Combining NLP, Speech Recognition, and Indexing: an AI-based Learning Assistant for Higher Education

Tobias Schmohl[1], Kathrin Schelling[1], Stefanie Go[1], Carolin Freier[2], Marianne Hunger[3], Franziska Hoffmann[4], Anne-Kathrin Helten[5], Florian Richter[6]

OWL Technical University of Applied Sciences and Arts, Germany[1]
Nuremberg Institute of Technology, Germany[2]
BayZiel – Bavarian Centre for Innovative Teaching, Germany[3]
Ansbach University of Applied Sciences, Germany[4]
Lutheran University of Applied Science, Germany[5]
Ingolstadt University of Applied Sciences, Germany[6]

Abstract

This paper presents the ongoing development of HAnS (Hochschul-Assistenz-System), an Intelligent Tutoring System (ITS) designed to support self-directed digital learning in higher education. Initiated by twelve collaborating German universities and research institutes, HAnS is developed 2021–2025 with the goal of utilizing artificial intelligence (AI) and Big Data in academic settings to enhance technology-based learning. The system employs AI for speech recognition and the indexing of existing learning resources, enabling users to search and compile these materials based on various parameters. Here, we provide an overview of the project, showcasing how iterative design and development processes contribute to innovative educational research in the evolving field of AI-based ITS in higher education. Notwithstanding the potential of HAnS, we also deliberate upon the challenges associated with ensuring a suitable dataset for training the AI, refining complex algorithms for personalization, and maintaining data privacy.

Keywords: *artificial intelligence in education (AIED), teaching/learning materials, intelligent tutoring system (ITS), adaptive learning, digital education*.

1. Introduction

In the era of rapid digital transformation, higher education is witnessing an increased adoption of technology-driven learning. As students turn towards technology, educational institutions need to keep pace with this development by providing Learning Assistance Systems (LAS) that are both effective and adaptable [1]. This progression, in turn, poses new challenges for researchers and policymakers alike: In order to use emerging technologies to foster self-directed learning, they need to fully comprehend the strengths and limitations of AI-based systems before they can make informed decisions about their efficient use in educational settings.

In this context, we present HAnS, an AI-driven Intelligent Tutoring System (ITS) for higher education that is currently being developed by eight German Universities of Applied Sciences (UAS) and four associated research institutions. Traditionally, an ITS is software that provides both instructions and immediate feedback without the need for human intervention, thereby helping learners understand academic concepts within personalized learning environments [2]. The HAnS project expands this concept by integrating Natural Language Processing (NLP) and speech recognition technologies with the indexing power of AI-based systems to create an ITS with the ability to adapt to users' individual learning strategies (micro-adaptivity) and select suitable instructional tasks for each student (macro-adaptivity) [3].

The AI component of HAnS is trained using authentic audio-visual teaching/learning materials (TLM) such as audio recordings, video clips, images, and texts, which are provided by instructors from various German universities. The ITS combines these materials with its AI-based framework to provide three main features that both instructors and learners can use to elevate the effectiveness of self-study in higher education. HAnS uses (1) automatic transcription and indexing of the TLM to allow users to search the recordings for specific topics. This intelligent search is complemented by (2) a recommender system that adaptively generates TLM dossiers for self-study, based on user

information and educational guidelines embedded in HAnS. Additionally, the ITS provides (3) automatically generated exercises based on predefined patterns and teaching materials.

This paper outlines the overall design and development process of HAnS, discusses the research methodologies employed within the project, and highlights potential challenges and future directions. Our objective is to shed light on the evolution of AI-based ITS and to contribute new insights into the possible role of AI-enhanced learning in higher education.

2. HAnS: A multi-faceted approach to AI-based ITS for higher education

The HAnS system's design and implementation take into account the pedagogical advantages of Artificial Intelligence in Education (AIEd) as underscored by Zawacki-Richter et al. [4]. Following the suggestions of Munir et al. [5], the project incorporates machine learning and deep learning algorithms to create an ITS that recognizes individual learning objectives and adaptively responds to them by providing learning resources that match the unique needs of each student. Shifting the paradigm from a traditional one-size-fits-all approach to digital teaching and learning towards more personalized and effective learning environments, HAnS aims to enhance the learning experience and empower students. By providing learners with access to a wide range of educational materials and offering personalized search results and exercises, the ITS fosters a sense of independence among students, enabling them to study at their own pace and to independently track their own progress.

