Intelligent Interaction Platform for Personalized Digital Tutor: Advancing Empathetic and Adaptive Learning Experiences

Kai Zhang
kkzhang08@ustc.edu.cn
University of Science and Technology of China, Hefei, China

Weiyin Gong
wgong@mail.ustc.edu.cn
University of Science and Technology of China, Hefei, China

Qi Liu✉
qiliuql@ustc.edu.cn
University of Science and Technology of China, Hefei, China

Junyu Lu
University of Science and Technology of China, Hefei, China

Meikai Bao
University of Science and Technology of China, Hefei, China

Xukai Liu
University of Science and Technology of China, Hefei, China

Linbo Zhu
University of Science and Technology of China, Hefei, China

Enhong Chen
University of Science and Technology of China, Hefei, China

Abstract

The rapid advancement of artificial intelligence and virtual digital human technologies is ushering in transformative changes in education. Virtual digital humans, empowered by multimodal information fusion and natural language processing capabilities, offer a more engaging and humanized interaction experience for online learning. However, general-purpose AI models for digital teaching often face notable challenges: 1) They often lack deep understanding of users, overlooking personalized learning needs; 2) they may struggle with knowledge reasoning, leading to unreliable and outdated content; 3) their human-machine interactions can feel unnatural, lacking emotional resonance and empathetic support.

To address these challenges, we developed an interaction platform focused on delivering personalized digital human services. This platform optimizes learner profiling, domain-specific data integration, and service frameworks to enhance the intelligence, personalization, and quality of digital teaching agents, effectively tackling issues related to shallow user understanding and unreliable reasoning. Through the integration of realistic digital human avatars and personalized interactive technology, the platform enables more natural and seamless human-machine interactions. Additionally, emotional speech synthesis and synchronized facial expression technology allow virtual teachers' facial expressions to dynamically align with their speech, enhancing expressiveness, vitality, and empathy, which together heighten users' sense of immersion and interaction quality. This development marks a significant step toward smarter, more personalized education.

CCS Concepts

• **Information systems** → **Chat**; **Multimedia streaming**.

WWW Companion '25, Sydney, NSW, Australia

ACM ISBN 979-8-4007-1331-6/25/04
https://doi.org/10.1145/3701716.3715182

Keywords

Virtual Digital Human, Online Education, Interactive Platform, Empathetic Interaction

ACM Reference Format:
Kai Zhang, Weiyin Gong, Qi Liu✉, Junyu Lu, Meikai Bao, Xukai Liu, Linbo Zhu, and Enhong Chen. 2025. Intelligent Interaction Platform for Personalized Digital Tutor: Advancing Empathetic and Adaptive Learning Experiences. In *Companion Proceedings of the ACM Web Conference 2025 (WWW Companion '25), April 28-May 2, 2025, Sydney, NSW, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3701716.3715182

1 Introduction

With the rapid development of large language models (LLM)[3, 8], artificial general intelligence (AGI), and digital technologies, digital humans, as a representative of intelligent education, are gradually becoming a rising force in the education domain[9, 11]. As a representation of intelligent interaction technology, virtual digital humans combine visual expressiveness with interactive intelligence, offering users a human-like service experience. Their applications go beyond replacing traditional interfaces, offering innovative ways to empower industries with intelligent technologies [6, 11].

In the field of education, virtual digital humans are regarded as an important component of future education models.[1, 2, 5]. Traditional education models mainly rely on two-way interaction between teachers and students. However, with the integration of virtual digital humans into the field of education, the teaching model is gradually shifting toward an interactive, three-way collaboration among teachers, machines, and students. Virtual digital humans can not only serve as "digital teachers" in the teaching process, helping students understand knowledge more effectively, but also provide real-time personalized feedback and guidance. Through this interactive approach, digital teachers offer a more immersive and interactive learning experience for online education, allowing students to receive guidance and support in virtual classrooms that is closer to that of traditional in-person teaching [6].

