TPRS: AI-Assisted Research Topic Refinement for Distance Learners

Nghia Duong-Trung[1,2](✉), Xia Wang[1], Rahul Rajkumar Bhoyar[1], Angelin Mary Jose[1], Silke Elisabeth Wrede[3], Lars van Rijn[3], Theresa Panse[3], Claudia de Witt[3], and Niels Pinkwart[1,4]

[1] German Research Center for Artificial Intelligence (DFKI), Alt-Moabit 91 C, 10559 Berlin, Germany
{nghia_trung.duong,xia.wang,rahul_rajkumar.bhoyar,angelin_mary.jose, niels.pinkwart}@dfki.de

[2] IU International University of Applied Sciences, Frankfurter Allee 73A, 10247 Berlin, Germany

[3] FernUniversität in Hagen, Universitätsstraße 27, 58097 Hagen, Germany
{silke.wrede,lars.vanrijn,theresa.panse,claudia.dewitt}@fernuni-hagen.de

[4] Humboldt-Universität zu Berlin, Unter den Linden 6, 10117 Berlin, Germany

Abstract. We present the Term Paper Recommendation System (TPRS), a hybrid AI-powered platform designed to scaffold the development of academic term papers in distance education. Tailored for students in a Bachelor of Arts program in Culture and Social Sciences, TPRS dynamically integrates large language models (LLMs), expert- and knowledge-based recommendation engines, and sentiment-driven routing to provide personalized formative feedback. A multi-shot prompting technique simulates high-fidelity tutoring interactions, trained on real supervision logs. Our system prioritizes transparency, student autonomy, and pedagogical alignment by combining structured validation, explainable recommendations, and relevance-based literature suggestions. A pilot deployment involving 18 students showed statistically significant improvements in submission quality (Hedge's $g = 0.44$, $p < .05$) and positive user reception across accuracy, usability, and trust dimensions, evaluated using the CRS-Que framework. This work contributes a modular, pedagogically informed approach to AI-supported academic writing, offering a promising direction for scalable, inquiry-driven support systems in higher education.

Keywords: Intelligent Tutoring · Generative AI · Distance Learning · Topic Refinement · CRS-Que

1 Introduction

Artificial Intelligence (AI) technologies have become integral to advancing instructional efficiency and addressing pedagogical challenges in higher education [1]. Applications of AI span a range of domains, including personalized

A. I. Cristea et al. (Eds.): AIED 2025, CCIS 2591, pp. 69–76, 2025.
https://doi.org/10.1007/978-3-031-99264-3_9

learning [6], intelligent tutoring [2], and automated feedback [13]. Pérez-Ortiz *et al.* conceptualize AI-based personal learning companions as tools capable of delivering adaptive, personalized support for learners [12]. A growing body of systematic reviews highlights the increasing adoption of AI technologies in higher education, particularly in promoting self-regulated learning and academic assistance [3,11]. The UNESCO AI Competency Frameworks for Teachers and Students emphasize the urgency of cultivating AI literacy, ethical awareness, and technical skills among educators and learners. The teacher framework focuses on AI pedagogy, ethics, and continuous professional development to ensure meaningful classroom integration [8], while the student framework centers on ethical responsibility, AI literacy, and foundational system design [9]. These initiatives collectively drive research into AI-driven educational systems, advocating for personalized learning, ethical AI deployment, and inclusive, sustainable educational transformation.

This work introduces the development prototype of TPRS, an AI-assisted tool designed to support German-speaking students in a Bachelor of Arts distance learning program in Culture and Social Sciences. The system aims to provide inspiring, personalized feedback on initial ideas for research topics and questions, which traditionally involve intensive one-on-one interactions between students and instructors. The existing workflow requires students to independently propose research topics and questions, which they submit via online platforms, such as Moodle, for iterative feedback from instructors. This process is labor-intensive and prone to delays. From a student's perspective, distilling complex academic materials into focused research ideas can be cognitively overwhelming, while long waiting times for feedback can fragment their workflow and reduce motivation. For instructors, the challenge lies in maintaining the depth and consistency of feedback while handling the administrative workload for large cohorts.

