese Struktur zudem die Kosten, da nicht immer der gesamte Kursinhalt betrachtet werden muss, sondern nur einzelne Kapitel.

Die einzelnen Dateien im Markdown-Format folgen einem einheitlichen Aufbau. Das Ziel war es, dem KI-Tutor möglichst alle im Kurs enthaltenen Informationen in textbasierter Form zur Verfügung zu stellen. Der Aufbau gliedert sich wie folgt:

1. Kapitelnummer und Titel

Kapitel 0.05: Nicht nur Äpfel...

Kursinhalt

Magst du Äpfel?

Nicht nur Äpfel...:

Wie sieht es mit Heidelbeeren, Pfirsichen, Pflaumen, Kiwis, Mangos, Melonen und Kirschen aus? Bist du verrückt nach Mandeln, Cashews oder Macadamias? Magst du es, anzuhalten und an Rosen zu riechen?

Wenn du eine dieser Fragen mit "Ja" beantwortet hast, kannst du dich bei den Honigbienen bedanken.

Warum brauchen wir Bienen?

Honigbienen sind für die Produktion von Honig und Bienenwachs bekannt, sowie für die große Rolle, die sie bei der Bestäubung von Pflanzen und Blumen spielen. Bienen übernehmen in gewissem Umfang die Bestäubung von fast 75% aller Pflanzenarten, die weltweit direkt für die menschliche Ernährung genutzt werden.

Wie viel von dem, was du diese Woche gegessen hast, wäre nicht mehr verfügbar, wenn die Bienen verschwinden würden?

Aufgaben

Ziehe Bilder von Produkten in die richtige Kategorie:

Aufgabenstellung

Ohne die Bienen würden wir viele Produkte verlieren, einige komplett und einige würden radikal reduziert werden. Ziehe die Produkte, von denen du denkst, dass wir sie zu 100 % und andere zu 90 % verlieren werden, in das richtige Feld:

Produkte: Honig, Heidelbeeren, Apfel, Karotte, Erdbeere, Mandeln, Gurke, Zwiebeln

Felder: 100% verlieren, 90% verlieren

Lösung (nicht herausgeben!):

Feld 100%: Honig, Mandeln

Feld 90%: Karotte, Gurke, Zwiebel, Apfel, Erdbeere, Heidelbeeren

Abbildung 5.2: Beispiel für den Aufbau der Markdown-Dateien in der Wissensdatenbank

2. Kursinhalt

3. Aufgaben

a) Fragen

b) Antworten

Dieser Aufbau kann auch an Abbildung 5.2 Anhand der Markdown Datei zu einem Kapitel aus dem Rettet die Bienen Kurs nachvollzogen werden.

Da die Aufgaben auf RockStartIT oft interaktiv gestaltet sind, beispielsweise durch Drag-and-Drop-Funktionen oder ähnliche Mechanismen, wurde diese Aufgabenarten so konkret wie möglich textuell beschrieben. Ebenso wurden die zugehörigen Lösungen in textlicher Form

bereitgestellt, um sicherzustellen, dass der KI-Tutor auch Fragen zu diese Aufgaben korrekt verarbeiten und beantworten kann. Die Lösungen wurden zusätzlich mit dem Hinweis „nicht herausgeben!“ versehen.

5.2.2 Instruktionen

Der wichtigste Bestandteil des KI-Tutors sind die Instruktionen. Diese werden der Assistant API übergeben und verwandeln das zugrundeliegende generische GPT-Modell in einen auf den Kurs zugeschnittenen KI-Tutor. Ein wesentlicher Teil dieser Arbeit bestand darin, durch *Prompt-Engineering* mit systematisches Ausprobieren und Dokumentieren herauszufinden, welche Instruktionen die gewünschten Ergebnisse liefern und welche nicht. Die zugehörige Dokumentation, bestehend aus dem Prompt-Verlauf und Anmerkungen zu den funktionierenden sowie nicht funktionierenden Ansätzen, ist im GitLab-Repository (siehe Kapitel 8) dieser Arbeit hinterlegt.

Mit den Instruktionen wurde versucht, die in Kapitel 4.1 definierten Anforderungen bestmöglich zu erfüllen. In Listing 5.2 sind die finalen Instruktionen dargestellt, und im Folgenden wird erläutert, welche Sektionen aus welchen Gründen enthalten sind.

```

1 This GPT is a chatbot tutor named KI Tutor, designed to assist students with questions
   related to an online course called "Rettet die Bienen!".
2
3 It should provide clear, concise answers, offer additional explanations when necessary, and
   be patient and encouraging. It should ALWAYS and only use the Information in the
   Knowledgebase to answer the questions. Only answer questions relevant to the course.
4
5 All responses should be written in German and kept concise. The students are around the age
   of 14 to 16 years old, so answer in that niveau.
6
7 When asked for solutions to a question that is in the 'Antworten' part of the files in the
   knowledge base, KI Tutor should NEVER give out the solution directly but instead guide
   and motivate students to find the answers themselves by giving tips, asking guiding
   questions, and encouraging critical thinking without directly providing the answers.

```

Listing 5.2: Instruktionen für den KI-Tutor

Der erste Teil „*This GPT is a chatbot tutor named KI Tutor, designed to assist students with questions related to an online course called ‚Rettet die Bienen!‘*“ definiert die Rolle des KI-Tutors. Hier wird die im Prompt-Engineering weit verbreitete Methode des Role-Prompting angewendet. In diesem Abschnitt wird seine grundlegende Aufgabe beschrieben und ihm wichtige Kontextinformationen, wie sein Name sowie der Name des Kurses in dem er eingesetzt wird, mitgeteilt. Dieser Abschnitt ist seit Anfang an unverändert geblieben, da er gut funktioniert hat und der Tutor daraus konstant gut seine Rolle ableiten konnte.

Im nächsten Abschnitt „*It should provide clear, concise answers, offer additional explanations when necessary, and be patient and encouraging.*“ wird sein Verhalten als KI-Tutor nochmals genauer definiert. Hierbei werden die Anforderungen *PA4.1: Motivationsförderung* und *PA2.1: Unterstützung bei Lerninhalten* beachtet.

Mit der Anweisung „*It should ALWAYS and only use the information in the Knowledgebase to answer the questions.*“ wird der KI-Tutor dazu angeleitet, stets auf die Wissensdatenbank zuzugreifen. Zu Beginn des Prompt-Engineering Prozesses griff der KI-Tutor häufig nicht auf die Wissensdatenbank zu, sondern beantwortete Fragen basierend auf dem allgemeinen Wissen im Modell oder erfand Kapitel, die im Kurs nicht existierten (ein Phänomen, das auch als Halluzinieren bezeichnet wird). Es war notwendig, verschiedene Ansätze auszuprobieren. Mit diesem Prompt konnte das Problem behoben werden, sodass der KI-Tutor seine Antworten verlässlicher aus der Wissensdatenbank bezieht.

Der KI-Tutor ließ sich außerdem leicht dazu verleiten, Fragen außerhalb des Kursinhalts zu beantworten. Dieses Problem konnte durch das Hinzufügen der Anweisung „*Only answer questions relevant to the course.*“ weitgehend behoben werden.

Im darauffolgenden Abschnitt „*All responses should be written in German and kept concise. The students are around the age of 14 to 16 years old, so answer in that niveau.*“ wird das Antwortverhalten des KI-Tutors weiter spezifiziert, wobei besonders die Anforderung PA1: Altersgerechte Sprache berücksichtigt wird. In den Tests gab der Tutor anfangs oft zu lange und teilweise zu komplizierte Antworten. Durch diesen Prompt konnte die Qualität und Verständlichkeit der Antworten deutlich verbessert werden.

Ein großes Problem, das sich lange nicht zufriedenstellend lösen ließ, war, dass der KI-Tutor die Lösungen zu den Aufgaben im Kurs direkt preisgab. Diese Lösungen wurden ihm über die Wissensdatenbank zur Verfügung gestellt, damit er die Lernenden zu den richtigen Ergebnissen führen kann, wobei er diese aber nicht direkt verraten sollte. Verschiedene Formulierungen, die den Tutor dazu aufforderten, die Lösungen nicht herauszugeben, wurden ausprobiert, jedoch oft ignoriert. Schließlich konnte das Problem größtenteils mit folgendem Abschnitt in den Instruktionen behoben werden:

„*When asked for solutions to a question that is in the ‚Aufgaben‘ part of the files in the knowledge base, KI Tutor should NEVER give out the solution directly but instead guide and motivate students to find the answers themselves by giving tips, asking guiding questions, and encouraging critical thinking without directly providing the answers.*“

Zusätzlich wurden die Lösungen in der Wissensdatenbank mit dem Hinweis „Lösung (nicht herausgeben!)“ versehen, was ebenfalls dazu beigetragen hat, das Problem weitgehend zu beheben. Diese Formulierung geht direkt auf die Anforderung PA2.2: Unterstützung bei Lerninhalten ein, indem sie den KI-Tutor dazu anweist, die Lernenden zur Lösung zu führen, ohne die Antworten direkt zu liefern.

5.3 Abläufe

Um die Interaktion der verschiedenen Komponenten der KI-Tutor-Web-App besser zu veranschaulichen und die beiden wichtigsten Abläufe, das Laden der Seite (Thread-Erstellung) und den Nachrichtenfluss (Nutzerinteraktion), detailliert zu erklären, werden für beide Abläufe jeweils Sequenzdiagramme mit einer Erklärung der einzelnen Schritte vorgestellt.

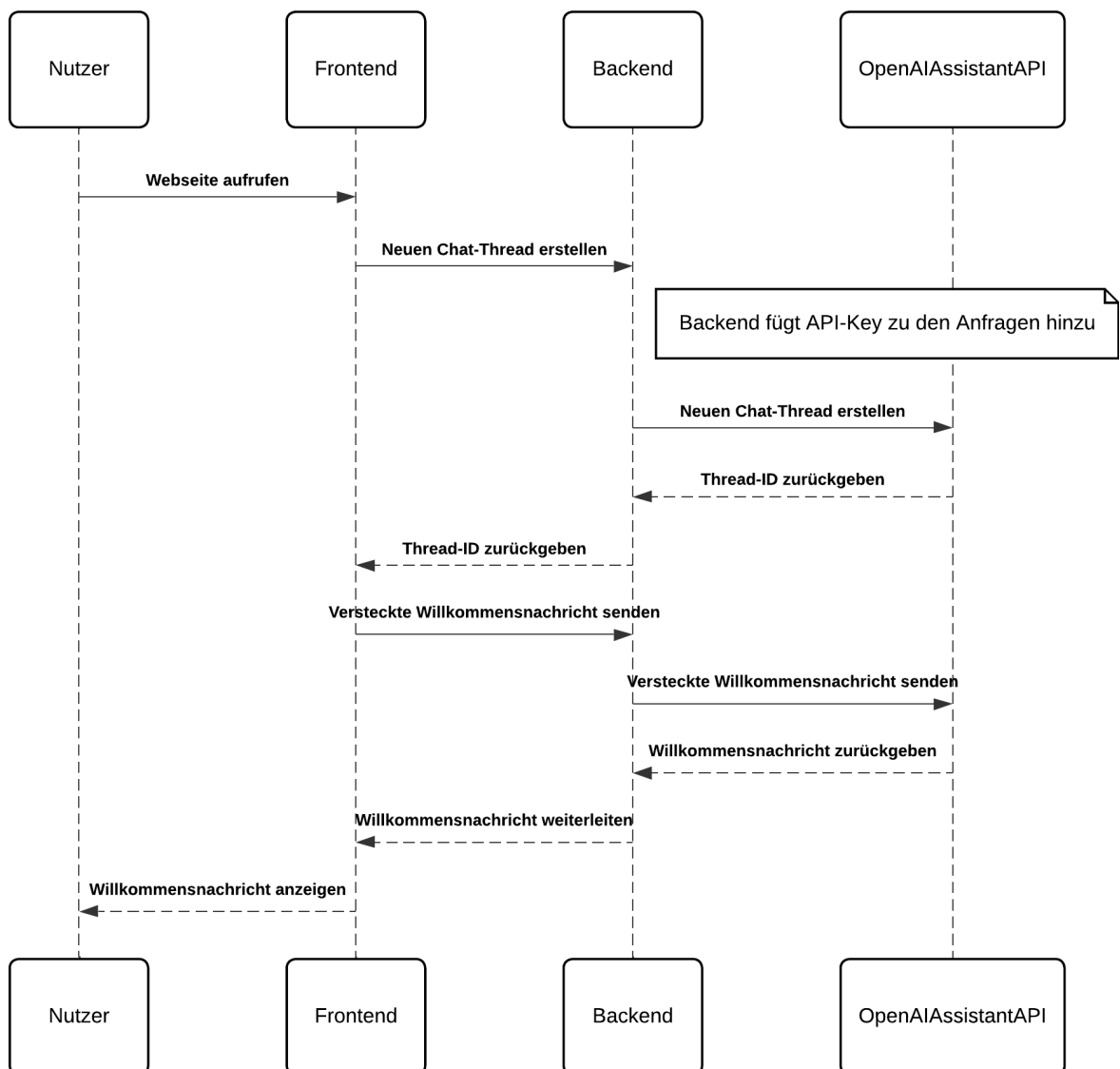


Abbildung 5.3: Ablauf Thread-Erstellung

5.3.1 Thread-Erstellung

Das Sequenzdiagramm in Abbildung 5.3 illustriert den Ablauf der Initialisierung eines neuen Chat-Threads sowie die Kommunikation zwischen dem Nutzer, dem Frontend, dem Backend und der Assistant API. Dieser Prozess wird eingeleitet, sobald ein Nutzer die Webseite aufruft und dadurch einen neuen Chat startet.

