LAF
https://doi.org/10.15302/J-LAF-2026-0031
Volume 14 / Issue 3 / 260031 / 2026
Received Date | 2025-06-03 Accepted Date | 2026-01-12

Research on the Development of a Large Language Model-Based Artificial Intelligence Teaching Assistant and Discipline-Based Knowledge Engine for Urban Planning: A Case Study of the New Science of Cities

Tian ZHENG, Junhao XIA, Yue MA, Xiaolan HUANG, Ying LONG*

School of Architecture, Tsinghua University, Beijing 100084, China

*CORRESPONDING AUTHOR
Address: No. 30 Shuangqing Road, Haidian District, Beijing 100084, China
Email: ylong@tsinghua.edu.cn

ABSTRACT

The rapid development of artificial intelligence (AI) is driving a profound transformation in higher education, shifting conventional teaching models toward intelligent instructional paradigms. This study explores the development of AI teaching assistant systems, their application in educational practice, and the construction of a discipline-based knowledge engine. The results demonstrate that, through the integration of domain knowledge injection, retrieval-augmented generation mechanisms, and optimized text-to-image generation capabilities, AI teaching assistants not only significantly improve the accuracy of professional knowledge question answering, but also exhibit high expressive precision and strong adaptability to Chinese local urban contexts in urban spatial scene image generation tasks. Building upon these findings, this study further proposes a pathway for constructing discipline-based knowledge engines grounded in large language models and deeply customized for disciplinary knowledge systems. A system architecture integrating text generation, knowledge graphs, and collaborative text-to-image generation is developed to facilitate the evolution from general-purpose language models to discipline-oriented intelligent engines. This research provides a replicable and transferable theoretical framework and technical pathway for the digital transformation of higher education, while also expanding the application scenarios and future directions of interdisciplinary research at the intersection of AI and education.

KEYWORDS

Artificial Intelligence; AI Teaching Assistant; Large Language Model; Retrieval-Augmented Generation; Text-to-Image Generation; Teaching Practice; Discipline-Based Knowledge Engine

HIGHLIGHTS

- Developed an AI teaching assistant system with retrieval-augmented generation and text-to-image generation capabilities
- Proposed a discipline-based knowledge engine framework and technical pathway for urban planning education
- Achieved the collaborative application of large language models in instructional interaction, spatial representation, and course practice

RESEARCH FUNDS

- The Construction of a Full-Process Digital and Intelligent Empowerment Case Database for Urban and Rural Planning and Design Courses Based on the Ministry of Education Virtual Teaching and Research Office, Undergraduate Education and Teaching Reform Project, Tsinghua University (No. DX02_34)
- Exploratory Project on the Application of Artificial Intelligence in Education and Teaching for the New Science of Cities, Undergraduate Education and Teaching Reform Project, Tsinghua University (No. DX08_05)

1 Introduction

The Fourth Industrial Revolution has introduced disruptive technologies represented by artificial intelligence (AI), driving continuous transformation in scientific research paradigms and teaching models. With rapid technological advancement, researchers are now able to integrate heterogeneous multi-source data such as text, images, and geospatial information to analyze urban issues in more fine-grained, dynamic, and visualized ways[1–4]. At the same time, higher education is undergoing profound structural adjustment and systemic reconfiguration.

Currently, higher education faces several challenges, including the uneven distribution of high-quality educational resources across regions, disciplines, and institutions, as well as the limitations of the conventional instructional model characterized by teacher-centered approaches, standardized curricula, and fixed learning paces in addressing students' diverse and individualized learning needs[5–7]. However, the pathways, mechanisms, key nodes, and underlying logic of deep transformation driven by technological advancement still require systematic clarification and empirical validation. Existing studies indicate that AI demonstrates significant potential in areas such as instructional content generation, knowledge graph construction, and human–machine collaborative learning. It is expected to reshape modes of knowledge acquisition and drive educational ecosystems toward greater intelligence, openness, and adaptability[8–12]. Recent studies have also validated the effectiveness of AI in applications such as virtual teaching assistants, writing support, lesson plan development, and assignment grading[13–15]. However, research on the system architecture, functional design, and integration of AI teaching assistants into instructional contexts remains in its early stages, lacking a mature theoretical framework and well-established practical paradigms. From an educational theory perspective, if AI teaching assistants are integrated with constructivist learning theory[16] and cognitive load theory[17], this may help explain the mechanisms through which they promote personalized learning and deeper cognitive engagement, thereby providing a theoretical foundation for their adoption in higher education.

This study takes the New Science of Cities, an undergraduate general education course at Tsinghua University, as a representative case. It aims to investigate the key mechanisms underlying the evolution of the AI teaching assistant from general-purpose language models to domain-specific intelligent engines, and to explore innovative pathways for integrating AI into teaching practice. The study seeks to provide a replicable and transferable theoretical framework and operational paradigm for promoting the deep integration of AI technologies and teaching practice in higher education.