TPRS addresses this challenge through AI-powered knowledge scaffolding [4,14], offering structured yet flexible support that preserves student autonomy. Rather than delivering fully automated outputs, the system guides students in refining their research focus while encouraging critical thinking [7,10]. By combining LLMs with expert- and knowledge-based recommendation strategies, TPRS analyzes student input, provides instant formative feedback, and suggests relevant academic literature to support topic development. Unlike purely generative AI tools, TPRS is designed to balance automation with pedagogical alignment. It acts as a cognitive scaffold, not a replacement for academic reasoning. In a pilot deployment, we applied the CRS-Que framework [5] to evaluate system usability, accuracy, transparency, and user trust. The results indicate that TPRS improves student engagement, accelerates the feedback cycle, and reduces instructor workload, while also raising important questions about explainability and responsible AI use. This study contributes to prototyping the development of AI-driven educational support tools by demonstrating how hybrid recommendation systems can enhance academic self-regulation and support research topic refinement in distance learning environments.

2 TPRS System Design and Development

2.1 Challenges in the Research-Oriented Term Paper Process

In the final semester of a Bachelor of Arts program in Culture and Social Sciences, students are required to write a term paper that synthesizes knowledge from prior coursework, particularly in Media Education and Media Pedagogy. This research-oriented semester is structured into three phases, yet each presents specific challenges for both students and instructors. During the Term Paper Preparation phase, students are expected to independently reflect on course materials, identify thematic interests, and formulate a coherent research topic along with corresponding research questions. This step is cognitively demanding: the process of narrowing broad content into a specific, academically viable focus often leads to overload and uncertainty. Many students struggle to define clear, well-scoped questions without substantial guidance, resulting in frustration and delays. The Student-Teacher Interaction phase typically unfolds via a Moodle discussion forum, where students submit their initial ideas for feedback. However, feedback is rarely immediate. These delays can interrupt students' momentum and diminish motivation. Moreover, the varying quality of submissions often necessitates multiple rounds of revision, turning the feedback loop into a prolonged and sometimes demotivating experience. Once a topic is approved, students proceed to the Term Paper Writing phase. Yet, time lost in earlier stages often compresses the actual writing window, which can negatively impact the final paper's quality. Instructors, meanwhile, face their own set of challenges. Providing individualized feedback through Moodle for large cohorts is time-consuming and difficult to scale. Maintaining consistency and depth of feedback across students with diverse skill levels and research interests is especially demanding. Some students need extensive support, while others require minimal input, leading to an uneven distribution of effort. Additionally, tracking multiple rounds of revisions creates an administrative burden that detracts from more pedagogically meaningful engagement.

2.2 System Architecture and Workflow

The TPRS workflow begins when a student submits an initial term paper proposal consisting of a research topic (RT) and one or more research questions (RQ_i, where $i \geq 0$). A sentiment analysis (SA) module evaluates the student's confidence in their submission. Based on this assessment, the system dynamically selects one of two recommender engines: the knowledge-based engine (KE) or the expert-based engine (EE). Confident submissions are routed to the KE, which recommends literature to deepen academic exploration. In contrast, submissions with low confidence activate the EE, which offers more foundational guidance, such as refining the topic or restructuring the questions. Students may also override the system's decision and manually select a preferred engine or language model (LLM), including ChatGPT-4o, Mixtral-8x7B-Instruct-v0.1, and Llama-3-8B. Following this, the system conducts structured validation using a set of evaluation modules (Eva_X) that assess the coherence and feasibility of the research topic ($\mathrm{Eva}_{\mathrm{RT}}$), each individual question ($\mathrm{Eva}_{\mathrm{RQ}_i}$), and topic-question pairs ($\mathrm{Eva}_{\mathrm{RT},\mathrm{RQ}_i}$). These validators are currently implemented using structured prompting techniques with ChatGPT-4o. Based on the validation results, the system generates positive or negative feedback accompanied by brief explanations to ensure transparency and trust. Two core recommendation scenarios are currently supported: (1) if the research topic is invalid but the questions are valid, the expert-based engine suggests a new topic aligned with the questions; (2) if both topic and questions are valid, the knowledge-based engine recommends up to three relevant academic sources to support further inquiry. Future iterations aim to expand this logic and generate more granular feedback across a broader range of submission patterns. To enhance recommendation quality, TPRS employs a multi-shot prompting strategy informed by historical student-teacher interactions. The system has been trained on a dataset of approximately 70 students from the same BA program, each with an average of 13âĂŞ15 exchanges. These interactions, collected in 2023 and the Winter 2023/2024 semester, are embedded into prompt templates to simulate high-fidelity tutoring. The resulting prompts are processed by a selected LLM, which generates structured, actionable feedback to help students refine their ideas. For literature recommendations, the system applies a cross-lingual semantic similarity algorithm. It generates a multilingual embedding of the student's topic using a model such as Multilingual BERT, then computes cosine similarity between this embedding and entries in a local academic repository. The most relevant references are retrieved and presented, ensuring that literature suggestions align closely with the student's research focus.