1. **Webseitenaufruf durch den Nutzer:** Der Prozess beginnt, wenn der Nutzer die Webseite aufruft. Diese Aktion wird an das Frontend der Anwendung weitergeleitet.
2. **Anfrage zur Erstellung eines neuen Chat-Threads:** Das Frontend sendet eine Anfrage an das Backend, um einen neuen Chat-Thread zu erstellen. Diese Anfrage ist

notwendig, um eine eindeutige und beständige Konversation zwischen dem Nutzer und dem KI-Tutor zu initiieren.

3. **Hinzufügen des API-Schlüssels:** Bevor die Anfrage an die Assistant API weitergeleitet wird, fügt das Backend den erforderlichen API-Schlüssel hinzu. Dieser Schritt ist entscheidend, um die Authentifizierung der Anfrage sicherzustellen und den Zugriff auf die Assistant API zu ermöglichen.
4. **Erstellung des Chat-Threads in der Assistant API:** Das Backend leitet die Anfrage zur Erstellung eines neuen Chat-Threads an die Assistant API weiter. Die API erstellt daraufhin einen neuen Chat-Thread und generiert eine eindeutige Thread-ID.
5. **Rückgabe der Thread-ID:** Die Assistant API sendet die generierte Thread-ID zurück an das Backend, welches diese wiederum an das Frontend weiterleitet. Diese ID wird für alle zukünftigen Interaktionen innerhalb dieses spezifischen Chats verwendet.
6. **Senden einer versteckten Willkommensnachricht:** Das Frontend initiiert eine versteckte Willkommensnachricht, die das Backend an die Assistant API weiterleitet. Diese Nachricht dient dazu, den KI-Tutor aufzufordern, sich dem Nutzer vorzustellen.
7. **Empfang und Weiterleitung der Willkommensnachricht:** Die Assistant API verarbeitet die versteckte Nachricht und generiert eine entsprechende Willkommensnachricht, die an das Backend zurückgesendet wird. Das Backend leitet diese Nachricht an das Frontend weiter.
8. **Anzeige der Willkommensnachricht:** Schließlich zeigt das Frontend die Willkommensnachricht dem Nutzer an. Diese Nachricht markiert den Abschluss der Initialisierungsphase und bereitet den Chat auf weitere Interaktionen vor.

5.3.2 Nutzerinteraktion

Das Sequenzdiagramm in Abbildung 5.4 beschreibt den Ablauf der Verarbeitung einer vom Nutzer eingegebenen Nachricht und zeigt die Kommunikation zwischen den verschiedenen Komponenten des Systems: dem Nutzer, dem Frontend, dem Backend und der Assistant API. Dieser Prozess umfasst das Senden der Nutzeranfrage an die API, das Streamen der Antwort in Teilen und die abschließende Anzeige der vollständigen Antwort.

1. **Nachrichteneingabe durch den Nutzer:** Der Prozess beginnt, wenn der Nutzer eine Nachricht in das Textfeld im Frontend eingibt.
2. **Weiterleitung der Nachricht an das Backend:** Sobald der Nutzer auf senden klickt, leitet das Frontend die eingegebene Nachricht an das Backend weiter.
3. **Senden der Nachricht an die Assistant API:** Das Backend empfängt die Nachricht und sendet sie an die Assistant API. Diese API ist dafür verantwortlich, die Eingabe des Nutzers zu verarbeiten und eine entsprechende Antwort zu generieren.

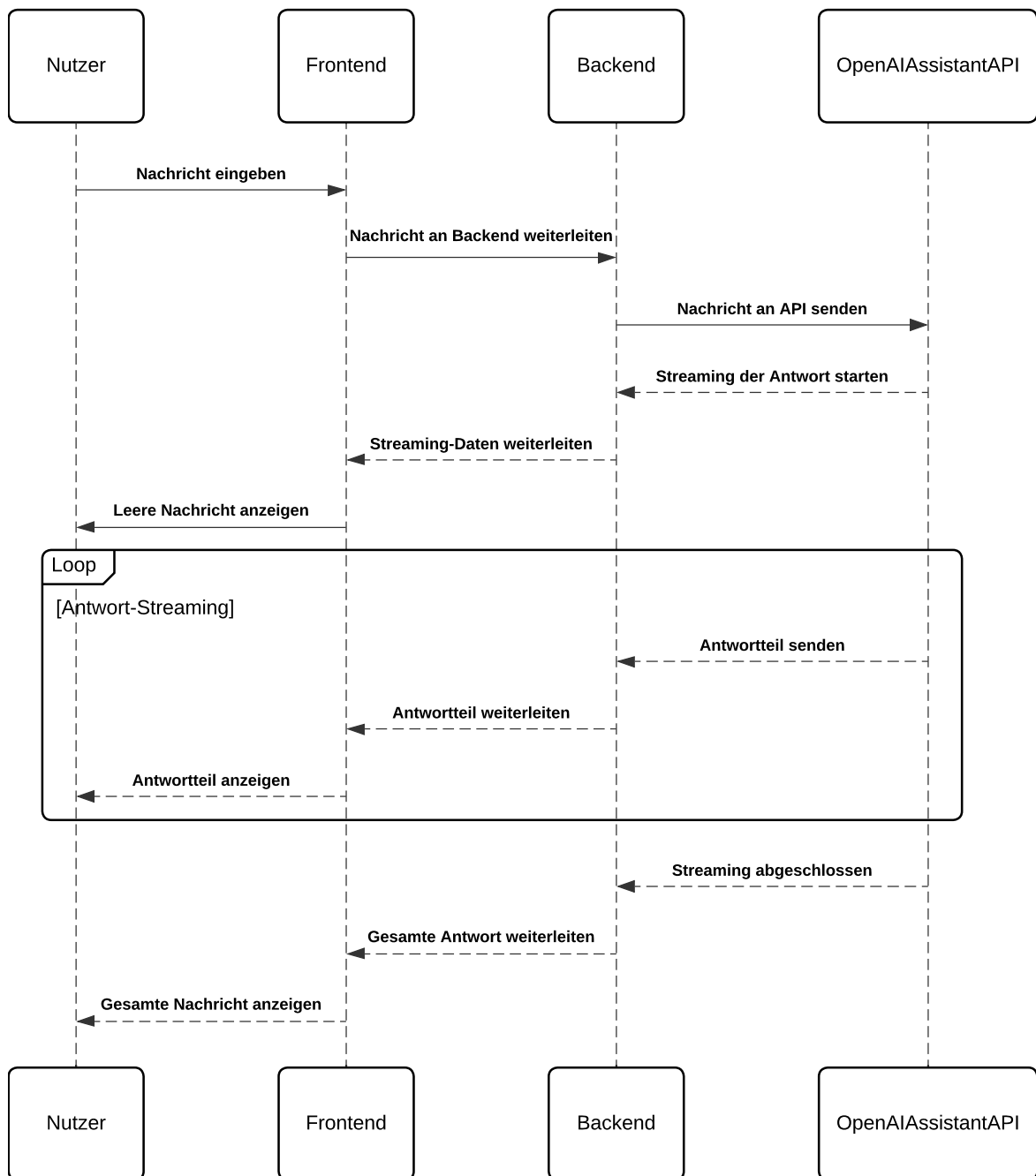


Abbildung 5.4: Ablauf Nachricht Antwort Fluss

4. **Start des Antwort-Streamings:** Die OpenAIAssistantAPI beginnt mit dem Streaming der Antwort. Anstatt die gesamte Antwort auf einmal zu senden, wird sie in kleinen Teilen (Chunks) gestreamt, um eine schnellere Reaktionszeit und eine bessere Nutzererfahrung zu ermöglichen.
5. **Weiterleitung der Streaming-Daten an das Frontend:** Das Backend empfängt die gestreamten Antwortteile und leitet diese sofort an das Frontend weiter. Dadurch kann das Frontend die empfangenen Teile der Antwort unmittelbar anzeigen.

6. **Anzeige einer leeren Nachricht:** Während die Antwortdaten gestreamt werden, zeigt das Frontend zunächst eine leere Nachricht an, um dem Nutzer zu signalisieren, dass die Antwort verarbeitet wird und in Kürze angezeigt wird.
7. **Schleife zur Anzeige der gestreamten Antwortteile:** In einer Schleife wird jeder empfangene Antwortteil vom Backend an das Frontend weitergeleitet. Das Frontend fügt diesen Antwortteil der Nachricht hinzu, wodurch die Nachricht nach und nach für den Nutzer sichtbar wird.
 - **Empfang der Antwortteile durch das Backend:** Die Assistant API sendet die Antwortteile in mehreren Paketen, die das Backend empfängt.
 - **Weiterleitung der Antwortteile an das Frontend:** Das Backend leitet jeden empfangenen Antwortteil unmittelbar an das Frontend weiter.
 - **Anzeige der Antwortteile:** Das Frontend fügt jeden Antwortteil der Antwort hinzu, wodurch die Antwort nach und nach für den Nutzer sichtbar wird.
8. **Abschluss des Antwort-Streamings:** Nachdem alle Teile der Antwort gesendet wurden, signalisiert die Assistant API das Ende des Streamings. Dieser Schritt beendet den Streaming-Prozess.
9. **Weiterleitung der gesamten Antwort an das Frontend:** Das Backend leitet die vollständige Antwort an das Frontend weiter.
10. **Anzeige der gesamten Nachricht an den Nutzer:** Das Frontend zeigt die vollständig zusammengesetzte Nachricht dem Nutzer an. Damit ist die Antwort auf die ursprüngliche Nutzeranfrage vollständig abgeschlossen und sichtbar.

5.4 Ergebnisse

Um einen besseren Eindruck von den Ergebnissen der Implementierung zu vermitteln, werden im Folgenden Screenshots der bereits oben beschriebenen Elemente der KI-Tutor-Web-App präsentiert.

In Abbildung 5.5 ist zu erkennen, wie der KI-Tutor sich beim Aufruf der Seite zunächst vorstellt und den Nutzer dazu ermutigt, Fragen zu stellen. Anschließend ist eine Beispielinteraktion zu sehen, bei der der Nutzer äußert, dass er nicht ganz verstanden hat, was eine Hypothese bedeutet. In der darauffolgenden Antwort des KI-Tutors wird sein Antwortverhalten deutlich: Zunächst erklärt er den Begriff und bietet anschließend ein Beispiel an, das auf den Kurskontext (Bienen) zugeschnitten ist. Am Ende seiner Antwort regt er den Nutzer erneut zum Nachfragen an.