Although virtual digital humans possess certain interactive capabilities, current digital teachers often struggle to accurately understand the needs of learners [5, 6]. Especially when facing open-ended conversations, digital teachers find it difficult to accurately

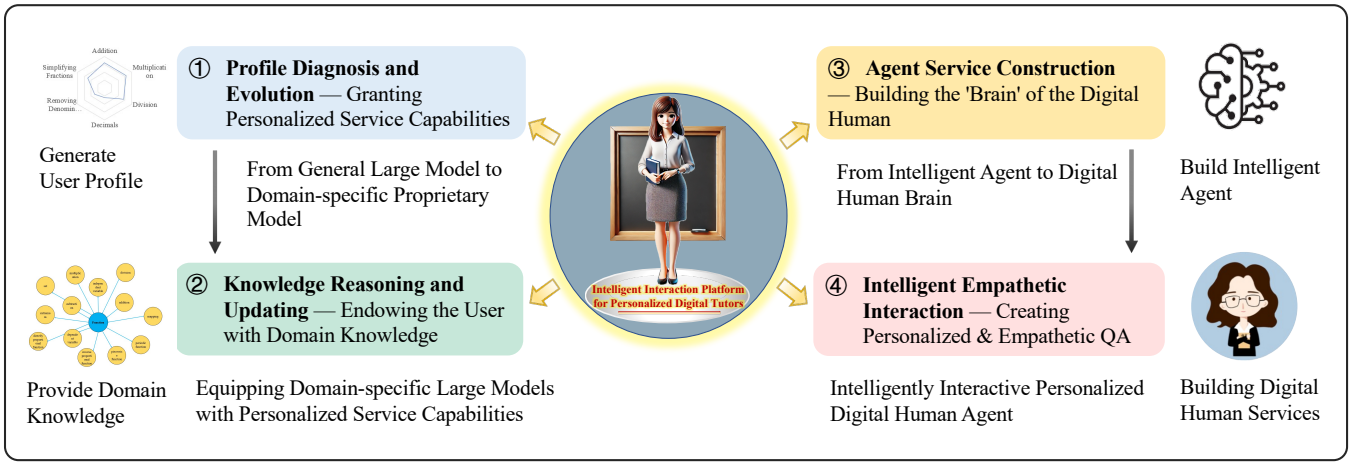

Figure 1: Intelligent Interactive Platform for Personalized Digital Tutors.

extract user intentions and learning preferences from the dialogue, leading to discrepancies in their understanding of the learner's cognitive state. Moreover, existing large models also face difficulties in aligning with learners' knowledge levels. As personalized long-term memory functions have not yet been fully developed, digital teachers also face shortcomings in tracking students' learning trajectories and summarizing learning preferences over time, thus limiting the effectiveness of personalized teaching.

Despite the remarkable achievements of large language model agents, they still face a series of challenges in practical applications, including the persistence of hallucination phenomena, limitations in understanding due to single-modal communication, and the lack of in-depth responses to personalized needs despite their broad generalization abilities[3, 8, 10]. Due to the lack of a validation mechanism, the automatically generated knowledge content sometimes contains inaccurate information, which is particularly disadvantageous in educational scenarios. Additionally, the insufficiently timely updates to the knowledge base can result in some generated content containing outdated, noisy, or incorrect information. Therefore, ensuring the accuracy and reliability of knowledge content has become one of the key issues in virtual digital human platforms. Additionally, virtual digital human technology also has limitations in interaction naturalness, particularly in emotional recognition and contextual understanding. Compared to real teachers, virtual digital humans lack emotional perception and expression capabilities, making it difficult to provide emotional support and empathetic care during interactions. In addition, current virtual digital humans often experience issues such as context switching and contradictory responses when handling continuous conversations, resulting in a lack of natural and smooth dialogue, which affects the user's sense of immersion and learning satisfaction.

To address these challenges, we have developed a virtual digital tutor platform[1] with intelligent, personalized, and high-quality interactive capabilities. On this platform, through multimodal data fusion and the creation of user profiles, we achieve a deep understanding of learners' behaviors and knowledge levels, supporting personalized teaching. We have designed a knowledge validation and update mechanism to ensure the accuracy of knowledge generation by the virtual digital human and dynamically adjust the knowledge base content through adaptive learning. Additionally, an emotional recognition and expression module has been integrated into the digital teacher system, allowing the virtual digital human to capture user emotions in real-time and provide a more human-like interactive experience. Through the implementation of these modules, we have created an intelligent virtual digital teacher open platform tailored for the education field, which can also provide support for multiple industries such as healthcare in the future.

2 Personalized Digital Tutor Platform

In this section, we introduce the intelligent interactive platform for personalized digital humans, as shown in Figure 1. Building on previous research, this platform focuses on learner ***profile diagnosis and evolution***, subject-specific multimodal ***knowledge reasoning and updating***, the integration of subject knowledge and learner profiles through large-model intelligent ***agent service constructions***, and ***intelligent empathetic interaction***.