2 Methods and Technical Implementation of the AI Teaching Assistant

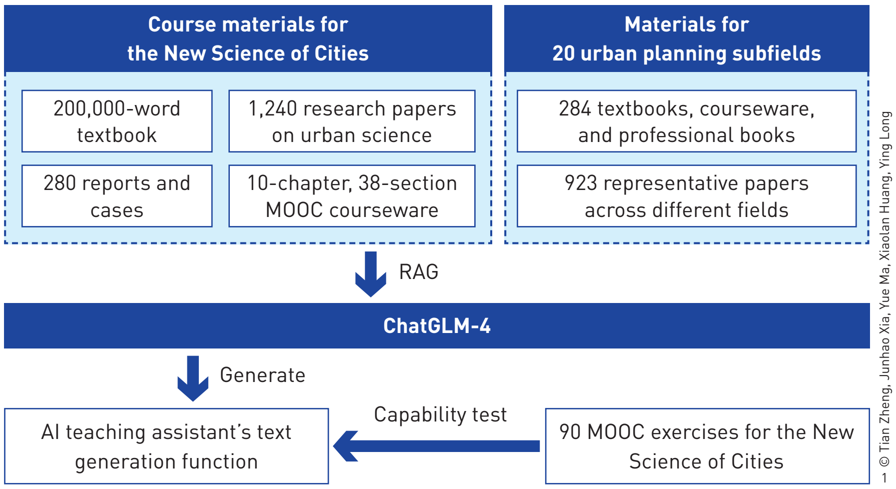
Current AI teaching systems in text-based interaction largely rely on keyword matching or outputs generated by a general-purpose large language model (LLM). As a result, their responses often lack in-depth understanding of domain-specific knowledge and the latest research developments. In addition, most existing systems either do not support image generation capabilities or rely solely on basic generative models, making it difficult to meet the expressive and cognitive requirements of professional teaching. Building upon existing technologies, the AI teaching assistant for the New Science of Cities is specifically adapted to course requirements and develops two core functions: text generation and image generation.

The text generation module enhances instructional relevance and domain alignment in course-based question answering by constructing a high-quality domain-specific corpus and incorporating a Retrieval-Augmented Generation (RAG) mechanism. The image generation module, on the other hand, leverages a large-scale image-text paired dataset and model fine-tuning to improve visual representation and cognitive support in the teaching context of future cities. Furthermore, these two core functions are integrated into the Zhipu Qingyan AI platform and the Hetang RainClassroom system. A functional card interface is introduced to provide prompt templates based on typical task scenarios, thereby improving usability and pedagogical applicability. Compared with general-purpose models (e.g., ChatGPT, Midjourney), the customized AI teaching assistant requires greater investment in corpus construction, model fine-tuning, and platform integration. However, it demonstrates significant advantages in instructional task alignment, response accuracy, and cultural contextualization.

2.1 Text Generation Function

In terms of text generation, the research team first focused on the systematic organization and construction of high-quality domain-specific data. The overall workflow is illustrated in Fig. 1, aiming to provide knowledge support with both disciplinary characteristics and educational value for course instruction. The textual data include textbooks, online course materials, seminal papers, and case-based literature. By constructing a comprehensive corpus based on the New Science of Cities and urban planning literature, the model's capabilities in knowledge understanding and contextual adaptation for teaching and academic research are significantly improved.

In terms of model capability optimization, the AI teaching assistant adopts ChatGLM-4[18] as the foundational LLM and incorporates RAG to achieve domain-specific enhancement for teaching scenarios. Compared with traditional fine-tuning approaches, RAG does not require updating model parameters