In Abbildung 5.6 ist zu sehen, wie die Benutzeroberfläche aussieht, während der KI-Tutor gerade eine Antwort formuliert. Zum einen sind bereits Teile der Nachricht des KI-Tutors sichtbar, die sich weiter füllen, bis die vollständige Antwort angezeigt wird. Zudem weist eine Animation mit dem Text „KI-Tutor schreibt...“ darauf hin, dass die Antwort noch

KI-Tutor

Hallo! Ich bin KI Tutor, dein Chatbot-Tutor für den Kurs "Rettet die Bienen!". Ich freue mich darauf, dir bei deinen Fragen zu helfen und gemeinsam mit dir mehr über Bienen und deren Schutz zu lernen. Zögere nicht, Fragen zu stellen – es gibt keine falschen Fragen! Lass uns zusammen Anliegen klären. 🐝🐝

Hi, ich verstehe nicht ganz was Hypothese bedeutet

Eine Hypothese ist eine Annahme oder Vermutung, die du aufstellst, um etwas zu erklären oder herauszufinden. Sie ist oft der Ausgangspunkt für ein Experiment oder eine Untersuchung. Man testet die Hypothese, um zu sehen, ob sie wahr oder falsch ist.

Zum Beispiel: Wenn du denkst, dass Bienen schneller fliegen, wenn es sonnig ist, kannst du diese Hypothese testen, indem du ihre Fluggeschwindigkeit bei verschiedenen Wetterbedingungen beobachtest.

Wenn du weitere Fragen hast oder ein Beispiel brauchst, frag einfach!

Stelle deine Frage...



Abbildung 5.5: Beispiel Interaktion mit dem KI-Tutor

KI-Tutor

Hallo! Ich bin KI Tutor, dein freundlicher Begleiter für das Online-Kurs "Rettet die Bienen!". Ich bin hier, um dir bei Fragen zu helfen und dir das Verständnis der Themen zu erleichtern. Zögere nicht, Fragen zu stellen –

... KI-Tutor schreibt...

Stelle deine Frage...



Abbildung 5.6: Benutzeroberfläche während der KI-Tutor antwortet

im Entstehen ist. Während dieser Zeit ist das Eingabefeld deaktiviert, erkennbar an der ausgegrauten Fläche.

5.5 Code Review

Ein Bestandteil der Arbeit war auch ein Code Review durch eine unabhängige dritte Person. Das grundlegende Feedback fiel positiv aus, jedoch wurden an einigen Stellen Anmerkungen gemacht, auf die im Folgenden eingegangen wird.

5.5.1 Markdown Unterstützung in KI-Tutor Antworten

Der KI-Tutor verwendete in seinen Antworten gelegentlich Markdown-Formatierungen, wie z. B. „**Beispieltext**“ für Fettschrift, die in einem Markdown-Reader korrekt angezeigt würden. Da der KI-Tutor jedoch keine Markdown-Unterstützung hatte, wurden diese Formatierungen für den Nutzer als störende oder verwirrende Zeichen dargestellt. Die Verwendung von Markdown durch den KI-Tutor liegt vermutlich daran, dass die Wissensdatenbank aus Markdown-Dateien besteht. Da die Markdown-Formatierung als potentiell sinnvoll zur besseren Übersicht erachtet wurde, wurde das Problem mithilfe der Bibliothek *react-markdown* [25] gelöst. Durch den Einsatz dieser Bibliothek werden die vom KI-Tutor mit Markdown formatierten Elemente korrekt als Fett-, Kursivschrift usw. dargestellt.

5.5.2 Quellen Angaben

In der Review wurde vom Reviewer angemerkt, dass es beim Testen des KI-Tutors nützlich sein könnte zu sehen, auf welche Dateien sich dieser bezieht, um bei Fehlverhalten besser nachvollziehen zu können, woran es gelegen haben könnte. Die mögliche Umsetzung dieser Funktionalität wurde nach der Review noch einmal überprüft, und tatsächlich bietet OpenAI diese Möglichkeit in der Assistant API an. Bei jeder Antwort des Assistants wird ein Annotations-Objekt mitgeschickt, das angibt, ob und welche Dateien aus der Wissensdatenbank als Quelle für die Antwort verwendet wurden. Allerdings werden diese in der aktuellen API-Version nur als kryptische File-IDs übermittelt, die nicht direkt menschenlesbar sind. Um diese Funktionalität nutzbar zu machen, wäre eine zusätzliche Anfrage erforderlich gewesen, um mithilfe der ID den entsprechenden Dateinamen abzurufen. Da diese Funktionalität für die vorliegende Arbeit nicht als essenziell angesehen wurde, wurde in Absprache mit dem Betreuer entschieden, auf die Implementierung dieser Funktion zu verzichten.

5.5.3 Magic Strings

Zum Zeitpunkt der Review waren in der Implementierung noch mehrere sogenannte „Magic Strings“ vorhanden, also Zeichenketten, die direkt im Code verwendet wurden, ohne in Variablen ausgelagert zu sein. Der Reviewer empfahl, diese zu überprüfen und Strings, die mehrfach im Code vorkommen, in Variablen auszulagern. Dies wurde im Anschluss an die Review überprüft, und entsprechende Stellen wurden behoben.

5.5.4 Umgebungsvariablen

Der Reviewer merkte an, dass es neben dem bereits als Umgebungsvariable gespeicherten API-Schlüssel auch sinnvoll wäre, die Organisations-ID sowie die Assistant-ID in Umgebungsvariablen auszulagern, um diese flexibler und einfacher austauschbar zu machen. Diese Anpassung wurde nach der Review ebenfalls umgesetzt.

5.5.5 Testen der Web-App

In der Review wurden auch Tests angesprochen, die zum Zeitpunkt der Review noch nicht vorhanden waren. Der Reviewer merkte an, dass aufgrund der nicht allzu komplexen Logik der Implementierung einfache End-to-End-Tests ausreichend seien. Dieser Empfehlung wurde gefolgt, und es wurden Tests mit dem Cypress Test Framework erstellt, um die wichtigsten Abläufe des KI-Tutors (siehe 5.3) zu prüfen. Dafür wurden zwei Testfälle entwickelt.

Im ersten Test wird das Laden der Seite überprüft. Es wird getestet, ob die Web-App einen neuen Thread erstellt und eine Willkommensnachricht generiert, sowie korrekt anzeigt.

Im zweiten Test wird die Nutzerinteraktion mit dem Frage-Antwort-Fluss geprüft. Dabei wird überprüft, ob der Nutzer eine Nachricht eingeben und abschicken kann, diese korrekt angezeigt wird, die Texteingabe während der Wartezeit auf die Antwort deaktiviert ist und ob die Web-App die Antwort des KI-Tutors korrekt anzeigt.

6 Evaluation

In diesem Kapitel wird die Evaluation der prototypischen Implementierung des KI-Tutors für einen Online-Kurs auf der Plattform RockStartIT erläutert. Ziel der Untersuchung ist es, die Akzeptanz und Nutzung des KI-Tutors durch Lehrende und Lernende mithilfe des Technology Acceptance Models (TAM) von Davis [11] zu evaluieren. Dabei wird besonderes Augenmerk darauf gelegt, wie der KI-Tutor wahrgenommen wird, welche Faktoren seine Akzeptanz beeinflussen und inwiefern er den Lernprozess unterstützt.

6.1 Design

Die Studie ist eine über einen Zeitraum von zwei Wochen durchgeführte Post-Use-Evaluation, bei der die Probanden den prototypischen KI-Tutor anhand eines Online-Kurses von RockstartIT ausprobieren konnten und anschließend einen Online-Fragebogen dazu ausfüllen sollten. Der Fragebogen bestand aus zwei Komponenten: einem quantitativen Teil, in dem mit dem TAM-Modell die Akzeptanz und Nutzungsabsicht evaluiert wurde, sowie einem qualitativen Teil mit offenen Fragen zur Nutzungserfahrung, um die Ergebnisse des ersten Teils besser interpretieren zu können und Verbesserungspotenzial abzuleiten.

Ziel der Studie war es, die Akzeptanz von Lehrenden und Lernenden gegenüber KI-Tutoren zu untersuchen. Um die Sicht der Lehrenden abzudecken, wurden Lehrkräfte und angehende Lehrkräfte per E-Mail zur Teilnahme eingeladen. Die Kontakte dafür stammten einerseits von meinem Betreuer und andererseits aus persönlichen Kontakten mit Lehrkräften im Studium.

Da es aus zeitlichen Gründen im Rahmen dieser Arbeit nicht möglich war, Schülerinnen und Schüler der entsprechenden Altersgruppe zu befragen, wurden Kommilitonen sowie Studierende anderer Studiengänge eingeladen, an der Studie teilzunehmen. Sie erhielten den Hinweis, dass der Kurs für Lernende der 7. bis 10. Klasse konzipiert wurde, und wurden gebeten, den KI-Tutor möglichst aus dieser Perspektive zu testen.

6.2 Fragebogen

In dieser Sektion wird der verwendete Fragebogen vorgestellt und die Auswahl der Fragen erläutert. Der Fragebogen besteht aus zwei Teilen: einem quantitativen Abschnitt mit Fragen, die auf dem TAM-Modell von Davis [11] basieren, sowie einem qualitativen Abschnitt mit

offenen Fragen, die darauf abzielen, die Erfahrungen der Probanden mit dem KI-Tutor zu erfassen.

6.2.1 Demografische Daten

Zunächst werden im Fragebogen demografische Daten erfasst:

1. Sind Sie als Lehrkraft tätig oder befinden Sie sich in der Ausbildung zur Lehrkraft?
2. Mit welchem Geschlecht identifizieren Sie sich?
3. Wie alt sind Sie?

Diese Daten werden erhoben, um mögliche Muster in den Antworten in Bezug auf Geschlecht und Alter zu erkennen. Besonders wichtig ist die Frage nach der Tätigkeit als Lehrkraft, da diese verwendet wird, um den Fragebogen anzupassen. Lehrkräfte erhalten zusätzliche Fragen, um die Perspektive der Lehrenden auf den KI-Tutor besser zu erfassen.

6.2.2 Quantitative Fragen

Die Fragen im quantitativen Teil des Fragebogens basieren auf dem TAM-Modell von Davis [11] und wurden an den Kontext des KI-Tutors angepasst. Die Fragen bestehen aus mehreren Aussagen, die jeweils einer der Komponenten des TAM-Modells zugeordnet sind: *Perceived Ease of Use* (wahrgenommene Benutzerfreundlichkeit), *Perceived Usefulness* (wahrgenommene Nützlichkeit) und *Intention to Use* (Nutzungsabsicht). Die Überfrage zu jeder dieser Kategorien lautete: „Bitte bewerten Sie folgende Aussagen auf einer Skala von „stimme gar nicht zu“ bis „stimme voll und ganz zu“.“

Die Probanden hatten die Möglichkeit, jede Aussage auf einer 5-stufigen Likert-Skala zu bewerten, mit den Abstufungen: „stimme gar nicht zu“, „stimme eher nicht zu“, „unentschieden“, „stimme eher zu“ und „stimme voll und ganz zu“. Alternativ konnten sie „Keine Antwort“ wählen, wenn sie keine Bewertung abgeben wollten.

Perceived Usefulness (PU):

- PU1 Der KI-Tutor unterstützt mich dabei, die Themen leichter zu verstehen.
- PU2 Das Antwortverhalten des KI-Tutors war zuverlässig.
- PU3 Durch den KI-Tutor wird das Lernerlebnis verbessert.
- PU4 Ich denke, dass der KI-Tutor Lernende sinnvoll bei Aufgaben unterstützen kann.
- PU5 Ich finde den KI-Tutor nützlich.

Zusätzlich für Lehrkräfte:

- PU6 Ich denke, dass ein KI-Tutor wie dieser nützlich für meine Arbeit/für den Unterricht sein kann.
- PU7 Durch den Einsatz von einem KI-Tutor im Unterricht wird meine Arbeit erleichtert.
- PU8 Der KI-Tutor kann eine nützliche Ergänzung zu meinen Lehrmaterialien sein.