2.1 Profile Diagnosis and Evolution

In response to changes in cognitive characteristics at different stages, we diagnose the user's cognitive level by acquiring long-term interaction records, mining behavioral pattern data, and performing knowledge tracking. This enables cross-stage alignment of user profiles and addresses the challenge of the intelligent agent's shallow understanding of the learner. We have designed profile diagnosis and evolution, as shown in Figure 2. This section dynamically updates the learner profile, optimizes learning path planning, and recommends personalized teaching content, thereby improving teaching effectiveness.

Specifically, through quantitative cognitive-level diagnosis based on structured dialogues, we complete diversified cognitive diagnostics. By combining cognitive diagnostic models[4, 7], we comprehensively analyze both the explicit and implicit cognitive traits of learners, accurately depicting their mastery and cognitive differences across different knowledge points. We dynamically track the

[1]https://github.com/BAOOOOOM/digital_human

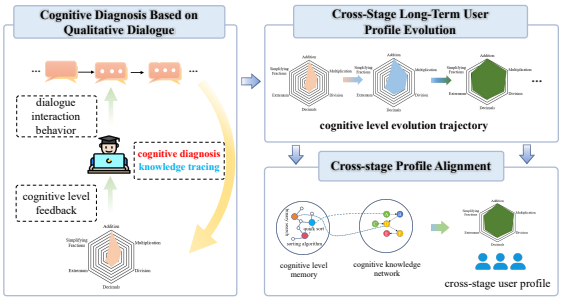

Figure 2: Profile Diagnosis and Evolution.

learner's cognitive state and knowledge acquisition level and utilize memory mechanisms to continuously update and store cognitive states, thereby forming a historical trajectory of cognitive evolution that provides data support for predicting future learning trends.

To better understand the learner's cognitive state, we continuously record the learner's interaction behaviors, learning outcomes, and changes in cognitive status over different time periods. By analyzing the dynamic changes in learning behavior patterns and knowledge acquisition, we uncover the cognitive development patterns and stage characteristics of the learner throughout the long-term learning process. This improves the long-term learning efficiency of adaptive learning systems and enhances the intelligent intervention capabilities of digital teachers. Furthermore, by analyzing the relationship between cognition and knowledge through cross-stage profile alignment, we map the learner's cognitive performance on different knowledge points to a knowledge graph. This constructs a dynamic linkage between knowledge and cognition, enabling the visualization of multidimensional learner profiles.

It clearly shows the learner's cognitive development path and knowledge mastery status. Profile diagnosis and evolution not only effectively solve the issue of the intelligent agent's shallow understanding of the learner but also provide strong support for personalized recommendations and learning path optimization.

2.2 Knowledge Reasoning and Updating

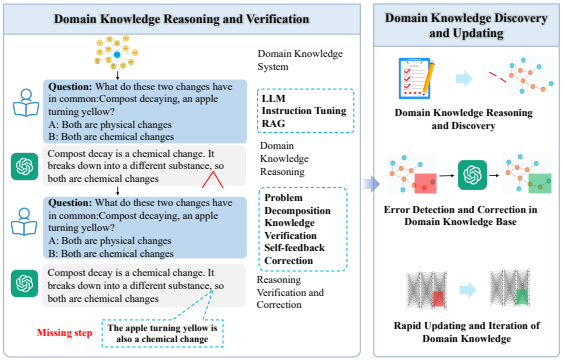

Figure 3: Knowledge-reasoning and Updating.

In the process of knowledge inference, the results may lack accuracy due to outdated knowledge bases or logical errors in the inference process, leading to untrustworthy knowledge inference. To address this issue, an innovative approach is proposed that combines domain knowledge inference and verification, enhancing the reliability of inference results through continuous discovery and updating of domain knowledge, as shown in Figure 3.

This approach constructs a high-quality disciplinary knowledge system, providing reliable support for the reasoning and services of digital teacher agents. By integrating large language models, instruction fine-tuning, and retrieval-augmented generation, it performs domain knowledge inference on the disciplinary knowledge system, breaking down complex disciplinary problems into subproblems to reduce inference complexity and improve accuracy. Existing knowledge graphs or standard knowledge bases are used to verify inference results, identifying potential logical gaps or knowledge inconsistencies. A self-feedback correction mechanism is also introduced to automatically adjust detected inference deviations while continuously monitoring external data sources and new developments in the field to dynamically add new knowledge or update existing entries. This dynamic maintenance mechanism enables the knowledge base to evolve continuously, staying responsive to new knowledge, and enhancing the inference reliability and service quality of the digital teacher agent.