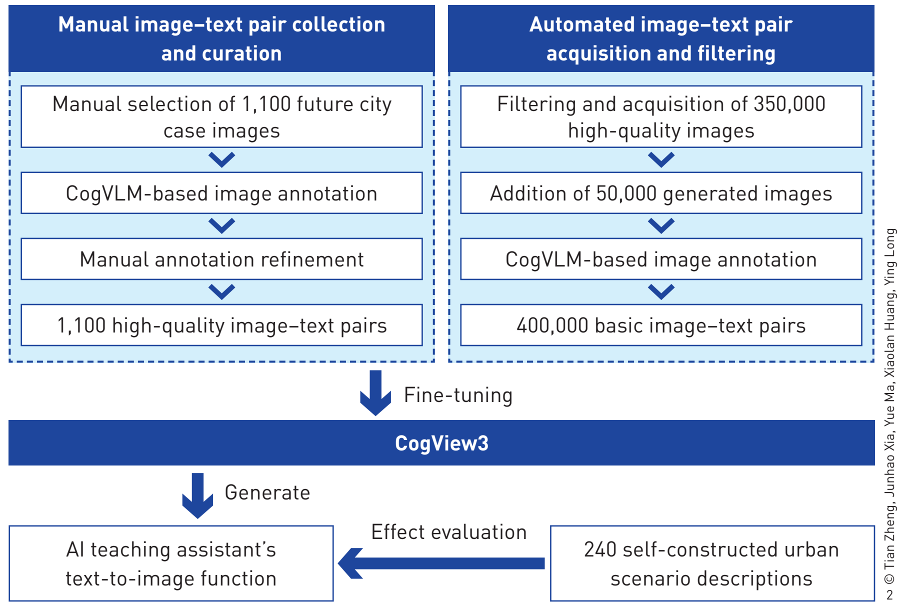


Fig. 1 Technical roadmap for text generation function development.
Fig. 2 Technical roadmap for text-to-image function development.

and therefore offers advantages such as faster iteration and more convenient knowledge updating. It is particularly suitable for domains where knowledge evolves rapidly or is frequently updated. While ensuring domain specificity, this approach also effectively mitigates the hallucination problem of LLM. During the response generation stage, prompt engineering is applied to explicitly require the LLM to adopt course-specific analytical perspectives and frameworks when structuring its responses, thereby guiding the AI teaching assistant to more fully and systematically utilize the retrieved corpus in generating answers.

2.2 Text-to-Image Generation Function

In terms of the image generation capability, to enhance the AI teaching assistant's support for visually representing urban scenarios, the research team fine-tuned the model on a self-constructed high-quality image-text paired dataset, improving its text-to-image generation performance (Fig. 2). The dataset consists of both manually curated and automatically collected components. The manual process focuses on representative cases of future urban spatial design, where 1,100 representative images were carefully selected and annotated by researchers with a background in design. For the automated component, 350,000 high-resolution real-world images were collected from international architectural design platforms, and an additional 50,000 high-quality AI-generated images were incorporated. To support annotation at scale, CogVLM[18], a vision-language model, was employed to automatically generate image descriptions. Ultimately, a dataset of 400,000 image-text pairs was constructed, covering multiple urban scenario categories such as residential spaces, office spaces, and public spaces, providing rich training samples for model development.

Based on the above dataset, the research team fine-tuned CogView3[18] as the backbone model, employing Low-Rank Adaptation (LoRA) to optimize key parameters, and conducted approximately 2,000 training steps. To further enhance the model's ability to understand prompts, an LLM was introduced prior to image generation to rewrite and expand user inputs. In this process, concise keywords containing domain knowledge were transformed into semantically enriched sentences with explicit spatial composition and complete scene descriptions, thereby improving both visual quality and semantic accuracy of generated images.

In the evaluation stage, 240 test samples were constructed across 30 categories of representative urban spatial scenarios (with 8 descriptions per category), covering both classical and future-oriented as well as indoor and outdoor environments. The results show that the latest CogView3-Plus model[18] achieved a user satisfaction rate of approximately 60% on the test set, whereas the fine-tuned model based on CogView3 performed better in terms of image content accuracy, visual style alignment, and cultural contextual representation (Fig. 3). Notably, it demonstrates a significant improvement in the representation of Chinese urban contexts, achieving a user satisfaction rate of 75%, which confirms its practical value in supporting creative expression.

2.3 Interface Development

The research team integrated the two core functions—text generation and image generation—into the Zhipu Qingyan AI