Perceived Ease of Use (PEOU):

- PEOU1 Die Interaktion mit dem KI-Tutor ist klar und verständlich.
- PEOU2 Ich finde die Bedienung des KI-Tutors insgesamt unkompliziert.
- PEOU3 Es ist einfach, den KI-Tutor zu verwenden, auch ohne technische Vorkenntnisse.
- PEOU4 Ich finde der KI-Tutor macht ohne Probleme das was ich möchte.
- PEOU5 Ich finde die Nutzung des KI-Tutors einfach.

Intention to Use (IU):

- IU1 Angenommen ich habe Zugang zu so einem KI-Tutor, dann beabsichtige ich ihn zu nutzen.
- IU2 Wenn ich Zugang zu so einem KI-Tutor habe, sage ich voraus, dass ich ihn nutzen werde.

Zusätzlich für Lehrkräfte:

- IU3 Ich kann mir vorstellen so einen KI-Tutor regelmäßig in meiner Arbeit/im Unterricht zu benutzen.

6.2.3 Qualitative Fragen

Die qualitativen Fragen des Fragebogens dienen dazu, die Ergebnisse des ersten Studienteils besser zu verstehen und potenzielle Verbesserungspotenziale des KI-Tutors zu identifizieren. Die Fragen decken dabei verschiedene Bereiche ab, die im Folgenden näher erläutert werden.

- Q1 Welche Herausforderungen sind Ihnen bei der Nutzung des KI-Tutors begegnet?

Diese Frage zielt auf die *Perceived Ease of Use* (PEOU)-Komponente des TAM-Modells ab. Sie soll mögliche Schwierigkeiten bei der Nutzung des KI-Tutors aufdecken, um potenzielle Probleme zu identifizieren, die in Zukunft verbessert werden könnten. Zudem können dadurch auch Ergebnisse aus dem ersten Teil der Studie besser erklärt werden, indem sie in den Kontext der Nutzererfahrung gesetzt werden.

Q2 Haben Sie Vorschläge, wie der KI-Tutor weiter verbessert werden könnte?

Diese Frage zielt darauf ab, konkrete Verbesserungsvorschläge der Probanden zu sammeln. Sie ermöglicht es, direktes Feedback von den Nutzern zu erhalten, um den KI-Tutor gezielt weiterzuentwickeln und auf die Bedürfnisse der Zielgruppe anzupassen.

Q3 Gab es Situationen, in denen Sie den KI-Tutor besonders hilfreich bzw. weniger hilfreich fanden?

Diese Frage hilft, die Wahrnehmung der *Perceived Usefulness* (PU)-Komponente des TAM-Modells zu bewerten. Durch das Erkennen hilfreicher oder weniger hilfreicher Situationen kann festgestellt werden, in welchen Szenarien der KI-Tutor besonders nützlich ist oder seine Funktionen verbessert werden müssen.

Q4 Denken Sie, dass KI-Tutoren wie dieser den Unterricht in Zukunft sinnvoll ergänzen könnten? Warum oder warum nicht?

Diese Frage adressiert bei den Lehrenden die *Intention to Use* (IU)-Komponente des TAM-Modells und soll erfassen, wie die Probanden den zukünftigen Einsatz von KI-Tutoren im Unterricht einschätzen. Hier das Potenzial von KI-Tutoren in Bildungskontexten hinterfragt, um zu verstehen, ob und wie sie den Unterricht potentiell bereichern könnten.

6.3 Ergebnisse

An der Studie nahmen insgesamt dreizehn Probanden teil (siehe Tabelle 6.1), davon sechs Lehrkräfte oder angehende Lehrkräfte und sieben Lernende. Von den Lehrkräften waren fünf männlich und eine weiblich, während unter den Lernenden ebenfalls fünf männlich waren und zwei weiblich. Die Lehrkräfte hatten ein Durchschnittsalter von 32,5 Jahren und die Lernenden ein Durchschnittsalter von 23,7 Jahren.

Daten	Lehrende	Lernende	Gesamt
Teilnehmer gesamt	6	7	13
Anzahl männlich	5	5	10
Anzahl weiblich	1	2	3
Altersdurchschnitt	32,5	23,7	27,8

Tabelle 6.1: Übersicht über die Probanden

6.3.1 Perceived Usefulness

In diesem Abschnitt werden die Ergebnisse der Kategorie *Perceived Usefulness* sowohl textuell beschrieben als auch grafisch in gestapelten Balkendiagrammen dargestellt. Zusätzlich sind

in Tabelle 6.2 die deskriptiven Statistiken zu den einzelnen Aussagen aufgeführt. Hierfür wurden den Bewertungen numerische Werte zugeordnet, von 1 „stimme gar nicht zu“ bis 5 „stimme voll und ganz zu“.

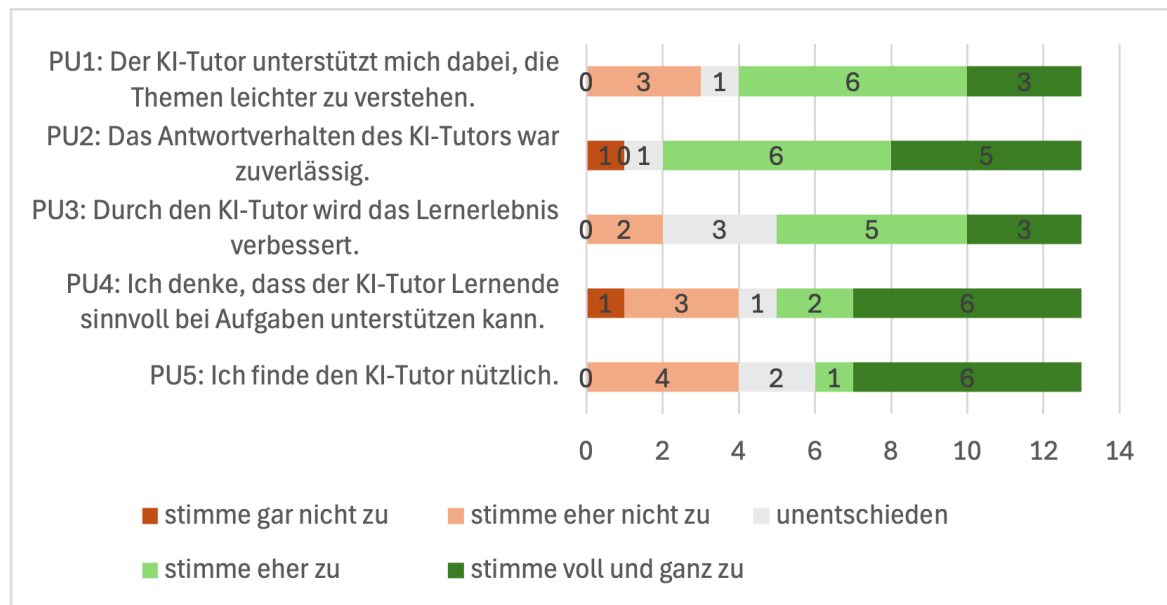


Abbildung 6.1: Perceived Usefulness (PU) von Lehrenden und Lernenden

Abbildung 6.1 zeigt die kombinierten Ergebnisse aller Probanden, sowohl Lehrende als auch Lernende. Die Mehrheit der Aussagen wurde überwiegend positiv bewertet. Besonders auffällig ist Aussage PU2, die sehr positiv beurteilt wurde und nur eine negative Bewertung erhielt. Aussage PU4 und PU5 wurden zwar ebenfalls mehrheitlich positiv bewertet, jedoch teilweise auch kritisch betrachtet: Beide Aussagen erhielten jeweils 4 von 13 Bewertungen auf der negativen Seite des Spektrums.

Wenn man die Bewertungen der Lehrenden (Abbildung 6.2) und Lernenden (Abbildung 6.3) separat betrachtet, zeigen sich folgende Ergebnisse: Die Lehrenden bewerteten die Aussagen insgesamt positiver als die Lernenden. Bei den Lehrenden wurden alle Aussagen überwiegend positiv beurteilt, wobei nur vereinzelt negative Bewertungen (ein bis zwei von sechs Probanden) auftraten. Besonders positiv fielen die Aussagen PU2 und PU8 aus, die durchweg positiv bewertet wurden, jeweils drei Mal mit „stimme eher zu“ und drei Mal mit „stimme voll und ganz zu“. Auch PU3 wurde sehr positiv bewertet und erhielt keine negativen Bewertungen. Am kritischsten wurde die Aussage PU6 bewertet, mit zwei Mal „stimme eher nicht zu“.

Bei den Lernenden fielen die Bewertungen ebenfalls mehrheitlich positiv aus, jedoch mit mehr negativen Bewertungen als bei den Lehrenden. Am positivsten schnitten die Aussagen PU1 und PU2 ab, während PU4 und PU5 am schlechtesten, jedoch immer noch überwiegend positiv, mit einem Mittelwert von 3.43 (siehe Tabelle 6.2), bewertet wurden.