The development of HAnS is an ongoing process. An interdisciplinary team continually reassesses and adapts the system through Design-Based Research (DBR) cycles [6], aiming to align the features of HAnS with the needs and learning habits of students in higher education and, in the long run, integrate the ITS into everyday teaching and learning processes. The following sections elaborate on the different research approaches within the HAnS project and illustrate how the research that guides the development process may also offer a broader perspective on the potential of AI in higher education.

2.1 Assessing the longitudinal impact of AI-based ITS on teaching and learning processes

Since AI-based ITS represent a new approach to digital learning, longitudinal data on their influence on teaching and learning processes is scarce. Researchers at the Lutheran UAS Nuremberg therefore use the HAnS project's iterative DBR processes to investigate the long-term effects of ITS in higher education. Utilizing an impact model [7] for the HAnS system, areas of competence—e.g., academic self-efficacy and digital competencies—are gauged via online questionnaires, which are administered to students and teachers twice per semester. The first survey takes place at the beginning of the semester, ideally before the users' first contact with HAnS, the second survey follows at the end of the semester. If feasible, a comparison group of students who did not use HAnS at all will also be surveyed in order to apply methods of modern causal analysis (propensity score matching, difference-in-difference approach) [8] and generate additional hypotheses concerning the effects of AI-based ITS on their users.

2.2 Incorporating user feedback into the development of HAnS

Within the DBR framework, the technical development of HAnS is adapted iteratively to prospective users' needs and their respective teaching and learning practices as they emerge during field tests. The research team at Nuremberg Institute of Technology applies two qualitative methods to create a feedback loop through which students' and instructors' requirements can be channeled into the further development of HAnS. The researchers conduct online group discussions with students and instructors and use documentary analysis [8] to reconstruct distinct user group experiences and value patterns regarding digital teaching and learning in general and the HAnS system in particular. Ethnographic case studies [9] provide further insight into users' actual learning practices. For this, students are observed while they interact with the HAnS prototype to either solve specific tasks or pursue individual learning goals while preparing for exams. Subsequent interviews explore the students' experience as well as their expectations concerning the system's features and functionality.

2.3 The role of educational theory and research in HAnS development

To develop an ITS that supports teaching and learning processes in higher education, HAnS also requires an appropriate didactic framework. The researchers at OWL UAS aim to ground this framework in educational theory and qualitative research. Unlike the quantitative and qualitative evaluations described above, this educational research does not focus on the use and impact of the HAnS system, but on its potential as a tool for effective digital teaching and learning. Preliminary investigations into the didactic scaffolding for HAnS as an AI-based ITS for higher education therefore engaged various stakeholders such as educational specialists and potential users. Ten problem-

centered interviews [10] explored students' perspectives on the potential benefits and risks associated with AI-based applications in higher education, and in a focus group setting, two mixed groups of stakeholders discussed their expectations and ideas regarding automatically generated exercises provided by AI-based systems. These preliminary studies were used to underpin the creation of use cases for HAnS which will now guide didactic scaffolding in the upcoming DBR cycle.

2.4 Further avenues of research

Within the HAnS project, three additional approaches to AI-based teaching and learning in higher education are currently under investigation. Each of these research groups, as equal partners in the project, contributes insights that drive the iterative design process. However, as their research is still in its preliminary stages, this section provides only a brief overview of their methods and objectives. Acceptance research at Ingolstadt UAS explores learners' expectations and needs concerning the further development of HAnS. Quantitative online surveys based on a validated Likert scale are used to gauge students' attitudes and experiences concerning AI in higher education. Researchers at Ansbach UAS employ the open-source web analytics platform Matomo to analyze students' click behavior. This quantitative evaluation of the system aims to identify which content, functions, and services are the most or least utilized, when users exit HAnS, and how their behavior evolves over time. Meanwhile, the team at the Bavarian Centre for Innovative Teaching (BayZiel) collaborates with the HAnS developers to create a tool based on Mayer's multimedia principles [11] which will allow the ITS to automatically assess whether the instructional design of a teaching video prevents cognitive overload by presenting information via different sensory modes. This way, HAnS provides feedback to content creators and enforces quality control without infringing on instructors' autonomy concerning course content.

