

Kompagnon für das Verfassen von Hausarbeiten

Nghia Duong-Trung ¹, Xia Wang ¹, Rahul Rajkumar Bhoyar ¹, Angelin Mary Jose ¹,
Silke Elisabeth Wrede ², Lars van Rijn ², Theresa Panse ², Claudia de Witt ², and
Niels Pinkwart ¹

Abstract: Der Beitrag präsentiert eine Demonstration des „Term Paper Writing Companion“, einem KI-gestützten akademischen Unterstützungssystem, das entwickelt wurde, um das Lernen von Studierenden bei forschungsorientierten Schreibaufgaben zu unterstützen. Ursprünglich für einen B.A.-Studiengang in den Kultur- und Sozialwissenschaften an der FernUniversität in Hagen entwickelt, integriert es Large Language Models (LLMs), expertenbasierte Feedback-Mechanismen und Literaturempfehlungs-Maschinen (recommender engines). Das System bietet personalisiertes, formatives Feedback zu möglichen Themen und Fragestellungen von Hausarbeiten und adressiert gleichzeitig Probleme wie kognitive Überlastung der Studierenden und zeitliche Ressourcenengpässe bei den Lehrkräften. Diese Demo zeigt die Architektur, den Arbeitsablauf und die Benutzererfahrung und veranschaulicht das Potenzial des Systems, das forschungsbasierte Lernen und die akademische Autonomie in verschiedenen Hochschulumgebungen zu verbessern.

Keywords: Generative KI, Recommendersystem, Akademisches Schreiben.

1 Einführung

Die Auswahl eines klar definierten Forschungsthemas ist eine entscheidende Herausforderung im akademischen Studium, insbesondere in Kursen, die unabhängiges Forschen und kritisches Denken in den Vordergrund stellen. Das Tool „Term Paper Writing Companion“ wurde entwickelt, um Studierende während dieses Prozesses zu unterstützen, indem es ihnen mithilfe generativer KI intelligente Hilfestellungen bietet. Das Tool hilft Lernenden dabei, Forschungsfragen zu verfeinern und geeignete wissenschaftliche Literatur zu identifizieren, indem es strukturiertes, auf Eingabeaufforderungen basierendes Feedback, Sentimentanalyse und Recommender-Techniken kombiniert. Im Gegensatz zu rein generativen KI-Ansätzen, die zu einer übermäßigen Abhängigkeit von automatisierten Ergebnissen führen können, wurde das Tool ausdrücklich entwickelt, um die Autonomie der Studierenden zu fördern,

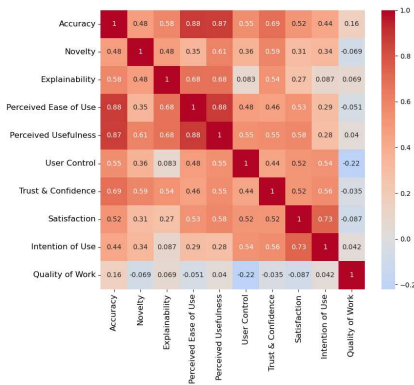
- 1 German Research Center for Artificial Intelligence (DFKI), Alt-Moabit 91C, 10559 Berlin, Germany, nghia_trung.duong@dfki.de,  <https://orcid.org/0000-0002-7402-4166>; xia.wang@dfki.de,  <https://orcid.org/0000-0003-0849-5645>; rahul_rajkumar.bhojar@dfki.de,  <https://orcid.org/0009-0005-7249-7305>; angelin_mary.jose@dfki.de,  <https://orcid.org/0009-0005-5026-8990>; niels.pinkwart@dfki.de,  <https://orcid.org/0009-0005-5026-0000>
- 2 FernUniversität in Hagen, Universitätsstraße 11, 58097 Hagen, Germany, silke.wrede@fernuni-hagen.de,  <https://orcid.org/0000-0002-4957-9792>; lars.vanrijn@fernuni-hagen.de,  <https://orcid.org/0000-0002-8381-2666>; theresa.panse@fernuni-hagen.de,  <https://orcid.org/0009-0006-5734-3161>; claudia.dewitt@fernuni-hagen.de,  <https://orcid.org/0000-0001-6478-8392>

indem es personalisierte, forschungsorientierte Unterstützung bietet. Seine Empfehlungen dienen als kognitive Unterstützung [KLC22; Mu23] und helfen den Studierenden, ihre Ideen zu entwickeln, mit Fortführung der eigenen akademischen Argumentation und Entscheidungsfindung. Um die Wirksamkeit des Tools zu bewerten, haben wir eine Pilotstudie im Rahmen des B.A.-Studiengangs durchgeführt und dabei das CRS-Que-Framework [Ji24] angewendet, um die Benutzerfreundlichkeit des Systems und seine pädagogische Wirkung zu bewerten.