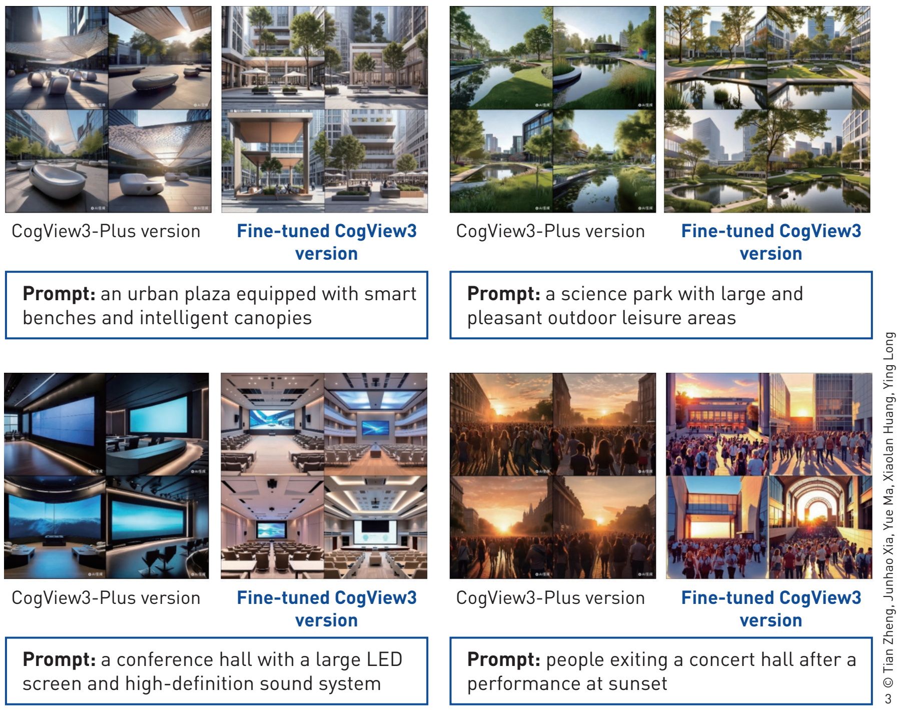
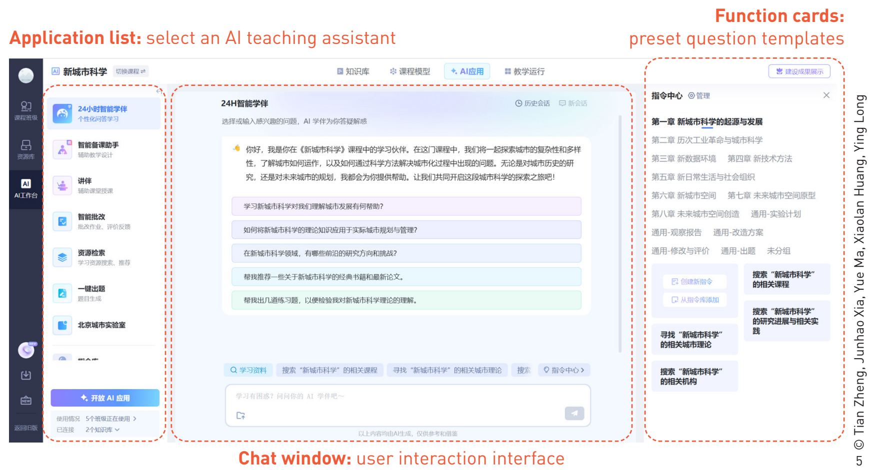


Fig. 3 Comparison between CogView3-Plus and fine-tuned CogView3 models.

platform and the Hetang RainClassroom system, enabling the application of AI in real-world teaching scenarios. The former served as a testing platform for students and was subsequently integrated into the latter after functional refinement, thereby better connecting existing course resources and improving the overall learning experience.

To improve students' operational efficiency and targeted questioning, the team designed and deployed 25 functional card modules on the Zhipu Qingyan AI platform, covering five application scenarios of experimental planning, observational reporting, transformation proposals, revision and evaluation, and question generation (Fig. 4). These functional cards received positive feedback from students and were subsequently migrated to the Hetang RainClassroom system to further extend their applicability in real teaching contexts. Based on teaching experience, the team systematically identified key and difficult points in the course, developed customized question templates, and embedded them into the system in the form of functional cards. This enables students to directly access structured and intelligent responses after completing each course unit (Fig. 5). In addition, by integrating knowledge graphs, MOOC resources, and course materials, the AI teaching assistant system establishes a multi-source learning resource linkage, which further supports students in achieving a deeper understanding of the course content (Fig. 6).



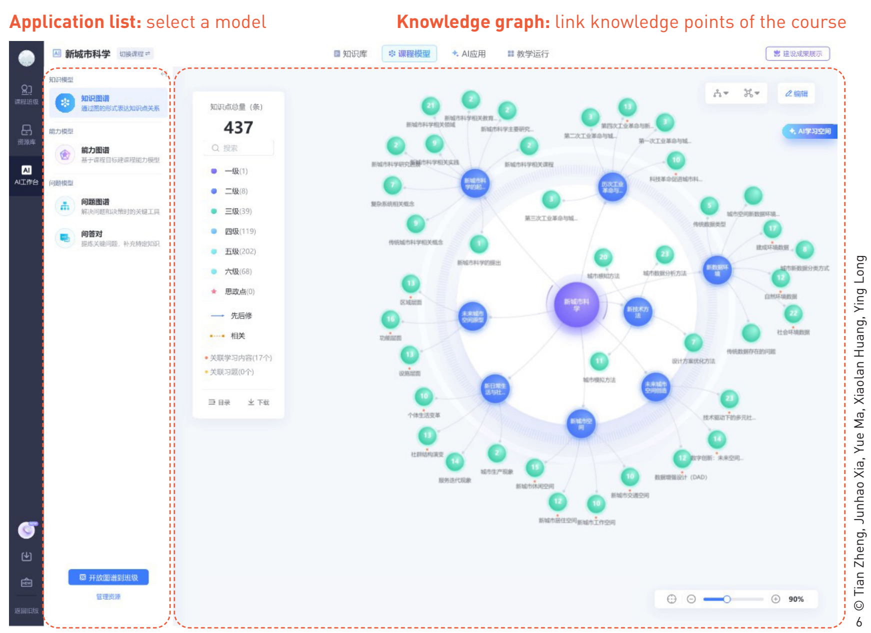


Fig. 4 Schematic illustration of functional card usage on the Zhipu Qingyan AI platform.
Fig. 5 Schematic illustration of functional card usage on the Hetang RainClassroom.
Fig. 6 Knowledge graph interface of the Hetang RainClassroom.

