Competence in Minutes: How Non-Expert Individuals Become Educators – A Live Experiment with AI-Powered Avatar Didactics

Thanh Vivien Le-Vu[1], Arnim Spengler[2], Rischert Enes[3], Khaoula Benmaarouf[4] and Felix Meckmann[5]

[1]Ruhr West University of Applied Sciences, Muelheim a. d. Ruhr, Germany, vivien.le-vu@hs-ruhrwest.de
[2]Ruhr West University of Applied Sciences, Muelheim a. d. Ruhr, Germany, postfach@arnim-spengler.de
[3]Ruhr West University of Applied Sciences, Muelheim a. d. Ruhr, Germany, rischert.enes@hs-ruhrwest.de
[4]Ruhr West University of Applied Sciences, Muelheim a. d. Ruhr, Germany, khaoula.benmaarouf@hs-ruhrwest.de
[5]Ruhr West University of Applied Sciences, Muelheim a. d. Ruhr, Germany, felix.meckmann@hs-ruhrwest.de

Abstract

This paper investigates the integration of Artificial Intelligence (AI) in higher education, exemplified by the KI4Edu project. It responds to students' growing expectations for digital, flexible, and personalized learning, alongside increasing institutional challenges such as staff shortages and rising demand. Large Language Models (LLMs) offer transformative potential but also introduce risks, including hallucinations, opacity, and limited contextual accuracy. This study presents a pedagogical framework and technical architecture for responsible AI deployment. Educators assume the role of mentors, ensuring quality assurance, critical reflection, and content curation, while AI supports knowledge accessibility, structuring, and scalability. The approach is implemented through a Retrieval-Augmented-Generation (RAG) system that consolidates verified teaching materials within a knowledge database, with an avatar serving as an interactive interface for personalized learning pathways.

ISSN 2991-9371

Key contributions include: (1) a systematic analysis of challenges in AI-assisted education, (2) a pedagogical model redefining educators' roles, and (3) a technical concept for avatar-based knowledge transfer. The paper discusses limitations, such as LLM output variability and social risks like "AI Psychosis," and highlights the need for monitoring mechanisms, scalable system architectures, and a balanced integration of technological innovation with human responsibility in teaching.

Keywords: *Artificial Intelligence in Higher Education; Role of Educators as Mentors; Didactic Guiding Principles in Digital Learning; Personalized Learning Paths; Large Language Models (LLM) in Education*

1. Introduction and Project Context

This paper presents current project results of the KI4Edu research project and provides an outlook on how these can be further developed.

The ongoing digitalization opens up new opportunities for the education sector but simultaneously poses major challenges for universities. Students increasingly expect location-independent, digitally accessible content and the use of modern technologies. Especially the use of Artificial Intelligence (AI) requires an adjustment of roles, methods, and processes [1]. Traditional attendance models are reaching their limits in view of staff shortages and growing demand for individualization of teaching. AI can support here by making content available faster, scaling interactions, and enabling personalized learning paths [2]. Nevertheless, the role of educators remains indispensable: didactic guidance, reflection, and securing specialized knowledge are still central tasks [3-5]. Against this background, the collaborative project KI4Edu of Ruhr-West University of Applied Sciences (HRW) and the University of Duisburg-Essen (UDE), funded by the Foundation for Innovation in University Teaching, investigates the responsible use of AI in teaching. The goal is to develop practical concepts that support educators in organization and mediation and simultaneously enable students to use AI reflectively. The focus is on changing roles of educators and learners as well as the didactic potentials and limitations of Large Language Models (LLM).

The article presents a case study on the development of an AI-supported system that makes teaching content digitally available and addresses weaknesses of current LLMs. Educators take on the role of knowledge providers and mentors who ensure the quality of content and guide didactic decisions. Students interact with a digital agent, prospectively with an avatar-based system, which serves as a low-threshold interface to complex knowledge. This creates a balance between technological scalability and pedagogical responsibility.

The paper provides three contributions: (1) a systematic analysis of the challenges of using AI in teaching, (2) a didactic guiding principle for redefining the role of educators, and (3) a technical concept that combines AI- supported data preparation with avatar-based mediation. Finally, limitations and perspectives for further development are discussed.