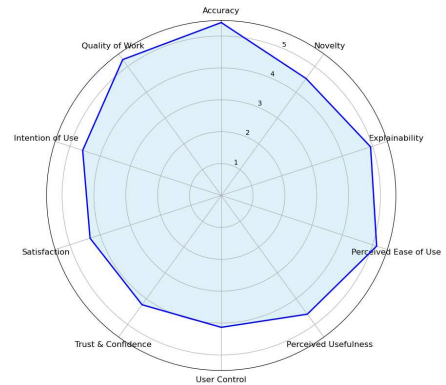
2 Systemübersicht und Architektur

Das System besteht aus drei Hauptkomponenten: einer Frontend-Webschnittstelle, einer Backend-Orchestrierungsschicht und einem von der Universität gehosteten LLM- Inferenzserver Flexi [Ze24]. Das Frontend bietet den Studierenden eine intuitive und personalisierte Oberfläche für die Einreichung von Forschungsvorschlägen, einschließlich erster Themenbeschreibungen und damit verbundener Forschungsfragen. Die Studierenden können zwischen mehreren LLMs wählen – wie Mixtral-8x7B, LLaMA-3-8B und anderen –, Feedback-Ergebnisse vergleichen und verschiedene Perspektiven erkunden. Sie können auch Feedback-Verläufe einsehen, Entwürfe überarbeiten und erneut einreichen sowie Empfehlungen steuern. Diese Oberfläche wurde entwickelt, um Transparenz zu fördern, iterative Überarbeitungen zu erleichtern und eine strukturierte Umgebung für akademische Erfahrungen zu bieten. Im Backend verarbeitet eine Reihe von Analysemodulen die Eingaben der Studierenden. Ein Modul zur Sentimentanalyse schätzt das Vertrauen der Studierenden in das eigene Vorhaben ein (confidence), woraufhin entschieden wird, ob eine wissensbasierte oder eine expertenbasierte Empfehlungsmaschine aktiviert wird. Bei Studierendeneinreichungen mit der Bewertung „hohes Vertrauensniveau“ empfiehlt das System akademische Literatur, die auf das vorgeschlagene Thema der Studierenden abgestimmt ist. Bei Einreichungen mit der Bewertung „geringes Vertrauensniveau“ bietet das Tool grundlegende Anleitungen, um den Studierenden zu helfen, ihren Forschungsschwerpunkt zu klären oder neu zu strukturieren. Diese Wege stellen sicher, dass die Studierenden eine Unterstützung erhalten, die ihrem aktuellen Vorbereitungs- und Vertrauensniveau entspricht. Die Eingaben werden anhand einer Reihe strukturierter promptbasierter Bewertungsmodule validiert, die die Klarheit, Durchführbarkeit und Übereinstimmung von Themen und Fragen bewerten. Diese Validierung umfasst individuelle Überprüfungen für jede Forschungsfrage und die Kohärenz zwischen Themen und Fragen. Der Prozess der Feedback-Generierung wird durch Multi-Shot-Prompting verbessert, bei dem das System historische Betreuungsdaten aus früheren Kohorten einbezieht, um nuancierte Antworten von Lehrkräften nachzuahmen. Diese Multi-Turn-Beispiele wurden aus echten Moodle-basierten Dialogen zwischen Studierenden und Lehrkräften abgeleitet, sodass das System Feedback generieren kann, das pädagogische Tiefe und akademische Erwartungen widerspiegelt. Alle LLM-Anfragen werden über Flexi geleitet, eine von der Universität kontrollierte Inferenzinfrastruktur, die die sichere lokale Ausführung vieler Open-Source-Modelle ermöglicht. Flexi macht die Abhängigkeit von APIs von Drittanbietern überflüssig und gewährleistet die Einhaltung der Datenschutzrichtlinien

der Einrichtung und der Europäischen Datenschutz-Grundverordnung (DSGVO). Parallel dazu ruft ein semantisches Recommender-Modul relevante akademische Literatur ab, indem es die Beiträge der Studierenden mit mehrsprachigen Satzkodierern einbettet und sie mit dem kuratierten Repository der Universität für akademische Texte abgleicht. Die Systemarchitektur wurde unter Berücksichtigung von Modularität und zukünftiger Integration entwickelt. Sie unterstützt Verbindungen zu Lernmanagementsystemen, Feedback-Dashboards und externen Bewertungsmodulen. Die Trennung zwischen Interaktionslogik und LLM-Verarbeitung verbessert nicht nur die Systemleistung, sondern stellt auch sicher, dass die studentischen Komponenten reaktionsschnell bleiben und gleichzeitig die Transparenz und das Vertrauen in die von KI generierten Vorschläge gewahrt bleiben. Erste Ergebnisse aus der Einführung im Sommersemester 2025 <https://tprs.fernuni-hagen.de/login>, inwieweit das System zur Optimierung der Betreuung beiträgt, die Feedback-Verzögerung reduziert und das engagement der Studierenden bei akademischen Schreibaufgaben fördert, liegen vor.