3 Application of the AI Teaching Assistant in Teaching Practice

3.1 Intelligent Classroom Interaction

In classroom teaching, instructors can use the AI teaching assistant to generate introductory content aligned with the course progression, combining academic rigor with pedagogical relevance to stimulate student interest. During lectures, the AI teaching assistant acts as a “third instructional agent,” participating in teacher–student interaction by providing exploratory and heuristic responses, thereby supporting instructors in guiding critical thinking. In discussion sessions, students can leverage the AI teaching assistant to clarify discussion directions and obtain supplementary knowledge and case materials, thereby enhancing both the quality and depth of classroom discussions (Fig. 7).

3.2 Post-Class Tutoring Support

In the post-class stage, the AI teaching assistant provides multidimensional support for knowledge consolidation and assignment assistance. For content not fully understood during class, students can quickly raise questions via functional cards or access a 24-hour intelligent learning companion for explanations (Fig. 8). For extended learning needs, relevant content can be retrieved from the knowledge base for further exploration. In addition, the system supports multiple interaction modalities, including voice-based interaction, screenshot-based Q&A, and voice-enabled mind map generation.

The AI teaching assistant is integrated throughout all stages of course projects. It assists students in constructing research frameworks, organizing ideas, and recommending appropriate research methodologies. By leveraging academic resources, the system provides analytical suggestions that help optimize logical structures, while also efficiently filtering relevant literature and cases to enhance information synthesis capabilities. In terms of visual representation, the image generation function supports the depiction of future scenarios, thereby strengthening the expressiveness of research outcomes. Finally, the system can further refine and polish written text, improving linguistic accuracy and academic expression quality.

3.3 Student and Instructor Evaluation

To evaluate the practical effectiveness of the AI teaching assistant, the research team analyzed student usage data from the Fall semesters of 2023 and 2024, and collected student feedback through questionnaires.

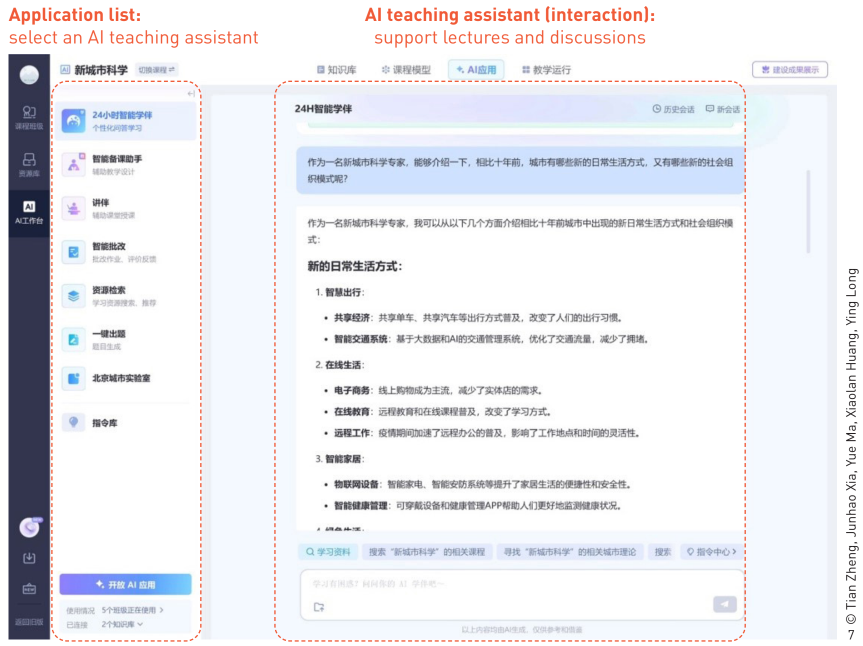


Fig. 7 Classroom Q&A with AI teaching assistant.
Fig. 8 24-hour intelligent learning companion.