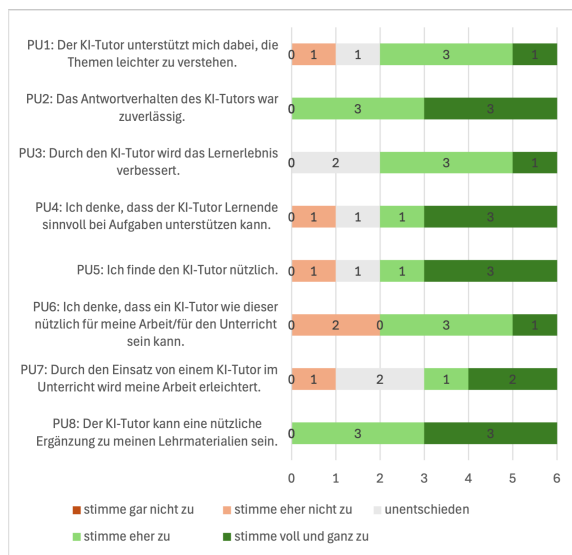


Abbildung 6.2: PU der Lehrenden

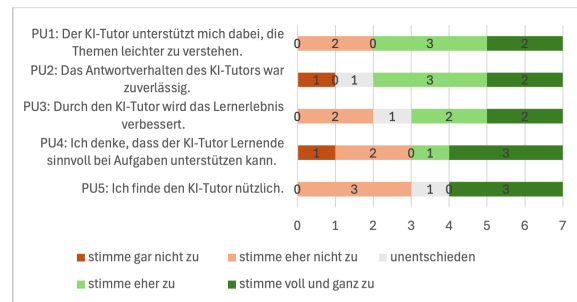


Abbildung 6.3: PU der Lernenden

6.3.2 Perceived Ease of Use

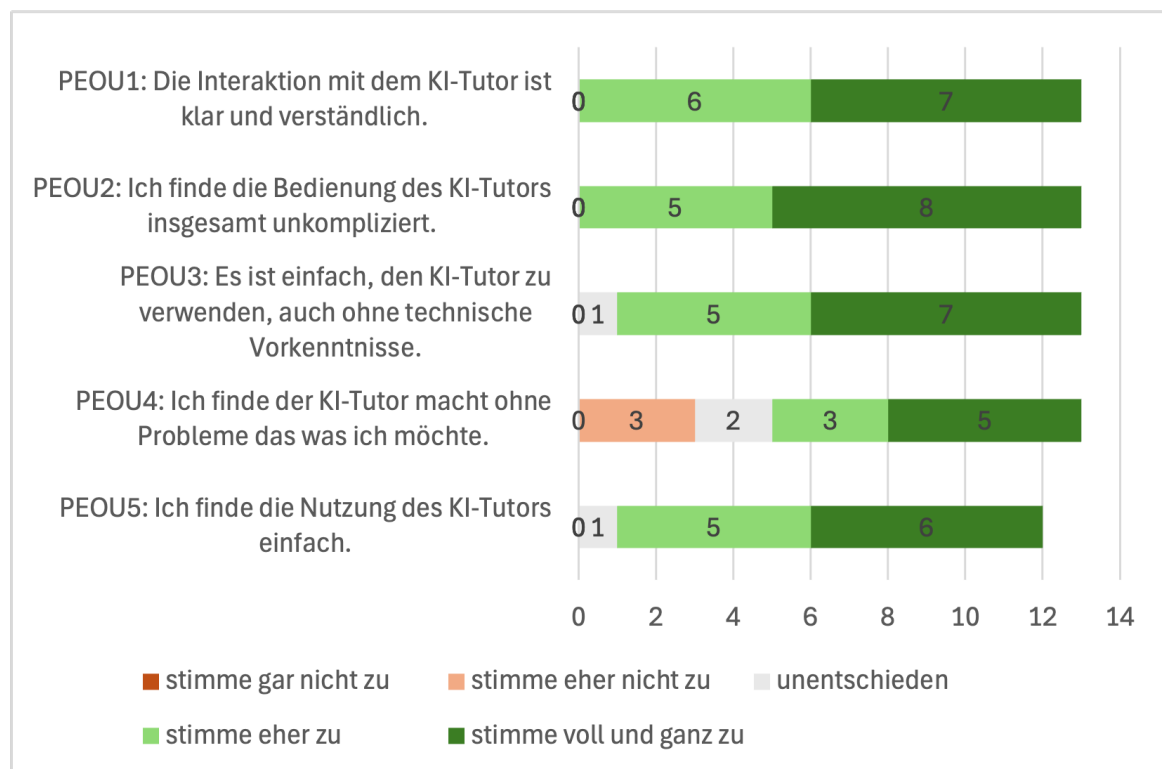


Abbildung 6.4: Perceived Ease of Use (PEOU) von Lehrenden und Lernenden

Bei der Komponente *Perceived Ease of Use* wurden die Aussagen fast ausschließlich positiv bewertet (siehe Abbildung 6.4), lediglich Aussage PEOU4: „Ich finde, der KI-Tutor macht ohne Probleme das, was ich möchte.“ erhielt überhaupt negative Bewertungen. Aussage

Aussage	n	Mittelwert	Median	Std.-abw.
PU1: Der KI-Tutor unterstützt mich dabei, die Themen leichter zu verstehen.	13	3.69	4	1.11
<i>nur Lehrende</i>	6	3.67	4	1.03
<i>nur Lernende</i>	7	3.71	4	1.25
PU2: Das Antwortverhalten des KI-Tutors war zuverlässig.	13	4.08	4	1.12
<i>nur Lehrende</i>	6	4.50	4.5	0.55
<i>nur Lernende</i>	7	3.71	4	1.38
PU3: Durch den KI-Tutor wird das Lernerlebnis verbessert.	13	3.69	4	1.03
<i>nur Lehrende</i>	6	3.83	4	0.75
<i>nur Lernende</i>	7	3.57	4	1.27
PU4: Ich denke, dass der KI-Tutor Lernende sinnvoll bei Aufgaben unterstützen kann.	13	3.69	4	1.49
<i>nur Lehrende</i>	6	4.00	4.5	1.27
<i>nur Lernende</i>	7	3.43	4	1.71
PU5: Ich finde den KI-Tutor nützlich.	13	3.69	4	1.38
<i>nur Lehrende</i>	6	4.00	4.5	1.27
<i>nur Lernende</i>	7	3.43	3	1.51
PU6: Ich denke, dass ein KI-Tutor wie dieser nützlich für meine Arbeit/für den Unterricht sein kann. (<i>nur Lehrende</i>)	6	3.50	4	1.23
PU7: Durch den Einsatz von einem KI-Tutor im Unterricht wird meine Arbeit erleichtert. (<i>nur Lehrende</i>)	6	3.67	3.5	1.21
PU8: Der KI-Tutor kann eine nützliche Ergänzung zu meinen Lehrmaterialien sein. (<i>nur Lehrende</i>)	6	4.50	4.5	0.55

Tabelle 6.2: Deskriptive Statistik der Perceived Usefulness Aussagen

PEOU2 schnitt am besten ab, mit acht Mal „stimme voll und ganz zu“ und fünf Mal „stimme eher zu“. Auch Aussage PEOU1 wurde ausschließlich positiv bewertet.

Bei getrennter Betrachtung der Bewertungen von Lehrenden (Abbildung 6.5) und Lernenden (Abbildung 6.6) ist zu erkennen, dass die Lernenden diese Komponente tendenziell besser bewertet haben. Aber auch die Bewertungen von den Lehrenden sind sehr positiv. Alle Aussagen, außer PEOU4, erhielten von den Lehrenden ausschließlich die Bewertungen „stimme eher zu“ und „stimme voll und ganz zu“. Aussage PEOU4 schnitt am schlechtesten ab, mit zwei Mal „stimme eher nicht zu“.

Bei den Lernenden erhielten nur PEOU1 und PEOU2 ausschließlich positive Bewertungen. PEOU1 ist dabei über beide Gruppen hinweg die am besten bewertete Aussage, mit sechs Mal „stimme voll und ganz zu“ und ein Mal „stimme eher zu“. Auch bei den Lernenden schnitt PEOU4 am schlechtesten ab.

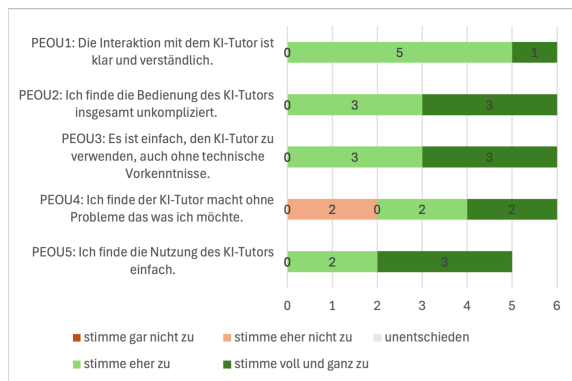


Abbildung 6.5: PEOU der Lehrenden

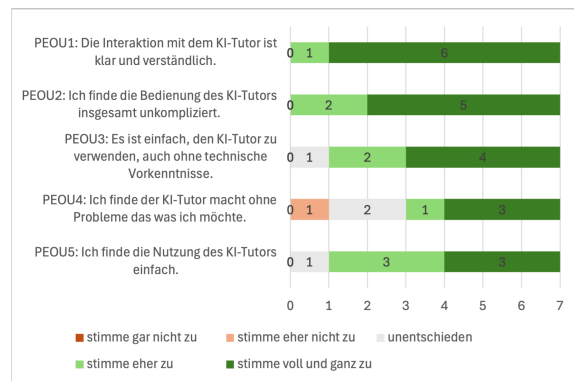


Abbildung 6.6: PEOU der Lernenden

6.3.3 Intention to Use

Bei der gemeinsamen Betrachtung der Bewertungen von Lehrenden und Lernenden (Abbildung 6.7) zur *Intention to Use*-Komponente der Studie wurden beide in diesem Abschnitt enthaltenen Aussagen insgesamt positiv bewertet. Aussage IU1 schnitt mit drei Mal „stimme eher zu“ und sechs Mal „stimme voll und ganz zu“ etwas besser ab als IU2. Allerdings erhielt IU1 auch die einzige „stimme gar nicht zu“-Bewertung in dieser Komponente.

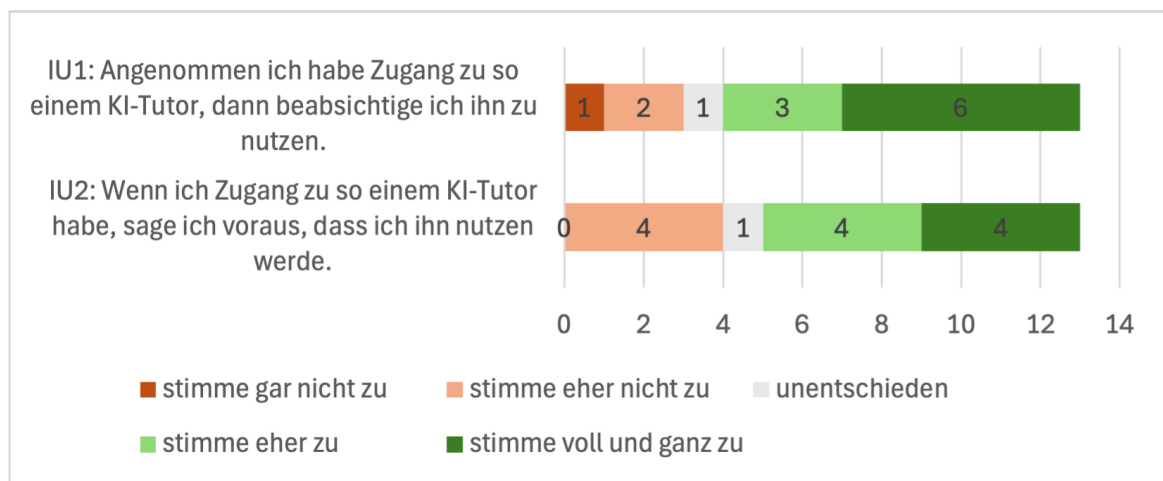


Abbildung 6.7: Intention to Use (IU) von Lehrenden und Lernenden

Wenn man nur die Lernenden betrachtet (Abbildung 6.9), zeigt sich dasselbe Muster wie in der Gesamtübersicht. Aussage IU1 schneidet etwas besser ab als IU2, weist jedoch eine größere Streuung der Meinungen auf, einschließlich einer „stimme gar nicht zu“-Bewertung. Bei den Lehrenden wurde ebenfalls Aussage IU1 am besten bewertet. Zusätzlich wurde bei den Lehrenden mit IU3 eine weitere Frage abgefragt, die genauso positiv bewertet wurde wie IU1.

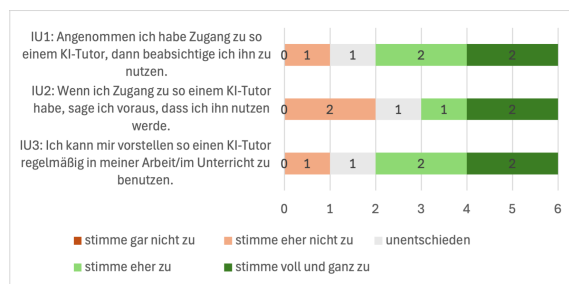


Abbildung 6.8: IU der Lehrenden

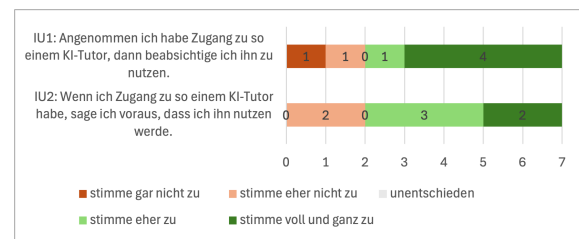


Abbildung 6.9: IU der Lernenden

6.4 Diskussion

Zunächst ist zu erwähnen, dass die Stichprobe mit 13 Probanden relativ klein ist, insbesondere wenn man die Gruppen der Lehrenden und Lernenden getrennt betrachtet, die mit jeweils sechs bzw. sieben Personen noch kleiner ausfallen. Zudem waren die Teilnehmenden überwiegend männlich, und die Lernenden entsprachen nicht dem Zielalter des Kursinhalts und des KI-Tutors. Diese Faktoren könnten die Ergebnisse möglicherweise beeinflusst haben. Trotz dieser Einschränkungen zeigen die Ergebnisse überwiegend eine positive Tendenz. Sowohl Lehrende als auch Lernende bewerteten die Anwendung des KI-Tutors in vielen Bereichen als hilfreich und nützlich. Im Folgenden werden die Ergebnisse im Detail ausgewertet und die spezifischen Aspekte der Nutzerakzeptanz genauer beleuchtet.