3 Pilot Evaluation

We conducted a pilot evaluation of TPRS within a Bachelor of Arts distance-learning program in Culture and Social Sciences at FernUniversität in Hagen. The study took place from 01.11.2024 to 20.12.2024 and included 18 student participants spanning six age groups. The majority identified as female (77.78%), with male (16.67%) and non-binary (5.55%) participants also represented. Most students (66.67%) were enrolled part-time, while the remainder studied full-time. Despite variations in demographic background and AI familiarity, participants had comparable academic progress, with most having completed 8 to 9 course modules.

To evaluate the system's usability and pedagogical impact, we applied the CRS-Que framework [5], which assesses conversational recommender systems along four dimensions: Perceived Qualities, User Beliefs, User Attitudes, and Behavioral Intentions. *Perceived Qualities* received good ratings across three criteria: accuracy, novelty, and explainability. Students found the recommendations relevant and helpful, with many highlighting their novelty and unexpected insights as a strength. The system's ability to explain its recommendations clearly contributed to user trust and transparency. *User Beliefs* were measured in terms of perceived ease of use, usefulness, and control. Participants reported that TPRS was intuitive and effective for academic planning tasks such as topic development and literature search. They appreciated having the option to choose among different engines and LLMs, reinforcing a sense of control. *User Attitudes*, including trust and satisfaction, were also positive. Students expressed confidence in the system's guidance and reported that it helped clarify their research direction. Satisfaction scores indicated a fair level of approval, particularly for feedback responsiveness. Finally, in the dimension of *Behavioral Intentions*, most students indicated a willingness to reuse TPRS for future assignments and recommend it to peers. Besides the CRS-Que framework, the authors added *Judgment on Quality of Work* dimension by asking the students: How would you rate the quality of your work on the three main tasks of finding a topic for your homework in the module? Please rate individually in the case study, theoretical references, selected topic, and central question.

As hypothesized, analysis of submission quality demonstrated that supporting students in their term paper planning process has the potential to increase submission quality. This increase in quality may have effects on the intensity of the following tutoring process and may yield efficiency gains in either time spent on or the depth of tutoring, benefiting both students and teachers in the long run. To assess academic impact, domain experts evaluated students' initial and final submissions using a structured rubric. The rubric included four academic criteria: use case description (3 points), theoretical background (3 points), research topic formulation (2 points), and quality of research questions (3 points), on a total 11-point scale. Two independent raters ensured interrater reliability, achieving substantial agreement (Cohen's $\kappa = 0.675$, $p < .001$). Discrepancies were resolved through consensus. A directional paired t-test revealed a significant improvement in submission quality from initial ($M_{t1} = 6.79$, $SD_{t1} = 2.07$) to final ($M_{t2} = 7.53$, $SD_{t2} = 1.74$) drafts, with a small to moderate effect size (Hedge's $g = 0.44$, $p < .05$). A Shapiro-Wilk test confirmed that the initial scores deviated from normality, but further analysis retained all data points. Importantly, demographic covariates (age, gender, study mode, academic progress) showed no significant correlation with the observed improvement. These findings indicate that TPRS positively influenced the quality of students' academic submissions while offering a scalable model for reducing instructor workload and supporting inquiry-based learning.

4 Technical Report

The TPRS has been developed as a fully integrated system comprising both frontend and backend components. The system facilitates seamless interactions between students and AI-powered feedback mechanisms, replacing the conventional "to be reviewed by teachers" and "teachers post feedback" blocks in the workflow with "to be reviewed by AI" and "AI's instance feedback," respectively. The TPRS system was designed and implemented from scratch to ensure scalability, modularity, and compliance with educational and data protection standards.