(a) Korrelations-Heatmap der Bewertungsdimensionen.



(b) Radardiagramm der Durchschnittswerte der Kategorien.

Fig. 1: Ergebnisse der Evaluation.

3 Bereitstellung und Evaluierung

Um die Wirksamkeit des Term Paper Writing Companion zu bewerten, wurde im Wintersemester 2024 eine Pilotstudie mit 15 Studierenden eines B.A.-Studiengangs durchgeführt. Die Bewertung erfolgte anhand des CRS-Que-Rahmenwerks, das 10 zentrale Bewertungsdimensionen mit insgesamt 32 Unterdimensionen umfasst. Die Studierenden bewerteten jeden Punkt auf einer 7-Punkte-Likert-Skala. Das Radardiagramm (siehe Abbildung 1b) vermittelt ein klares visuelles Profil der wahrgenommenen Stärken des Systems. Dimensionen wie Genauigkeit (5,42), Arbeitsqualität (5,27) und wahrgenommene Benutzerfreundlichkeit (5,13) wurden hoch bewertet, was darauf hindeutet, dass die Studierenden das System als effektiv und benutzerfreundlich empfanden. Im Gegensatz dazu erhielten Benutzerkontrolle (4,13) und Vertrauen & Zuversicht (4,23) vergleichsweise niedrigere Bewertungen,

was auf Verbesserungsmöglichkeiten hinsichtlich Transparenz und Interaktionsflexibilität hindeutet. Ergänzend dazu zeigte die in Abbildung 1a dargestellte Korrelations-Heatmap aussagekräftige Wechselbeziehungen zwischen den Bewertungsmetriken. So war beispielsweise die Erklärbarkeit stark mit Vertrauen und wahrgenommener Nützlichkeit verbunden, was die Bedeutung verständlicher KI-generierter Rückmeldungen für die Förderung von Vertrauen und nachhaltiger Nutzung unterstreicht. In ähnlicher Weise zeigte die Benutzerfreundlichkeit positive Korrelationen mit Zufriedenheit und Nutzungsabsicht, was die Benutzerfreundlichkeit als zentralen Treiber für das Engagement.

4 Abschluss

Diese Demonstration stellt das Term Paper Writing Companion System vor, ein modulares KI-gestütztes Unterstützungssystem, das Studierenden durch den kognitiv anspruchsvollen Prozess der Formulierung und Verfeinerung akademischer Themen begleitet. Durch die Integration einer strukturierten, promptbasierten Bewertung, hybrider Empfehlungsmechanismen und einer datenschutzkonformen LLM-Inferenz über den FLEXI-Server schließt das System die Lücke zwischen autonomem und angeleitetem Lernen. Die Pilotbewertung unter Verwendung des CRS-Que-Frameworks zeigt vielversprechende pädagogische Ergebnisse, insbesondere in Bezug auf die wahrgenommene Genauigkeit, Benutzerfreundlichkeit und Qualität der studentischen Arbeit. Gleichzeitig zeigen Bereiche wie Benutzerkontrolle und Vertrauen wichtige Richtungen für die weitere Entwicklung auf. Unsere Ergebnisse deuten darauf hin, dass KI-gesteuerte Scaffolding-Tools wie unser vorgeschlagenes System das forschungsbasierte Lernen im Fernunterricht sinnvoll unterstützen können, indem sie zeitnahes formatives Feedback liefern und gleichzeitig die Belastung der Lehrenden reduzieren.

Literaturverzeichnis

- [Ji24] Jin, Y.; Chen, L.; Cai, W.; Zhao, X.: CRS-Que: A user-centric evaluation framework for conversational recommender systems. *ACM Transactions on Recommender Systems* 2 (1), S. 1–34, 2024.
- [KLC22] Kim, J.; Lee, H.; Cho, Y. H.: Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies* 27 (5), S. 6069–6104, 2022.
- [Mu23] Munshi, A.; Biswas, G.; Baker, R.; Ocumpaugh, J.; Hutt, S.; Paquette, L.: Analysing adaptive scaffolds that help students develop self-regulated learning behaviours. *Journal of Computer Assisted Learning* 39 (2), S. 351–368, 2023.
- [Ze24] Zesch, T.; Hanses, M.; Seidel, N.; Aggarwal, P.; Veiel, D.; De Witt, C.: Flexible LLM Experimental Infrastructure (FLEXI)–Enabling Experimentation and Innovation in Higher Education Through Access to Open LLMs. In: *2024 21st International Conference on Information Technology Based Higher Education and Training (ITHET)*. IEEE, S. 1–8, 2024.