2. Challenges of Increased AI Use in Teaching

The increasing use of Artificial Intelligence in higher education and its acceptance opens up new possibilities but is also associated with a number of challenges. These concern didactic and epistemological questions as well as social and psychological dimensions, which are listed below and then discussed [6,7]:

- Student expectations
- Problematic dimensions of AI knowledge
- Unreliability of AI knowledge
- Unknown knowledge sources
- Hallucinations
- Not exactly the knowledge to be conveyed
- Missing individual knowledge and focus
- Missing real-world knowledge of specialists
- Didactic guidance by specialists
- Black-box problem
- Social and psychological challenges: "AI Psychosis" [8]

A central tension arises from the growing expectation of students to take digital and AI-supported offerings for granted and to actively use them privately. Universities are thus under innovation pressure, while educators often lack the necessary technical and didactic competencies. Moreover, the behavior and outputs of LLMs are difficult to assess. The lack of transparency of AI-generated knowledge is problematic. Origin and sources often remain unclear or lack academic character, which contradicts the academic demand for traceability. In addition, there is the susceptibility to "hallucinations" and the tendency to deliver only generic knowledge. Subject-specific "real-world knowledge" – practical examples, research experiences, or critical reflections – remains irreplaceable and cannot currently be mapped by LLMs, or only incompletely.

Didactic limitations also emerge: AI conveys content but can only limitedly control learning processes or make pedagogical decisions. Educators are more than mere knowledge transmitters – they set learning goals, promote motivation, and reflect on learning progress, which remains essential for sustainable competence acquisition [9].

Another problem is the black-box nature of AI systems. Educators often do not know which content is conveyed and by what criteria it is selected. Moreover, outputs can vary greatly and improve or worsen with each new LLM generation. These varying responses from language models raise questions about reproducibility and reliability, which contradict scientific standards. Instructions to the LLM do not offer an adequate means to address this challenge.

Furthermore, social and psychological risks exist. Phenomena such as "AI Psychosis" – the perception of AI as a friend or authority – can promote social isolation and loss of reality. Universities therefore bear a special responsibility to train students in the reflective use of AI and to ensure real interactions [10].

Even if these challenges are not exhaustive, it is clear that the use of AI in teaching must be reflected not only technically but also in terms of content, didactics, ethics, social aspects, and psychology. Only in this way can potentials be utilized without compromising quality, reliability, and responsibility. AI systems currently cannot do this, and humans can only do so to a limited extent.

3. Consequences and Needs for Action

From the described challenges and the didactic guiding principle, consequences for university teaching and needs for action at institutional and individual levels arise. The integration of AI must not be understood only as a technological task but requires pedagogical, organizational, and societal rethinking. Central is the development of new didactic concepts that consider the strengths and weaknesses of AI. While AI supports the provision and structuring of information, reflection, categorization, and discussion remain inextricably linked to the role of educators. In addition, information and source criticism must be systematically strengthened. Students should learn to critically examine AI-generated content and classify it scientifically. Finally, the role of educators changes: instead of pure knowledge transfer, they increasingly assume control and moderation tasks to ensure that AI content is used reflectively and embedded in scientific discourses.

The following table summarizes central needs for action and clarifies their relation to the previously mentioned challenges:

Challenge	Need for Action	Brief Explanation
Student Expectations	Development of new didactic concepts	Combination of AI-supported learning phases with human supervision (e.g., blended learning).
Unreliability & Hallucinations	Strengthening information and source criticism	Development of media literacy for examining and verifying AI content.

Missing Didactic Guidance	Higher responsibility of educators	Teachers become instances of reflection and orientation, not just knowledge transmitters.
Black-Box Problem	Transparency & Regulation of AI systems	Institutional guidelines, transparent data spaces, and logging of AI interactions.
Missing Real-World Knowledge	Integration of subject-specific and experiential knowledge	Lecturers specifically incorporate practical and research experiences into teaching.
Social Risks ("AI Psychosis")	Prevention & Education	Creation of discourse spaces and education about psychological dangers.
Different Competence Levels among Educators	Further training programs	Training in didactics, technology, and critical use of AI.

Table 1 Challenges and approaches to AI-supported teaching
(*Source: own illustration*)

The following matrix estimates the risks of the mentioned challenges. It aims to show to what extent the challenge can be controlled or compensated.

Figure 1 Risk Matrix: Challenges of AI in Higher Education,
(*Source: Own illustration, created with ChatGPT (OpenAI, 2025).*)

ISSN 2991-9371

Risk Matrix:

- Y-axis (Relevance): How strongly the challenge influences the quality of teaching.
- X-axis (Controllability): How well the challenge can be controlled or compensated.