Perceived Usefulness

Wenn man die *Perceived Usefulness*, also die wahrgenommene Nützlichkeit, über beide Teilnehmergruppen hinweg betrachtet (Abb. 6.1), ist eine klare positive Tendenz erkennbar. Alle Fragen wurden mehrheitlich positiv bewertet. Besonders Aussage PU2: „Das Antwortverhalten des KI-Tutors war zuverlässig.“ wurde sehr positiv bewertet und ist mit einem Mittelwert von 4,08 die am besten bewertete Aussage. Dies lässt darauf schließen, dass das Verhalten des KI-Tutor größtenteils den Erwartungen der Nutzer während der Interaktion entsprochen hat. Allerdings gibt es bei dieser Aussage auch einen Ausreißer, bei dem die Bewertung auf „stimme gar nicht zu“ lautete. Dieser Proband hat jedoch in den qualitativen Fragen keine Rückmeldung gegeben, was diese negative Bewertung ausgelöst haben könnte.

Wenn man nur die Bewertung der Lehrenden (Abb. 6.2) betrachtet, zeigt sich, dass diese in fast allen Kategorien sehr positiv ausfällt. Besonders hervorzuheben ist auch hier Aussage PU2, die ausschließlich positive Bewertungen erhalten hat, was darauf hindeutet, dass die Lehrenden mit dem grundlegenden Frage-Antwort-Fluss des KI-Tutors keine größeren Probleme hatten. Auch die Aussage PU3 wurde von den Lehrenden sehr gut bewertet: Abgesehen von zwei Lehrenden, die „unentschieden“ waren, sahen alle den KI-Tutor tendenziell als Bereicherung für das Lernerlebnis an. Besonders interessant sind die drei zusätzlichen Fragen, die nur den Lehrenden gestellt wurden und konkret abfragen, wie nützlich sie den KI-Tutor für ihre eigene Arbeit bzw. den Unterricht erachten. Dabei fällt die Aussage PU8: „Der KI-Tutor kann eine nützliche Ergänzung zu meinen Lehrmaterialien sein“ mit

ausschließlich positiven Bewertungen besonders auf. Im Kontrast dazu steht Aussage PU6: „Ich denke, dass ein KI-Tutor wie dieser nützlich für meine Arbeit/für den Unterricht sein kann“, die zwar immer noch gut bewertet wurde, aber mit einem Mittelwert von 3,5 die niedrigste Bewertung innerhalb der *Perceived Usefulness*-Komponente der Lehrenden erhielt. Dies könnte darauf hindeuten, dass einige Lehrende zwar grundsätzlich Interesse daran haben, einen solchen KI-Tutor auszuprobieren, jedoch noch nicht vollständig vom konkreten Nutzen im Alltag überzeugt sind. Die größte Unentschlossenheit zeigt sich bei Aussage PU7: „Durch den Einsatz eines KI-Tutors im Unterricht wird meine Arbeit erleichtert“, die nur drei positive Bewertungen erhielt. Dies könnte auf die oben erwähnte Vermutung hindeuten, dass zwar Interesse an der Technologie besteht, der tatsächliche Nutzen zur Arbeitserleichterung jedoch noch teilweise skeptisch betrachtet wird.

Bei den Lernenden (Abb. 6.3) ist eine größere Meinungsstreuung erkennbar. Die Bewertungen weisen zwar immer noch eine positive Tendenz auf, jedoch sind auch vermehrt negative Bewertungen vorhanden, was darauf hindeutet, dass der Nutzen des KI-Tutors für diese Gruppe nicht so offensichtlich war wie für die Lehrenden. Am besten schnitten die Aussagen PU1 und PU2 ab, wobei dennoch zwei Personen angaben, dass der KI-Tutor sie eher nicht beim Verständnis der Themen unterstützen konnte. Das Antwortverhalten des KI-Tutors wurde ebenfalls überwiegend als zuverlässig bewertet, mit dem bereits erwähnten Ausreißer aus der Gesamtbetrachtung. Am schlechtesten wurden die Aussagen PU4: „Ich denke, dass der KI-Tutor Lernende sinnvoll bei Aufgaben unterstützen kann“ und PU5: „Ich finde den KI-Tutor nützlich“ bewertet, wobei der Mittelwert immer noch eine tendenziell positive Bewertung von 3,43 aufweist. Auffällig ist hierbei, dass es eine größere Spaltung in den Bewertungen gab: Die Zustimmung zu diesen Aussagen war entweder sehr stark, im Sinne von „stimme voll und ganz zu“, oder tendenziell negativ. Eine mögliche Erklärung für dieses Muster könnte darin liegen, dass die Befragung mit Universitätsstudierenden durchgeführt wurde und nicht mit der eigentlichen Zielgruppe des Kurses und KI-Tutors (7.-10. Klasse). Obwohl die Probanden aufgefordert wurden, den KI-Tutor aus der Perspektive dieser Zielgruppe zu bewerten, könnte der Tutor weniger nützlich erschienen sein, da die Aufgaben auf ein jüngeres Publikum und somit auf ein einfacheres Niveau ausgelegt waren. Es gab jedoch auch Antworten in den qualitativen Fragen, die Kontext zu den Bewertungen liefern. Ein Proband, der diese Aussagen negativ bewertete, nannte als Grund, dass der KI-Tutor sehr generische Antworten gegeben habe und nicht genug auf den Kursinhalt eingegangen ist. Eine weitere Person äußerte, dass sie es vermisst habe, dass der KI-Tutor keine konkreten Werte zu den Aufgaben herausgab. Dies könnte allerdings auch am Design des KI-Tutors liegen, da dieser bewusst so konzipiert wurde, keine direkten Lösungen zu liefern.

Perceived Ease of Use

Bei der *Perceived Ease of Use*-Komponente, also der wahrgenommenen Benutzerfreundlichkeit, schnitt der KI-Tutor insgesamt sehr positiv ab (siehe Abb. 6.4). Die meisten Aussagen erhielten überwiegend „stimme eher zu“ und „stimme voll und ganz zu“-Bewertungen. Besonders die Bedienung des KI-Tutors wurde als unkompliziert wahrgenommen, und auch die Nutzung insgesamt wurde durchgehend positiv bewertet, mit lediglich einem „unentschieden“ als Ausnahme. Dies deutet darauf hin, dass die Benutzeroberfläche des KI-Tutors

als Chat-Oberfläche gut funktioniert hat und keine größeren Verständnisprobleme bei der Bedienung aufgetreten sind.

Interessant ist jedoch, dass Aussage PEOU4: „Ich finde, der KI-Tutor macht ohne Probleme das, was ich möchte“ einen Ausreißer darstellt. Diese Aussage erhielt als einzige in diesem Segment negative Bewertungen. Im Vergleich dazu wurde Aussage PEOU1: „Die Interaktion mit dem KI-Tutor ist klar und verständlich“ durchweg positiv bewertet. Beide Aussagen zielen auf die Interaktion mit dem KI-Tutor ab, und während die grundsätzliche Funktionalität offenbar gut funktioniert und der KI-Tutor klare und verständliche Antworten liefert, hatten einige Probanden Schwierigkeiten damit, dass der KI-Tutor nicht immer genau das machte, was sie beabsichtigten. Oder Sie andere Erwartungen an diesen hatten.

Beim Vergleich der beiden Teilnehmergruppen, Lehrende und Lernende, decken sich die Meinungen weitgehend, und die Bewertungen fallen in beiden Gruppen überwiegend positiv aus. Allerdings gaben die Lernenden häufiger die Bewertung „stimme voll und ganz zu“ ab. Dies könnte darauf hindeuten, dass die Lernenden etwas besser mit der Benutzeroberfläche und der Interaktion mit dem KI-Tutor zurechtkamen. Ein möglicher Grund hierfür könnte der geringere Altersdurchschnitt der Lernenden sein, da sie von klein auf mit Chatoberflächen aufgewachsen sind und daher möglicherweise eine höhere Vertrautheit mit dieser Art der Interaktion besitzen. Ein weiterer Einflussfaktor könnte sein, dass die Lernenden in dieser Studie größtenteils von einer technischen Universität stammen und daher möglicherweise über eine größere digitale Kompetenz verfügen.

Intention to Use

Auch die *Intention to Use*-Komponente, also die Nutzungsabsicht, wurde überwiegend positiv bewertet (siehe Abb. 6.7). Allerdings gibt es hier auch einige negative Bewertungen, was auf eine gewisse Unsicherheit der Teilnehmer hinweisen könnte. Beim Vergleich der Gruppen zeigt sich, dass die Lernenden tendenziell eine stärkere positive Intention zur Nutzung des Tutors haben, was möglicherweise auf ihre größere Vertrautheit im Umgang mit digitalen Tools zurückzuführen ist. Besonders interessant ist die Aussage IU3, die nur den Lehrenden gestellt wurde und gezielt danach fragt, ob sie sich vorstellen können, einen solchen KI-Tutor regelmäßig im Unterricht einzusetzen. Diese Frage ist besonders relevant, da die Lehrenden letztlich darüber entscheiden, ob der KI-Tutor im Unterricht genutzt wird oder nicht. Die Aussage wurde ebenfalls überwiegend positiv bewertet, mit nur einer negativen Bewertung („stimme eher nicht zu“). Dies könnte ein Indikator dafür sein, dass Lehrende grundsätzlich offen für den Einsatz eines KI-Tutors im Unterricht sind und Potenzial in der Technologie sehen.

Auch die qualitativen Fragen bestätigen, dass die befragten Lehrenden Potenzial in KI-Tutoren wie diesem sehen. Fast alle Lehrenden beantworteten die Frage Q4: „Denken Sie, dass KI-Tutoren wie dieser den Unterricht in Zukunft sinnvoll ergänzen könnten? Warum oder warum nicht?“ positiv. Mehrfach wurde darauf hingewiesen, dass es zu viele Schüler pro Lehrkraft gibt und eine individuelle 1:1-Betreuung nicht möglich ist, weshalb gerade in diesem Aspekt der Vorteil von KI-Tutoren gesehen wird. Dies ist in den folgenden Zitaten ersichtlich:

- „Ja absolut. Die Nutzung des KI-Tutors kommt einer 1 zu 1 Betreuung gleich, was im normalen Unterricht durch eine Lehrkraft nicht möglich ist. Es entlastet die Lehrkraft und bereichert die Schüler/innen“
- „Ja, noch intelligentere Tutoren könnten die Lehrkräfte entlasten und bei Inklusion und Differenzierung (bis hin zur Individualisierung) helfen. Es sind zu viele Kinder pro Lehrkraft um den Kindern gerecht zu werden.“
- „Das kann sie, da die Lehrkraft und die Lehrmaterialien die Inhalte nicht immer für alle Lernenden verständlich erklären können. Die KI kann hierbei eine interessante Alternative sein, um die Lehrinhalte auf eine andere Weise und mit anderen Worten zu verdeutlichen und somit die Inhalte ggf. mehr Lernenden insgesamt verständlich zu machen.“

Eine der Antworten der Lehrenden war dem KI-Tutor im Kontext des Beispiellurses, mit dem die Studie durchgeführt wurde, eher kritisch gegenüber eingestellt. Der Proband betonte, dass in streng geführten Formaten mit engen Antworten der Nutzen des KI-Tutors eher gering sei, da zusätzliche Informationen auch über Suchmaschinen gefunden werden könnten. „Es fehlt beim sich einfachen beantworten lassen einer Frage irgendwie die kognitive Aktivierung, so dass man gezwungen ist sich selbst damit auseinander zu setzen.“ In offeneren Formaten könnte der Tutor seiner Meinung nach jedoch von größerem Nutzen sein, da das Stellen der richtigen Fragen bereits eine intensivere Auseinandersetzung mit dem Inhalt erfordert. In dieser Arbeit wurde bereits darauf geachtet, dass der KI-Tutor den Lernenden nicht einfach die Lösungen präsentiert, aber zukünftige Arbeiten könnten diesen Bereich noch genauer untersuchen.