In terms of usage, 30 students participated in the feedback survey in Fall 2023, and approximately 70% reported using the AI teaching assistant in tasks such as information retrieval, concept understanding, research framework construction, and text refinement. In Fall 2024, all 17 students used ChatGLM, the underlying LLM, for assistance. The total number of interactions reached 432, with a maximum of 108 queries submitted by a single user and a maximum usage duration of 24 h 11 min (Fig. 9). Students

Fig. 9 Statistical overview of students' usage of Zhipu Qingyan in the fall semester 2024.

with different learning styles exhibited significant differences in their usage patterns. Students who tended toward independent thinking primarily used the system in a critical and verification-oriented manner, focusing on information retrieval and rapid validation, and thus showed relatively shorter usage durations. In contrast, students who preferred interactive learning were more inclined to use the system as a conversational cognitive tool, engaging in continuous dialogue to expand cognitive boundaries and deepen understanding, resulting in longer usage times. In addition, students in the top 20% of usage duration generally achieved "excellent" course grades, indicating a potential positive correlation between AI teaching assistant usage intensity and learning performance. However, this relationship may also be influenced by factors such as overall learning engagement, and therefore requires further validation.

Meanwhile, students' feedback indicates that the AI teaching assistant plays a significant role in improving research thinking and writing quality. However, several limitations remain, including insufficient depth of responses, unstable image generation quality, slow response speed, and incomplete literature retrieval functionality. Some students suggested enhancing the system's understanding of urban space and campus environments, improving contextual memory and update capabilities, reducing templated expressions, and strengthening cross-domain adaptability. In addition, existing studies have shown that excessive reliance on intelligent tools may weaken learners' autonomous learning and problem-solving abilities[19]. Therefore, in teaching practice, it is necessary to clearly define the auxiliary role of AI teaching assistants and maintain students' critical thinking and reflective capabilities through appropriate pedagogical guidance. Future development of the AI teaching assistant should not only improve the depth and flexibility of responses at the technical level, but also enhance interpretability and transparency at the design level, thereby developing students' ability to effectively interact with AI rather than becoming dependent on it.

From the instructors' perspective, this practice involves both the construction and application of the AI teaching assistant. Instructors serve not only as system designers and builders, but also as users and observers in the teaching process. During the system development, key challenges include accurately identifying instructional requirements, systematically constructing the knowledge base, and ensuring deep alignment between the system and course content. In practical application, several key operational considerations are highlighted. First, the AI teaching assistant should be introduced appropriately during classroom teaching, emphasizing its supportive rather than substitutive role, so as to stimulate student interest and enhance classroom engagement. Second, in the post-class stage, the system's knowledge-based question-answering capability should be fully utilized, allowing it to handle routine and repetitive inquiries, thereby enabling instructors to focus more on guiding students' systematic thinking. Third, for individualized questions from students, instructors should integrate AI-generated feedback with professional expertise to provide more in-depth analysis and guidance.

4 Exploration of AI-Enabled Discipline-Based Knowledge Engine Construction

4.1 From AI Teaching Assistant to Domain-Specific Vertical Foundation Model

Tsinghua University has taken the lead in exploring AI-enabled disciplinary engine construction based on its experience in online education development and AI-supported teaching practices, aiming to promote the transformation of higher education paradigms. The AI teaching assistant for the New Science of Cities can serve as a starting point for gradually advancing toward the construction of a disciplinary engine in the field of urban planning. The technical architecture of the disciplinary engine consists of four main layers. The knowledge graph layer organizes multi-source knowledge in the field of urban planning—including core concepts, methods, and cases—into a structured graph format, enabling logical reasoning and cross-domain retrieval. The multimodal fusion layer integrates text, images, and tabular data, and through cross-modal alignment and representation learning, enables the model to jointly understand and process visual, statistical, and textual information. The model evolution layer builds upon general-purpose LLM and incorporates RAG, gradually evolving into a domain-specific vertical foundation model for urban planning through domain corpus fine-tuning, knowledge injection, and human-in-the-loop feedback. The application interface layer supports diverse application scenarios, including personalized theoretical learning and planning design guidance. Through this architecture, the AI teaching assistant is no longer limited to a supportive instructional tool, but is further developed into a system with generalized reasoning and domain-specific intelligence capabilities in urban planning.

Specifically, the team further expanded the input content and application scope of the AI teaching assistant (Fig. 10). Building upon the original AI assistant model, 20 subfields of urban planning were identified based on the undergraduate training program in urban planning. A comprehensive corpus covering the New Science of Cities and urban planning disciplines with broad knowledge scope was constructed, providing a solid domain-specific knowledge foundation for the LLM. Meanwhile, 49 sets of real and mock tests for certified urban–rural planner examination were collected for model evaluation. The results show that the domain-specific vertical model outperformed other general-purpose models on the corresponding test tasks (Table 1).

Fig. 10 Construction of a domain-specific vertical foundation model.