The risk matrix shows which problems are highly relevant but difficult to control (e.g., black-box problem, student expectations) – these are the most critical areas for action.

Overall, it is clear that AI does not lead to the displacement of traditional university teaching but to its transformation. Educators remain the central authority, but their role changes towards didactic steering and quality assurance. This requires, on the one hand, institutional support through guidelines, infrastructures, and further training programs, and on the other hand, a stronger anchoring of reflection, transparency, and media literacy in the students' learning process.

4. Technical Implementation

The technical implementation of the concept is based on procedures developed in the KI4Edu project for the systematic preparation and use of teaching content. The goal is to create a verifiable knowledge base that feeds the avatar as a didactic interface and reduces the risks of generic AI outputs.

A first step is data preparation: materials such as audio and video recordings, presentations, or scripts are automatically collected, standardized, and transferred into a consistent structure. This creates a solid foundation for the further technical components.

The core component is the Retrieval-Augmented Generation (RAG) system [10]. Instead of generating answers exclusively from a generic language model, it first accesses a project-specific knowledge database. For this purpose, the teaching materials are broken down into smaller semantic units, provided with vectors (embeddings), and stored in a vector database. This approach enables:

- a fact-based search that ensures the relevance of content,
- the linking with metadata (e.g., slides, video clips, transcripts),
- the reduction of hallucinations and content deviations.

User Prompt Template (Server-Assembled)

Listing 3: User message template (assembled with query + hints).

```
user_prompt = f"""
    ## Carefully review the above checklist. YOU MUST include and
       explain each formula and image listed.
    <check>{checklist}</check>


    ##  Question: <query>{query}</query>

    ## Context
    <context>
    {build_context_for_generation(retrieved_docs)}
    </context>
    The generated HTML should be well formated ensure the closing tags
       are right and that they respect the rules.
    Generate clean HTML response in German:
"""
```

Figure 2 Code example for inserting a user message
(Source: KI4Edu [1])

Based on this system, the generation and testing of answers take place. Every AI-generated output is validated with regard to three central criteria:

- content accuracy,
- coherence in the technical context,
- didactic suitability.

Validation is carried out by specialists who, in their role as mentors, retain curation authority and thus ensure the alignment of AI outputs with learning objectives and quality standards. After successful testing, specific learning LLMs for individual courses or topics can be released. They form the basis for avatar-based teaching scenarios and enable individualized, asynchronous student support. Continuous monitoring and feedback integration ensure the ongoing development of the systems.

Figure 3 System – Frontend
(*Source: Benmaarouf* [10])

At the same time, the technical implementation brings its own challenges. These include:

- the reproducibility of answers, as language models can vary depending on prompting,
- the scalability of the system, especially with large and heterogeneous data volumes,
- the integration into an avatar interface that not only reproduces text but also communicates in a didactically meaningful way.

Thus, the technical implementation represents a central building block for integrating AI-supported didactics into university teaching in a practical and reliable manner. It creates the prerequisites for supplying the avatar system developed in the project with reliable and quality-assured content and simultaneously mitigating the black- box problem [10].

5. Mentor as a Central Authority

The system concept combines technological innovation with a clear didactic framework. At its center is the educator as a mentor, who takes over didactic control and ensures the quality of learning processes. An avatar could be used as a human-computer interface. The avatar fulfills several functions: On the one hand, it marks curated content and prepares and presents content in an accessible, interactive form. In addition to knowledge transfer, it can also take on didactic tasks, such as asking comprehension questions, explaining content at different levels of complexity, or moderating short interaction sequences.

The educator in the role of a mentor, however, remains the central authority of the system. Their tasks can be divided into three core areas:

- Curation Authority: Educators determine which content flows into the knowledge pool and which learning objectives are pursued.
- Command Post: Through a monitoring interface, educators can view the avatar's interactions with students to create transparency regarding the didactic concept chosen by the AI.
- Didactic Intervention: Educators intervene specifically when misunderstandings arise, additional motivation is required, or learning objectives need to be adjusted.

The concept also enables the involvement of non-experts in teaching. Through the presented concept, tutors or other persons without comprehensive expertise can control the avatar, which is supervised by the mentors in the background. Since this avatar exclusively accesses verified content from the RAG system and is curated by

the mentor, the professional quality is maintained while educators are relieved of routine tasks. At the same time, the system supports both the scaling and individualization of teaching. The avatar addresses large student groups, while AI-supported learning paths open up personalized support. Educators, in their mentor role, retain the ability to provide individual feedback and guide reflection processes.