Die qualitativen Fragen wurden auch gestellt, um herauszufinden, in welchen Bereichen der KI-Tutor besonders hilfreich war und wo noch Verbesserungspotenzial besteht. Der KI-Tutor wurde in verschiedenen Situationen als nützlich empfunden, insbesondere bei der Beseitigung von Unklarheiten im Kursmaterial. Sowohl Lehrende als auch Lernende lobten seine Fähigkeit, Wissenslücken zu schließen und Schwächen im Kurs zu kompensieren. Besonders häufig wurde der Nutzen bei der Erklärung unbekannter Begriffe, sowie der Bereitstellung von Zusatzinformationen hervorgehoben. Zudem wurde positiv erwähnt, dass der KI-Tutor dabei hilft, Denkfehler zu erkennen und zu korrigieren, sich nicht auf Abschweifungen einließ und stets im Kurskontext blieb. Auch die Fähigkeit, Fragen trotz grober Rechtschreibfehler oder fehlender Wörter richtig zu verstehen, wurde hervorgehoben, wodurch auch Lernende mit Lese-Rechtschreib-Schwäche inkludiert werden.

Die folgenden Punkte wurden als Bereiche genannt, in denen der KI-Tutor noch Verbesserungspotenzial aufweist. Am häufigsten wurde erwähnt, dass der KI-Tutor gerade bei Aufgaben oft nicht genügend ins Detail ging und eher „um den heißen Brei geredet“ hat. Dies war jedoch im Bereich der Aufgaben beabsichtigt, da der KI-Tutor die Lösung nicht direkt verraten, sondern zur Lösung hinführen sollte. In Zukunft könnte allerdings geprüft werden, ob dieser Ansatz weiter verbessert werden kann. Zudem wurde bemängelt, dass der KI-Tutor manchmal zu komplexe Antworten für jüngere Lernende gab und dass das Arbeiten mit zwei Fenstern (eines für den Kurs und eines für den KI-Tutor) für Kinder ungewohnt sei. Diese Einschränkung ergab sich aus der fehlenden Integration des Tutors in den

Kurs, da dies den Umfang dieser Arbeit überschritten hätte. Dies sollte aber in zukünftigen Versionen berücksichtigt werden. Darüber hinaus wurde erwähnt, dass der KI-Tutor bei Aufgaben mit Bildern oder grafischen Elementen weniger hilfreich war. Dies liegt zumindest teilweise an den aktuellen Einschränkungen der OpenAI-API, die nur mit textbasierten Informationen arbeitet. Diese Funktionalität könnte sich in Zukunft verbessern. Bis dahin könnte geprüft werden, ob grafische Elemente des Kurses noch besser textuell beschrieben werden können, um dem KI-Tutor mehr Kontext zu bieten. Ein Proband bemängelte zudem, dass das Nachfragen des KI-Tutors am Ende einer Antwort erzwungen wirkte und zu oft stattfand.

Mit Q2: „Haben Sie Vorschläge, wie der KI-Tutor weiter verbessert werden könnte?“ wurden die Teilnehmenden nach konkreten Verbesserungsvorschlägen gefragt. Dabei wurden folgende Punkte genannt: Zum einen wurde der Wunsch nach einer stärkeren Integration des KI-Tutors in den Kurs geäußert, eventuell mit der Weitergabe des aktuellen Fortschritts oder Kapitels an den KI-Tutor. Dies ist ein Vorschlag, der bereits in Betracht gezogen wurde, aber den Rahmen dieser Arbeit überschritten hätte. Für zukünftige Versionen ist dieser Aspekt jedoch von großer Bedeutung, und die aktuelle Implementierung wurde bewusst offen gestaltet, um eine Erweiterung in diese Richtung zu ermöglichen. Ein weiterer Wunsch war die Integration von Spracherkennung und Tonausgabe. Außerdem wurde vorgeschlagen, Beispielfragen anzuzeigen, um den Nutzern zu zeigen, wie der KI-Tutor am besten verwendet werden kann.

7 Limitationen und zukünftige Arbeiten

7.1 Limitationen

Eine wichtige Limitation, die bereits zu Beginn der Evaluation erwähnt wurde, ist die geringe Stichprobengröße mit 13 Probanden. Diese fällt noch kleiner aus, wenn man die Gruppen der Lehrenden und Lernenden getrennt betrachtet, mit jeweils sechs bzw. sieben Personen. Auch wenn die Studie interessante Einblicke zur Einstellung gegenüber KI-Tutoren gab, ist eine Verallgemeinerung der Ergebnisse auf Grund der geringer Stichprobengröße sehr beschränkt. Eine weitere Einschränkung betrifft die Teilnehmenden auf Seiten der Lernenden: Diese entsprachen nicht der eigentlichen Zielgruppe des Kurses, der für Schülerinnen und Schüler der 7.-10. Klasse konzipiert wurde, sondern es handelte sich um Universitätsstudierende. Dies könnte die Wahrnehmung der Nützlichkeit und Benutzerfreundlichkeit des KI-Tutors beeinflusst haben, da der Kurs für die Studierenden zum einen sehr einfach war und sie zum anderen bereits mehr Erfahrung im Umgang mit digitalen Programmen haben. Eine weitere Limitation besteht darin, dass der getestete KI-Tutor eine erste prototypische Implementierung darstellte, die in vielen Aspekten noch ausbaufähig ist. Ein Beispiel hierfür ist die fehlende direkte Integration des KI-Tutors in den Kurs. In der aktuellen Version musste der KI-Tutor in einem separaten Fenster geöffnet werden, was die Benutzerfreundlichkeit möglicherweise einschränkte. Zukünftige, verbesserte Versionen könnten daher potenziell bessere Ergebnisse liefern. Zudem stellt der verwendete Kurs, auf dem der KI-Tutor basierte, eine weitere Einschränkung dar. Der getestete Abschnitt stammte aus dem Einführungsbereich und enthielt weniger tiefgehende Fragen und Inhalte. Dies könnte sich ebenfalls auf die wahrgenommene Nützlichkeit ausgewirkt haben, da die behandelten Themen insgesamt recht einfach und stark strukturiert waren. Eine technische Limitation ergab sich aus den Einschränkungen der verwendeten Assistant API von OpenAI. Diese konnte, zum Zeitpunkt dieser Arbeit, nur textuelle Informationen in der Wissensdatenbank verarbeiten. Da jedoch viele Aufgaben der Kurse auf RockStartIT auch grafische Elemente beinhalten, stellte dies eine Einschränkung dar. Obwohl versucht wurde, diese grafischen Aufgaben so gut wie möglich textuell zu beschreiben, bleibt dies eine Limitation, die möglicherweise die wahrgenommene Nützlichkeit des KI-Tutors beeinträchtigt hat, was auch in Kommentaren in den qualitativen Fragen angedeutet wurde.

7.2 Zukünftige Arbeiten

Zukünftige Arbeiten könnten sich mit der Weiterentwicklung und Verbesserung des KI-Tutors beschäftigen. Im Rahmen dieser Arbeit wurde der Tutor zunächst prototypisch implementiert und weist noch Verbesserungspotenzial auf. Die derzeitige Implementierung ist jedoch bewusst einfach und offen gestaltet, um Erweiterungen zu erleichtern. Eine mögliche Verbesserung wäre eine tiefere Integration des KI-Tutors in den Kurs. In der aktuellen Version ist der Tutor noch als externe Webseite konzipiert. Eine direkte Einbindung auf der Kursseite, möglicherweise mit der Übertragung des aktuellen Kurses und Kapitels an den KI-Tutor, wäre eine sinnvolle Erweiterung. Auch das Anzeigen von Beispielfragen, die der Nutzer stellen könnte, wäre eine hilfreiche Ergänzung. Zudem äußerten die Teilnehmenden den Wunsch nach einer Sprach-Ein- und Ausgabe, die ebenfalls in zukünftigen Versionen umgesetzt werden könnte.

Bezüglich der Akzeptanz von KI-Tutoren gibt es weitere interessante Forschungsfelder, die genauer untersucht werden könnten. Zum einen war die Teilnehmerzahl dieser Studie begrenzt, und es wäre aufschlussreich, diese in einem größeren Rahmen zu wiederholen, um belastbarere Erkenntnisse zu gewinnen. Zudem wäre eine Studie von Interesse, die die Akzeptanz und Benutzerfreundlichkeit bei Lernenden der tatsächlichen Zielgruppe des Kurses, also Schülerinnen und Schülern der 7.-10. Klasse, evaluiert, da in dieser Arbeit aus Zeitgründen nur Studierende befragt werden konnten. Darüber hinaus könnte eine Untersuchung basierend auf einem anspruchsvolleren Kurs wertvolle Einblicke bieten, um zu ermitteln, welche Auswirkungen dies auf die wahrgenommene Nützlichkeit des KI-Tutors hat.

8 Fazit

Die individuelle Förderung von Schülern stellt in modernen Bildungssystemen aufgrund des oft hohen Lehrenden-Schüler-Verhältnisses eine Herausforderung dar. Besonders nach Unterrichtschluss fehlt häufig die notwendige Unterstützung, was den Lernerfolg beeinträchtigen kann. KI-Tutoren bieten hier eine vielversprechende Lösung, da sie personalisierte Lernunterstützung rund um die Uhr ermöglichen. Um das Potenzial von KI-Tutoren besser zu verstehen, wurde in dieser Arbeit die Akzeptanz von Lehrenden und Lernenden gegenüber solchen KI-Tutoren untersucht.

Zu diesem Zweck wurde zunächst prototypisch ein KI-Tutor für einen Kurs von RockStartIT entwickelt. Dies umfasste sowohl die Entwicklung einer Web-App als auch die Erstellung des KI-Tutors. Für den KI-Tutor mussten Anforderungen definiert und Anweisungen für das KI-Modell durch Ausprobieren und Dokumentieren entwickelt werden, um diese Anforderungen möglichst gut zu erfüllen. Zusätzlich musste der KI-Tutor mit den Kursinhalten und den Aufgaben vertraut gemacht werden, damit er effektiv unterstützen kann. Hierfür wurde für die Wissensdatenbank des KI-Tutors ein strukturierter Aufbau des Kursinhalts entworfen und die Inhalte eines Kurses in dieses Format übertragen.

Anhand dieses KI-Tutors und des Kurses wurde die Akzeptanz von Lehrenden und Lernenden gegenüber KI-Tutoren mithilfe des Technologie-Akzeptanz-Modells (TAM) untersucht. Die Teilnehmenden der Studie testeten den KI-Tutor und füllten anschließend einen Fragebogen aus, der sowohl quantitative Fragen basierend auf dem originalen TAM-Fragebogen als auch qualitative Fragen enthielt. Die qualitativen Fragen dienten insbesondere dazu, die Ergebnisse genauer zu interpretieren und sowohl die Stärken als auch das Verbesserungspotenzial der prototypischen Implementierung des KI-Tutors zu erkennen.

Die Ergebnisse der Akzeptanzuntersuchung fielen überwiegend positiv aus. Besonders die *Perceived Usefulness*-Komponente wurde von den Lehrenden sehr gut bewertet. Jeder der befragten Lehrenden stimmte der Aussage PU8: „Der KI-Tutor kann eine nützliche Ergänzung zu meinen Lehrmaterialien sein.“ zu. Die Bewertungen der Lernenden waren zwar ebenfalls positiv, jedoch gab es dort häufiger kritische Stimmen, die den KI-Tutor in seiner aktuellen Implementierung und für den Beispielkurs als weniger nützlich empfanden. In den qualitativen Fragen wurde als möglicher Grund angeführt, dass der KI-Tutor oft zu generische Antworten gab und bei den Aufgaben keine konkreten Werte lieferte. Bei den Aufgaben war dies zumindest absichtlich so konzipiert, da der KI-Tutor keine Lösungen direkt vorgeben sollte. Zukünftige Arbeiten könnten untersuchen, ob dieser Ansatz weiter verbessert werden kann.

Die Ergebnisse der *Perceived Ease of Use*-Komponente fielen in beiden Gruppen sehr positiv aus. Dies deutet darauf hin, dass die Umsetzung des KI-Tutors und die Benutzeroberfläche

als benutzerfreundlich wahrgenommen wurden. Allerdings gab es bei der Frage PEOU4: „Ich finde, der KI-Tutor macht ohne Probleme das, was ich möchte.“ einige negative Rückmeldungen, was darauf hindeutet, dass die Antworten des KI-Tutors nicht immer den Erwartungen der Nutzer entsprachen. Dies ist ein Aspekt, der in zukünftigen Arbeiten genauer untersucht werden sollte.