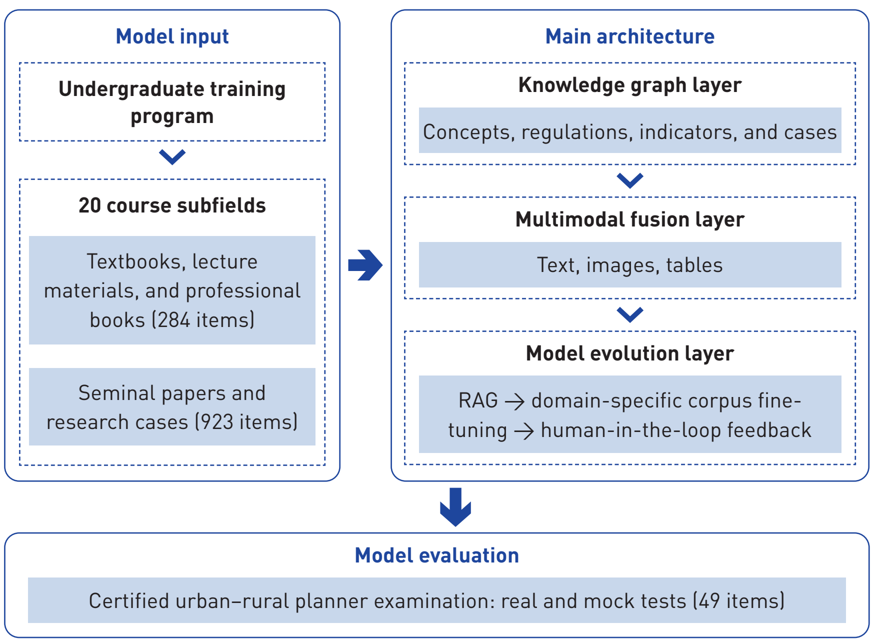


Table 1: Performance comparison between the domain-specific vertical model and general-purpose LLM

LLM	Multiple-choice question (933 items)	True/false question (217 items)
ChatGLM-4 + specialized knowledge	77.4%	72.6%
ChatGLM-4	75.0%	71.3%
GPT-4o	70.7%	72.2%
Claude 3.5 Sonnet	70.2%	70.0%
Gemini-Pro	54.7%	58.8%

NOTES

1. The percentages represent the matching accuracy between model responses and the standard answers.
2. The base model ChatGLM-4 used in this study was released in January 2024. In comparison, large models such as DeepSeek-R1 and Qwen2.5-72B were released in the second half of 2024, after the construction stage of the domain-specific vertical model in this study. Therefore, they were not included in the performance comparison.

4.2 Future Perspectives on Discipline-Based Knowledge Engine Construction

Starting from the AI teaching assistant, the construction of a discipline-based knowledge engine for urban planning will reshape the development landscape of the discipline across multiple dimensions, including knowledge, functionality, and application. It will open new pathways for cultivating innovative planning professionals who are capable of addressing future urban development challenges, solving complex urban planning problems, and advancing frontier research in the field (Fig. 11).

In terms of knowledge sources, the AI teaching assistant periodically integrates textbooks and course materials, standards and regulations, as well as the latest literature and case studies. New knowledge is embedded into the existing knowledge graph according to thematic units and course chapters, enabling a stable mapping from authoritative sources to specific clauses, indicators, and representative cases. This ensures that the model can continuously learn and dynamically update its knowledge base, thereby developing more accurate understanding and stronger logical reasoning over disciplinary concepts and rules.

In terms of knowledge fusion, existing textbooks and regulatory documents, recent research findings, and practical cases are first segmented, deduplicated, anonymized, and version-labeled. An indexing system is then constructed based on course syllabi and