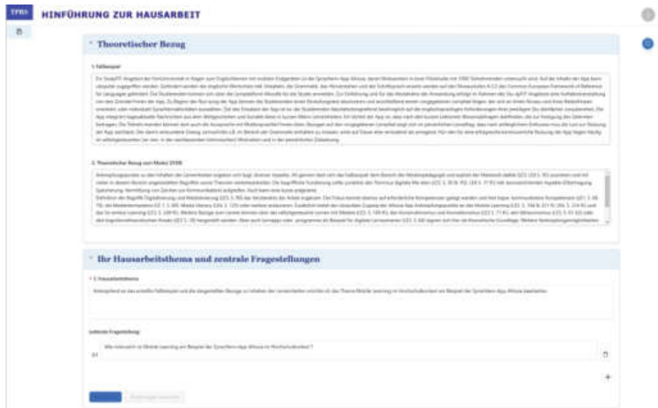

Fig. 1. The TPRS interface: Students' term paper ideas, reflection on learning resources, and several research questions.

During the pilot study, students interacted with TPRS using three selectable LLMs. The most frequently chosen model was Mixtral-8x7B-Instruct-v0.1 (54 interactions), followed by ChatGPT-4o (41 interactions) and Llama-3-8B (37 interactions). On average, each interaction involved approximately 2,606 input tokens and 472 output tokens. Mixtral produced the most consistent outputs, typically around 512 tokens, while ChatGPT-4o generated more variable responses, with a peak of 648 tokens. Llama-3-8B generally produced shorter outputs, reflecting a more conservative response style. The total cost of deploying TPRS during the pilot was $1004.59, primarily incurred through AWS services. The largest expenses stemmed from OpenSearch, SageMaker, and EC2 instances, which together accounted for over 80% of the total. OpenSearch alone made up nearly 40%. Other services used–such as Lambda, CloudWatch, and Secrets Manager–were part of AWS's free tier. The cost analysis confirms that the system is feasible for medium-scale deployment in educational settings.

The TPRS frontend provides students with an intuitive interface to submit proposals, select models, and view feedback history. Built using HTML, CSS, JavaScript, and Node.js, the interface adheres to accessibility standards and supports personalized feedback tracking. Package management is handled via YARN, and user data is stored using MongoDB. The backend is implemented with FastAPI for efficient RESTful API handling, and LangChain is used to manage interactions across multiple LLMs. ChatGPT-4o is accessed via OpenAI's API, while open-source models are integrated through Hugging Face. A modular architecture allows easy integration of new models or APIs in future iterations. Continuous integration and deployment (CI/CD) pipelines are in place to support iterative improvements. The system is hosted on AWS infrastructure within the EU to ensure compliance with GDPR and other data protection regulations. Several screenshots of TPRS are illustrated in Figs. 1 and 2. TPRS also integrates with the university library system, enabling retrieval of academic references aligned with students' topics. This connection ensures that literature recommendations are pedagogically grounded and domain-specific.

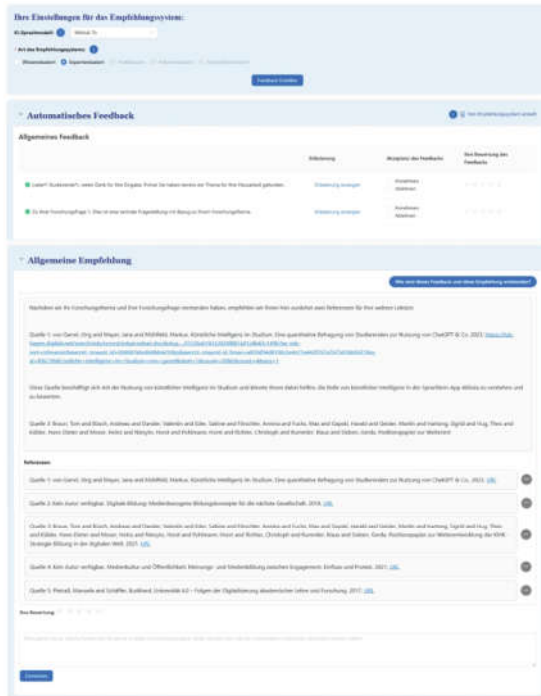

Fig. 2. The TPRS interface: Automatic feedback and recommendation.