2.3 Agent Service Construction

Currently, intelligent agents based on large language models, with their deep neural network architecture, have demonstrated exceptional capabilities in natural language understanding and complex reasoning tasks and are widely applied in fields such as education, justice, and healthcare. However, these intelligent agents still face several challenges in practical applications, such as the difficulty in eliminating the "hallucination" phenomenon, the limitations of single-modal understanding, and insufficient responses to personalized needs. To enhance the effectiveness of large model-based intelligent agents in personalized teaching, this platform integrates a multi-turn dialogue intelligent service framework, combining domain knowledge and learner profiles to support intelligent agents in making precise decisions during the teaching process.

First, through knowledge-integration pre-training techniques and by combining deep learning, transfer learning, and adaptive algorithms, the platform integrates general knowledge and domain-specific knowledge into the large model to enhance its teaching reasoning capabilities. Additionally, we propose a knowledge-data-driven personalized decision-making framework, which combines student data and subject knowledge to support accurate personalized recommendations from intelligent agents during the teaching process. Furthermore, to enhance the interactivity and adaptability of the digital teacher, memory, planning, and reflection modules are employed, enabling the intelligent agent to engage in multi-turn dialogues and continuously optimize teaching strategies.

The platform allows the digital teacher to dynamically adjust teaching content based on feedback, achieving personalized teaching while improving student engagement and learning efficiency.

2.4 Intelligent Empathetic Interaction

To address the issue of unnatural human-computer interaction, we employ emotional cognition technology to perform real-time analysis and assessment of learners' emotions, as shown in Figure 4. First, the digital teacher's intelligent interaction capabilities

are enhanced by optimizing the learning experience from multiple dimensions. Emotional cognition technology enables real-time analysis of the learners' emotional states during the learning process, helping the digital teacher accurately detect changes in their emotions. By identifying both verbal and non-verbal signals from learners, the system precisely understands their emotional state and provides personalized feedback tailored to their emotional needs.

Additionally, to strengthen guidance and effectiveness in teaching, the guided interactive teaching technique allows the digital teacher to emulate a real teacher's ability to lead. Specifically, the system can intelligently analyze a learner's current comprehension level, learning progress, and knowledge acquisition, especially identifying areas of confusion. For these areas, the digital teacher poses progressively guiding questions to help learners gradually understand and master the relevant knowledge points, thereby enhancing their engagement in the learning process and boosting their motivation for self-directed learning.

Moreover, the platform integrates digital human visualization, emotional speech synthesis, and synchronized facial expression technologies, allowing the virtual teacher's facial expressions to change in sync with the audio. This significantly improves the expressiveness and vitality of the digital teacher, enabling learners to experience more authentic emotional feedback and visual interaction. The UI interface of the platform is shown in Figure 5, which integrates digital human visualization, emotional speech synthesis, and synchronized facial expressions. The virtual teacher's expressions and voice emotions sync with the audio, offering a more engaging, lifelike learning experience with authentic emotional feedback and enhanced visual interaction.

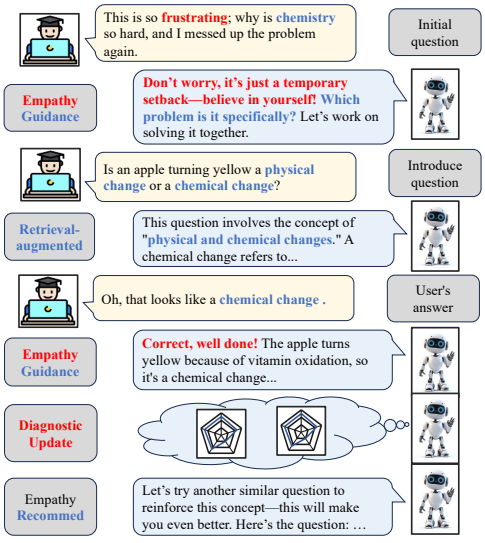

Figure 4: Intelligent Empathetic Interaction.