3. From research to implementation: Shaping the HAnS system

As the digital era continues to unfold, the development of effective and adaptable LAS becomes increasingly important. The HAnS project, as described in this article, addresses the need for a new generation of tools for higher education by equipping an ITS with AI-based features that promote personalized learning and learner autonomy. This way, HAnS contributes to the ongoing evolution of higher education by demonstrating the potential of AI and Big Data to enhance technology-based learning and by providing multi-faceted research on current trends in digital teaching and learning.

While HAnS represents a significant advancement in the field of AI-based ITS and will, hopefully, contribute to the development of new pedagogical approaches that support self-directed learners in a digital age, the project itself is still ongoing. Until the end of 2025, the features of HAnS will be continually reassessed and adapted over the course of several DBR cycles. This iterative process, guided by user feedback and educational research, aims to match the system's functionality to the needs of students and instructors in higher education.

Due to the current lack of best practices regarding the creation of AI-based tools for higher education, however, there are still many obstacles to overcome. As with any AI-based system, data privacy issues are of the utmost importance. Ensuring that user data is securely stored and used in a manner that respects users' privacy is a non-negotiable requirement for the successful implementation of HAnS in everyday teaching and learning practices at any educational institution. But this is only one of many concerts since the AI-based features of the ITS pose additional challenges that must be addressed by developers and educational researchers alike. For one, the effectiveness of HAnS strongly depends on the quality and variety of the dataset used to train the AI. If the TLM that are used as training materials lack breadth in terms of subject matter or depth regarding the complexity of the topics discussed, the system's ability to generate comprehensive and reliable learning resources may be compromised. Additionally, while the personalization aspect of HAnS is promising, it is also contingent upon the system's ability to accurately interpret and respond to individual learning objectives. To ensure that HAnS provides truly personalized learning experiences,complex AI algorithms must be continually tested and refined, and the automatic generation of exercises requires similarly careful monitoring and evaluation. If HAnS is to provide learners with immediate feedback and opportunities for personalized practice, the quality of the automatically generated exercises is crucial. If they are not aligned with the learning objectives or lack pedagogical rigor, these exercises may hinder rather than facilitate learning.

4. Conclusion

The HAnS project is an approach to the digitization of higher education that emphasizes self-directed, personalized learning that is supported by the application of emerging technologies. This paper contributes to the field of AI-based ITS by detailing the development process of HAnS and highlighting

current challenges as well as future directions. Our findings provide valuable insight for researchers and practitioners interested in leveraging AI and Big Data to enhance self-directed learning in higher education. As the HAnS project continues, we anticipate its potential to contribute to the creation of digital learning environments that are student-centered, adaptable, and effective.

5. References

[1] Kobysheva, L., Luginina, A., Gafiatulina, N., & Artamonova, Y. (2021). Organization of higher education in context of digitalization: Online learning experience at pandemic, development trends. E3S Web of Conferences, 273.

[2] Soofi, A. A., & Moiz, U. A. (2019). A systematic review of domains, techniques, delivery modes and validation methods for intelligent tutoring systems. International Journal of Advanced Computer Science and Applications (IJACSA), 10(3), 99–107.

[3] Rus, V., D'Mello, S., Hu, X., & Graesser, A. (2013). Recent advances in conversational intelligent tutoring systems. AI Magazine, 34(3), 42–54.

[4] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39.

[5] Munir, H., Vogel, B., & Jacobsson, A. (2022). Artificial intelligence and machine learning approaches in digital education: A systematic revision. Information, 13(4), 203.

[6] Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? Educational Researcher, 41(1), 16-25.

[7] Ottmann, S., & König, J. (2018). Was wirkt wie? Konzeptionelle Überlegungen zur Messung und Analyse von Wirkungen in der Sozialen Arbeit. Der Wirkungsradar des Instituts für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg.

[8] Bohnsack, R. (2010). Documentary method and group discussions. In R. Bohnsack, N. Pfaff, & W. Weller (Eds.), Qualitative analysis and documentary method in international educational research (pp. 99–124). Opladen: Budrich.

[9] Harwati, L. N. (2019). Ethnographic and case study approaches: Philosophical and methodological analysis. International Journal of Education and Literacy Studies, 7(2), 150–155.

[10] Witzel, A. & Reiter, H. (2012). The Problem-centred Interview. Bd.: Principles and practice. London: Sage.

[11] Mayer, R. (2021). Evidence-based principles for how to design effective instructional videos. Journal of Applied Research in Memory and Cognition, 10(2), 229–240.