5 Conclusion

This paper presented TPRS, an AI-assisted system designed to scaffold the development of term paper topics and research questions in distance learning environments. Combining LLMs with knowledge-based and expert-driven recommendation strategies, TPRS adapts its feedback based on students' confidence levels and leverages historical supervision data to simulate high-fidelity guidance. Structured validation modules and semantic similarity algorithms ensure that feedback and literature recommendations are contextually relevant and academically sound. Through a pilot deployment with students in a Bachelor of Arts program, TPRS demonstrated its ability to improve submission quality while reducing instructor workload. Evaluation using the CRS-Que framework confirmed high levels of perceived accuracy, usability, and trust. Technically, the system integrates modern web technologies, modular LLM orchestration via LangChain, and scalable cloud deployment on GDPR-compliant infrastructure.

Acknowledgments. The authors kindly appreciate the support of CATALPA, Center of Advanced Technology for Assisted Learning and Predictive Analytics, FernUniversität in Hagen, through the project "AI.EDU Research Lab 2.0".

References

1. Al-Zahrani, A.M., Alasmari, T.M.: Exploring the impact of artificial intelligence on higher education: the dynamics of ethical, social, and educational implications. Humanit. Soc. Sci. Commun. **11**(1), 1–12 (2024)
2. Bailón, A.B., Contreras, W.F., Molina-Solana, M.: Intelligent tutoring system, based on video e-learning, for teaching artificial intelligence. In: Pérez, J.B., et al. (eds.) Trends in Practical Applications of Agents, Multi-Agent Systems and Sustainability - The PAAMS Collection, 13th International Conference, PAAMS 2015, Salamanca, Spain, June 3–4, 2015, Special Sessions. Advances in Intelligent Systems and Computing, vol. 372, pp. 215–224. Springer (2015)
3. Castillo-Martínez, I.M., Flores-Bueno, D., Gómez-Puente, S.M., Vite-León, V.O.: AI in higher education: a systematic literature review. Front. Educ. **9**, 1391485 (2024)
4. Ertugruloglu, E., Mearns, T., Admiraal, W.: Scaffolding what, why and how? a critical thematic review study of descriptions, goals, and means of language scaffolding in bilingual education contexts. Educ. Res. Rev. 100550 (2023)
5. Jin, Y., Chen, L., Cai, W., Zhao, X.: CRS-que: a user-centric evaluation framework for conversational recommender systems. ACM Trans. Recommender Syst. **2**(1), 1–34 (2024)
6. Kaswan, K.S., Dhatterwal, J.S., Ojha, R.P.: AI in personalized learning. In: Advances in Technological Innovations in Higher Education, pp. 103–117. CRC Press (2024)
7. Kim, J., Lee, H., Cho, Y.H.: Learning design to support student-AI collaboration: perspectives of leading teachers for AI in education. Educ. Inf. Technol. **27**(5), 6069–6104 (2022)
8. Miao, F., Cukurova, M.: AI competency framework for teachers (2024)
9. Miao, F., Shiohira, K.: AI competency framework for students (2024)
10. Munshi, A., Biswas, G., Baker, R., Ocumpaugh, J., Hutt, S., Paquette, L.: Analysing adaptive scaffolds that help students develop self-regulated learning behaviours. J. Comput. Assist. Learn. **39**(2), 351–368 (2023)
11. Ouyang, F., Zheng, L., Jiao, P.: Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020. Educ. Inf. Technol. **27**(6), 7893–7925 (2022)
12. Perez-Ortiz, M., Novak, E., Bulathwela, S., Shawe-Taylor, J.: An AI-based learning companion promoting lifelong learning opportunities for all. arXiv preprint arXiv:2112.01242 (2021)
13. Shi, H., Aryadoust, V.: A systematic review of AI-based automated written feedback research. ReCALL 1–23 (2024)
14. Zeng, J., Zhang, P., Zhou, J., Shang, J., Black, J.B.: The impact of embodied scaffolding sequences on stem conceptual learning. Educ. Technol. Res. Dev. 1–26 (2024)