3 Conclusion

This paper presents an intelligent interaction platform tailored for a personalized digital tutor. The platform utilizes long-term user interaction records to assess cognitive levels, mine behavioral patterns, and perform knowledge tracking, achieving cross-stage user profile alignment and addressing the agent's insufficient understanding of learners. By integrating domain knowledge inference and validation methods, the platform enhances the reliability of inference results through continuous discovery and updates of domain knowledge. It monitors external data sources and advancements in academic fields to dynamically add new knowledge, ensuring the knowledge base's timeliness and accuracy. Finally, by conducting real-time analysis of learners' emotions, the platform provides personalized services and feedback. This implementation marks a significant step toward more intelligent and personalized online education.

Acknowledgements. This research was partially supported by the NSFC (No.62406303), the Key Technologies R & D Program of Anhui Province (No.202423k09020039,202423k09020010), Anhui Provincial Natural Science Foundation (No.2308085QF229) and the Fundamental Research Funds for the Central Universities.

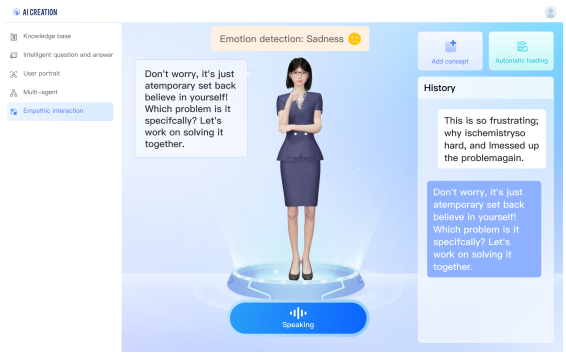

Figure 5: UI for Personalized Digital Tutor Platform.

References

[1] Shruti Agrawal and Satheesh Abimannan. 2023. Conversational virtual educator using Google's dialogflow. In *AIP Conference Proceedings*, Vol. 2755. AIP Publishing.

[2] Maria V Arkhipova, Natalia V Shutova, Olga A Orlova, Oksana R Zhernovaya, Yulia N Zintsova, and Irina M Udalova. 2024. Digital transformations in education: Augmented reality in the classroom. In *AIP Conference Proceedings*, Vol. 2969. AIP Publishing.

[3] Callie Y Kim, Christine P Lee, and Bilge Mutlu. 2024. Understanding large-language model (llm)-powered human-robot interaction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 371–380.

[4] Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. 2019. Ekt: Exercise-aware knowledge tracing for student performance prediction. *IEEE Transactions on Knowledge and Data Engineering* 33, 1 (2019), 100–115.

[5] Mei Qian, Qunqiang Chen, Rujuan Huang, Hao Zheng, and Chen Chen. 2024. Classroom AI question-answering system based on machine reading comprehension model. In *International Conference on Automation Control, Algorithm, and Intelligent Bionics (ACAIB 2024)*, Vol. 13259. SPIE, 255–260.

[6] Yanan Sun, Zhiyao Sun, Yu-Hui Wen, Sheng Ye, Tian Lv, Minjing Yu, Ran Yi, Lin Gao, and Yong-Jin Liu. 2023. Generation of virtual digital human for customer service industry. *Computers & Graphics* 115 (2023), 359–370.

[7] Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yu Yin, Shijin Wang, and Yu Su. 2022. NeuralCD: a general framework for cognitive diagnosis. *IEEE Transactions on Knowledge and Data Engineering* 35, 8 (2022), 8312–8327.

[8] Jiaqi Wang, Zihao Wu, Yiwei Li, Hanqi Jiang, Peng Shu, Enze Shi, Huawen Hu, Chong Ma, Yiheng Liu, Xuhui Wang, et al. 2024. Large language models for robotics: Opportunities, challenges, and perspectives. *arXiv preprint arXiv:2401.04334* (2024).

[9] Xiaodong Wei and Haixuan Yan. 2024. Research on Teaching Activities in University Blended Technology Development Courses with Support from Generative Virtual Digital Humans. In *2024 International Symposium on Educational Technology (ISET)*. IEEE, 230–234.

[10] Bowen Zhang and Harold Soh. 2023. Large language models as zero-shot human models for human-robot interaction. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 7961–7968.

[11] Rui Zhen, Wenchao Song, Qiang He, Juan Cao, Lei Shi, and Jia Luo. 2023. Human-computer interaction system: A survey of talking-head generation. *Electronics* 12, 1 (2023), 218.