Die *Intention to Use*-Komponente fiel ebenfalls positiv aus, jedoch etwas differenzierter. Dies könnte darauf hindeuten, dass grundsätzliches Interesse an KI-Tutoren besteht, der konkrete Nutzen jedoch noch nicht für alle direkt erkennbar ist und es Verbesserungspotenzial gibt. Besonders interessant ist in dieser Kategorie die Aussage IU3: „Ich kann mir vorstellen, so einen KI-Tutor regelmäßig in meiner Arbeit/im Unterricht zu benutzen.“, die nur den Lehrenden gestellt wurde und von der Mehrheit positiv bewertet wurde. Dies deutet darauf hin, dass Lehrende potenziell offen für den Einsatz eines KI-Tutors im Unterricht sind und das Potenzial der Technologie erkennen. Auch die Antworten auf die qualitative Frage Q4: „Denken Sie, dass KI-Tutoren wie dieser den Unterricht in Zukunft sinnvoll ergänzen könnten? Warum oder warum nicht?“ waren fast ausschließlich positiv. Hier wurde mehrfach das Problem des hohen Schüler-Lehrkraft-Verhältnisses hervorgehoben und betont, dass KI-Tutoren in diesem Bereich eine Entlastung für Lehrkräfte und eine Bereicherung für die Schülerinnen und Schüler sein könnten.

Zusammenfassend lässt sich sagen, dass die Erkenntnisse dieser Arbeit das Potenzial von KI-Tutoren, wie dem hier prototypisch implementierten, bestätigen. Die TAM-Untersuchung zeigt, dass bereits in der aktuellen Implementierung von einer Mehrheit der Teilnehmenden ein Nutzen erkannt wurde und eine gute Grundlage für zukünftige Arbeiten bietet. Allerdings gibt es noch Verbesserungspotenzial, wie beispielsweise eine tiefere Integration in den Kurs, was in der Sektion 7.2 für zukünftige Arbeiten näher beleuchtet wird. Auch war die Stichprobe in dieser Arbeit mit 13 Probanden relativ klein, weshalb die Ergebnisse in zukünftigen Studien mit größeren Stichproben validiert werden sollten.

Datenverfügbarkeit

Alle Forschungsdaten (u.a. Umfrageantworten und Ressourcen), sind im SDQ GitLab-Abschlussarbeiten-Repository hinterlegt (GitLab-Repository).

Der Quellcode der Implementierung ist auf GitHub unter folgenden Link gespeichert:
github.com/Rockstartit/ki-tutor

Literatur

- [1] Ahlam Mohammed Al-Abdullatif. „Modeling Students’ Perceptions of Chatbots in Learning: Integrating Technology Acceptance with the Value-Based Adoption Model“. In: *Education Sciences* 13.11 (Nov. 2023). Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, S. 1151. ISSN: 2227-7102. DOI: 10.3390/educsci13111151. URL: <https://www.mdpi.com/2227-7102/13/11/1151> (besucht am 01.06.2024).
- [2] Tarek Ait Baha u. a. „The impact of educational chatbot on student learning experience“. In: *Education and Information Technologies* 29.8 (1. Juni 2024), S. 10153–10176. ISSN: 1573-7608. DOI: 10.1007/s10639-023-12166-w. URL: <https://doi.org/10.1007/s10639-023-12166-w> (besucht am 12.09.2024).
- [3] *Assistants Playground - OpenAI API*. URL: <https://platform.openai.com/playground/assistants> (besucht am 25.09.2024).
- [4] Jose Belda-Medina und José Ramón Calvo-Ferrer. „Using Chatbots as AI Conversational Partners in Language Learning“. In: *Applied Sciences* 12.17 (Jan. 2022). Number: 17 Publisher: Multidisciplinary Digital Publishing Institute, S. 8427. ISSN: 2076-3417. DOI: 10.3390/app12178427. URL: <https://www.mdpi.com/2076-3417/12/17/8427> (besucht am 01.06.2024).
- [5] Ghazala Bilquise, Samar Ibrahim und Sa’Ed M. Salhieh. „Investigating student acceptance of an academic advising chatbot in higher education institutions“. In: *Education and Information Technologies* 29.5 (1. Apr. 2024), S. 6357–6382. ISSN: 1573-7608. DOI: 10.1007/s10639-023-12076-x. URL: <https://doi.org/10.1007/s10639-023-12076-x> (besucht am 01.06.2024).
- [6] *chatscope/chat-ui-kit-react*. original-date: 2020-08-26T17:56:46Z. 29. Aug. 2024. URL: <https://github.com/chatscope/chat-ui-kit-react> (besucht am 05.09.2024).
- [7] Ken Nee Chee u. a. „Exploring the Trend and Potential Distribution of Chatbot in Education: A Systematic Review“. In: *International Journal of Information and Education Technology* 13 (1. Jan. 2023), S. 516–525. DOI: 10.18178/ijiet.2023.13.3.1834.
- [8] Banghao Chen u. a. *Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review*. 5. Sep. 2024. DOI: 10.48550/arXiv.2310.14735. arXiv: 2310.14735[cs]. URL: <http://arxiv.org/abs/2310.14735> (besucht am 24.09.2024).
- [9] Yu Chen u. a. „Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success“. In: *Information Systems Frontiers* 25.1 (1. Feb. 2023), S. 161–182. ISSN: 1572-9419. DOI: 10.1007/s10796-022-10291-4. URL: <https://doi.org/10.1007/s10796-022-10291-4> (besucht am 10.04.2024).

- [10] Raquel Chocarro, Mónica Cortiñas und Gustavo Marcos-Matás. „Teachers’ attitudes towards chatbots in education: a technology acceptance model approach considering the effect of social language, bot proactiveness, and users’ characteristics“. In: *Educational Studies* 49.2 (4. März 2023), S. 295–313. ISSN: 0305-5698. DOI: 10.1080/03055698.2020.1850426. URL: <https://doi.org/10.1080/03055698.2020.1850426> (besucht am 21.05.2024).
- [11] Fred D. Davis. „Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology“. In: *MIS Quarterly* 13.3 (1989). Publisher: Management Information Systems Research Center, University of Minnesota, S. 319–340. ISSN: 0276-7783. DOI: 10.2307/249008. URL: <https://www.jstor.org/stable/249008> (besucht am 21.05.2024).
- [12] Xinjie Deng und Zhonggen Yu. „A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education“. In: *Sustainability* 15.4 (Jan. 2023). Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, S. 2940. ISSN: 2071-1050. DOI: 10.3390/su15042940. URL: <https://www.mdpi.com/2071-1050/15/4/2940> (besucht am 10.04.2024).
- [13] Arzu Deveci Topal, Canan Dilek Eren und Aynur Kolburan Geçer. „Chatbot application in a 5th grade science course“. In: *Education and Information Technologies* 26.5 (2021), S. 6241–6265. ISSN: 1360-2357. DOI: 10.1007/s10639-021-10627-8.
- [14] Harry Barton Essel u. a. „The impact of a virtual teaching assistant (chatbot) on students’ learning in Ghanaian higher education“. In: *International Journal of Educational Technology in Higher Education* 19.1 (15. Nov. 2022), S. 57. ISSN: 2365-9440. DOI: 10.1186/s41239-022-00362-6. URL: <https://doi.org/10.1186/s41239-022-00362-6> (besucht am 03.06.2024).
- [15] *Fine-tuning - OpenAI API*. URL: <https://platform.openai.com/docs/guides/fine-tuning/when-to-use-fine-tuning> (besucht am 03.09.2024).
- [16] Sánchez-Vera Fulgencio. „Developing Effective Educational Chatbots with GPT: Insights from a Pilot Study in a University Subject“. In: *Trends in Higher Education* 3.1 (März 2024). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, S. 155–168. ISSN: 2813-4346. DOI: 10.3390/higheredu3010009. URL: <https://www.mdpi.com/2813-4346/3/1/9> (besucht am 12.04.2024).
- [17] Aishwarya Gupta und Divya Hathwar. „Introduction to AI Chatbots“. In: *International Journal of Engineering Research and* V9.7 (11. Juli 2020), IJERTV9IS070143. ISSN: 2278-0181. DOI: 10.17577/IJERTV9IS070143. URL: <https://www.ijert.org/introduction-to-ai-chatbots> (besucht am 10.06.2024).
- [18] Julia Hirschberg und Christopher D. Manning. „Advances in natural language processing“. In: *Science* 349.6245 (17. Juli 2015). Publisher: American Association for the Advancement of Science, S. 261–266. DOI: 10.1126/science.aaa8685. URL: <https://www.science.org/doi/full/10.1126/science.aaa8685> (besucht am 08.05.2024).

-
- [19] Enkelejda Kasneci u. a. „ChatGPT for good? On opportunities and challenges of large language models for education“. In: *Learning and Individual Differences* 103 (1. Apr. 2023), S. 102274. ISSN: 1041-6080. DOI: 10.1016/j.lindif.2023.102274. URL: <https://www.sciencedirect.com/science/article/pii/S1041608023000195> (besucht am 08.05.2024).
- [20] Reena Malik u. a. „Adoption of Chatbots for Learning among University Students: Role of Perceived Convenience and Enhanced Performance“. In: *International Journal of Emerging Technologies in Learning (iJET)* 16.18 (20. Sep. 2021). Publisher: International Journal of Emerging Technology in Learning, S. 200–212. ISSN: 1863-0383. URL: <https://www.learntechlib.org/p/220124/> (besucht am 01.06.2024).
- [21] David Mhlanga. „Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning“. In: *FinTech and Artificial Intelligence for Sustainable Development: The Role of Smart Technologies in Achieving Development Goals*. Hrsg. von David Mhlanga. Cham: Springer Nature Switzerland, 2023, S. 387–409. ISBN: 978-3-031-37776-1. DOI: 10.1007/978-3-031-37776-1_17. URL: https://doi.org/10.1007/978-3-031-37776-1_17 (besucht am 10.09.2024).
- [22] Mai Neo. „THE MERLIN PROJECT: MALAYSIAN STUDENTS’ ACCEPTANCE OF AN AI CHATBOT IN THEIR LEARNING PROCESS“. In: *Turkish Online Journal of Distance Education* 23.3 (1. Juli 2022). Number: 3 Publisher: Anadolu University, S. 31–48. ISSN: 1302-6488. DOI: 10.17718/tojde.1137122. URL: <https://dergipark.org.tr/en/pub/tojde/issue/70682/1137122> (besucht am 27.08.2024).
- [23] *openai/openai-node*. original-date: 2021-12-14T22:32:58Z. 5. Sep. 2024. URL: <https://github.com/openai/openai-node> (besucht am 05.09.2024).
- [24] Dimitrios Ramandanis und Stelios Xinogalos. „Designing a Chatbot for Contemporary Education: A Systematic Literature Review“. In: *Information* 14.9 (Sep. 2023). Number: 9 Publisher: Multidisciplinary Digital Publishing Institute, S. 503. ISSN: 2078-2489. DOI: 10.3390/info14090503. URL: <https://www.mdpi.com/2078-2489/14/9/503> (besucht am 10.04.2024).
- [25] *remarkjs/react-markdown*. original-date: 2015-05-18T23:59:17Z. 9. Sep. 2024. URL: <https://github.com/remarkjs/react-markdown> (besucht am 09.09.2024).
- [26] Bayan Shawar und Eric Atwell. „Chatbots: Are they Really Useful?“ In: *LDV Forum* 22 (1. Juli 2007), S. 29–49. DOI: 10.21248/jlcl.22.2007.88.
- [27] Joseph Weizenbaum. „ELIZA—a computer program for the study of natural language communication between man and machine“. In: *Communications of the ACM* 9.1 (1. Jan. 1966), S. 36–45. ISSN: 0001-0782. DOI: 10.1145/365153.365168. URL: <https://dl.acm.org/doi/10.1145/365153.365168> (besucht am 08.06.2024).