6. Discussion, Limitations & Outlook

A promising approach lies in further developing a mentoring system. This opens up three central areas of action for educators:

- Curation authority over the knowledge corpus and the definition of learning objectives,
- Live insight into avatar interactions to create transparency regarding the black-box problem,
- Didactic intervention to clarify misunderstandings, promote motivation, and adequately support performance.

In addition to these opportunities, risks and limitations remain. These include:

- limited reproducibility of LLM responses and the risk of invalid or inconsistent content,
- potential security risks such as prompt injection attacks that can undermine the system's functionality,

- the still existing black-box problem, even if it is partially mitigated by monitoring,
- social risks such as the phenomenon of "AI Psychosis," where students view AI systems as social reference persons and neglect real interactions.

Several future perspectives emerge for further development:

- Establishment of command posts and monitoring systems for educators to increase transparency and controllability,
- Stronger scaling of teaching by relieving routine tasks, allowing educators more time for support and reflection,
- In-depth research on the balance between technical efficiency and the preservation of human responsibility in educational processes.

Overall, it is clear that the presented methodology can make an important contribution to the further development of AI-supported teaching. At the same time, it makes clear that responsible integration can only succeed if technological innovations are combined with clear didactic guidelines and a strong role for educators.

7. Acknowledgements

The authors would like to acknowledge the "Stiftung Innovation in der Hochschullehre" (Foundation for Innovation in Higher Education) for funding the KI4Edu project, which provided the empirical data for this paper.

8. References

[1] Enes, R., Meckmann, F. (2025). KI4Edu: A Study on the Use of Artificial Intelligence in Higher Education. Mülheim an der Ruhr: Hochschule Ruhr West. https://www.researchgate.net/publication/392518262_KI_in_der_Lehre_- _Eine_aktuelle_Ubersicht_uber_die_Nutzung_von_KI_im_Studienkontext accessed on Jul, 23, 2025

[2] Canada Global Academy, "Bridging the Gap: Democratizing Access to Quality Education with AI," https://canadaglobalacademy.com/bridging-the-gap-democratizing-access-to-quality-education-with-ai/, Accessed Jul, 7, 2025.

[3] Kudirat A., Abdullahi S., Rafiu Dimeji S & Upeksha M (2023). Impact of Artificial Intelligence on Construction Project Management. https://www.researchgate.net/publication/373755997_The_Impact_of_Artificial_Intelligence_on_Construction_Co sting_Practice Journal of Engineering and Applied Sciences, 10(2), 112-125. Accessed Aug, 25, 2025

[4] T. Schmohl, J. Löffl, and G. Falkemeier, "Künstliche Intelligenz in der Hochschullehre," In: T. Schmohl (Hrsg.) and D. Schäffer (Hrsg.), Lehrexperimente der Hochschulbildung. Didaktische Innovationen aus den Fachdisziplinen, Volume 2, Bielefeld: wbv, 2019, pg. 117-122.

[5] MarketScale, "AI in Education Can Democratize Expertise — But Only If Systems Evolve," https://marketscale.com/industries/education-technology/ai-in-education-can-democratize-expertise-but-only-if- systems-evolve/, accessed on May, 14, 2025

[6] Spengler, A.J.; Le-Vu, T.V.; Enes, R.; Meckmann, F. (2025a). "Projektreport KI4Edu: Technische Auswahlkriterien für den Einsatz von Large Language Models (LLM) in der Lehre", Hochschule Ruhr-West, unpublished.

[7] Spengler, A.J.; Le-Vu, T.V.; Enes, R.; Meckmann, F. (2025b). "Projektreport KI4Edu: Eignungsuntersuchungen für den Einsatz von Large Language Models (LLM) in der Lehre", Hochschule Ruhr-West, unpublished.

[8] BBC. Microsoft boss troubled by rise in reports of 'AI psychosis, https://www.bbc.com/news/articles/c24zdel5j18o, Accessed Aug.23, 2025.

[9] KMZ Esslingen, "Entwicklung eines KI-gestützten Lern-Mentors," https://kmz-es.de/ki-textgeneratoren-in-schule- und-unterricht-teil-3-entwicklung-eines-ki-gestuetzten-lern-mentors/, Accessed Jul, 10, 2025.

[10] Benmaarouf, K. (2025). Retrieval-Augmented Generation (RAG) Project Report, Technologies, Data Processing, Indexing, and Streaming Interaction, unpublished.