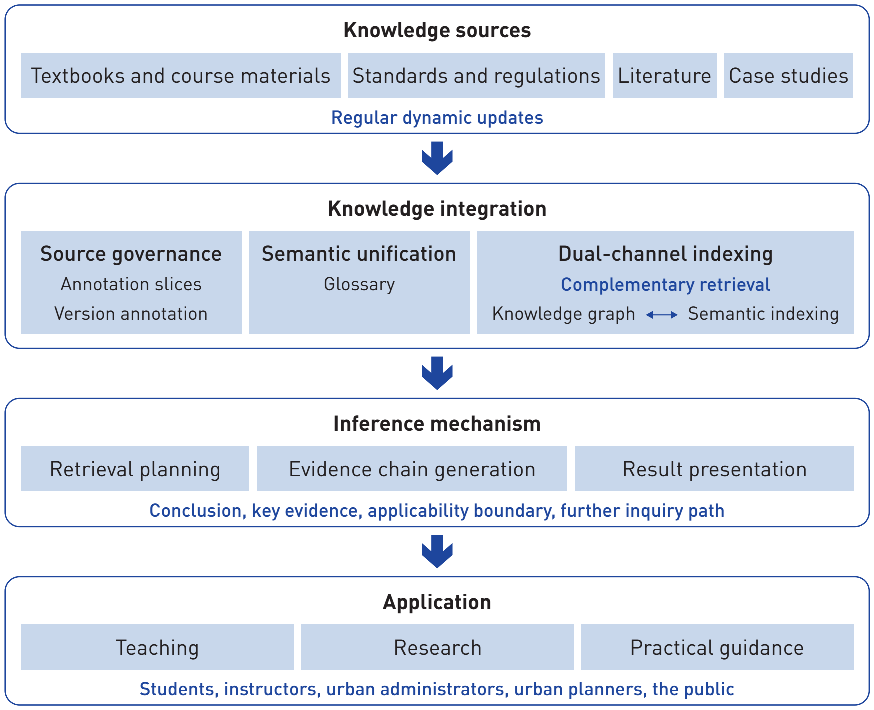


Fig. 11 Architecture of the discipline-based knowledge engine.

thematic tags. Second, a terminology glossary and conceptual network are used to constrain the usage boundaries of synonyms and semantically similar concepts. Finally, without altering the original usage of the knowledge system, new knowledge is precisely injected and dynamically updated.

In terms of reasoning mechanisms, retrieval order and scope can be dynamically adjusted according to task objectives. In research-support scenarios, the system strengthens retrieval of recent literature and case studies. In practical guidance scenarios, it simultaneously retrieves regulatory clauses and relevant cases. The evidence chain generation module links extracted clauses, parameters or conditions, applicability scopes, and citation sources at each retrieval step, and performs alignment and conflict resolution when necessary, ensuring that the intermediate reasoning process remains grounded in disciplinary evidence. Ultimately, the outputs are presented in a standardized structure comprising conclusions, key evidence, scope of applicability, and avenues for further verification.

Expanding application scenarios is also an important future direction for discipline-specific foundation models in urban planning. Urban managers can use such specialized models to predict urban development trends under different planning policies, thereby supporting decision-making. Planners and designers can leverage the models' capabilities in creative inspiration and regulatory compliance review to enhance both the innovativeness and compliance of design proposals. Meanwhile, the public may access personalized spatial planning recommendations through open interfaces.

5 Conclusions and Outlook

The study takes the New Science of Cities course as a case study and systematically reviews the development pathway and teaching practice of an AI teaching assistant. Focusing on text generation, image generation, and platform integration, it establishes an intelligent teaching support system for urban planning education. By constructing a high-quality domain-specific corpus and introducing a RAG mechanism along with fine-tuning of image-text models, the system significantly enhances AI capabilities in knowledge retrieval and visual representation. Meanwhile, platform integration further improves instructional interaction and enriches students' learning experience.

On this basis, the study proposes the concept of a "discipline-based knowledge engine" and constructs a multimodal educational system centered on LLM, integrating domain-specific corpora, knowledge graphs, and generative model optimization. This system not only supports classroom teaching but can also be extended to research training, competency assessment, and interdisciplinary collaboration, serving as an important intelligent infrastructure for advancing academic innovation and educational transformation.

However, the adoption of AI in teaching still faces several challenges, including insufficient understanding of teaching contexts, excessive student dependence, and limited system transparency. Therefore, the construction and application of the AI teaching assistant should not remain at the level of technological feasibility alone, but must be continuously reflected upon and optimized in relation to educational objectives and knowledge construction logic. Future work should further refine learning pathways and instructional interaction mechanisms under the guidance of constructivist learning theory and cognitive load theory, forming a hybrid paradigm in which technology and educational theory mutually reinforce each other, and promoting the transformation of AI from an assistive tool to a deeply embedded teaching partner.

In addition, ethical governance is a key issue for the sustainable development of AI-based education. In the future, more comprehensive regulatory frameworks need to be established for data privacy protection, algorithmic fairness, and system interpretability. Only with the dual support of technological innovation and ethical governance can AI teaching assistants truly become a driving force for the intelligent and sustainable transformation of higher education.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the technical support provided by the research team of Professor Hongning Wang from the Department of Computer Science and Technology, Tsinghua University, and by XuetangX.

Competing interests | The authors declare that they have no competing interests.

REFERENCES

[1] Long, Y., & Zhang, E. (2024). City laboratory: Criticism and prospect based on new data, new elements, and new pathways. *Journal of Urban and Regional Planning*, *16*(1), 30–46.

[2] Zhai, W., Jiang, Z., Meng, X., Zhang, X., Zhao, M., & Long, Y. (2022). Satellite monitoring of shrinking cities on the globe and containment solutions. *iScience*, *25*(6), 104411.

[3] Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274–279.

[4] Long, Y., Zhao, H., & Zhang, Y. (2024). The new science of cities: Technology, computation, evolution, and application. *City Planning Review*, *48*(7), 4–15.

[5] Xie, S. (2025). Problems in higher education management and optimization strategies in the context of big data. *Knowledge Library*, *41*(6), 128–131.

[6] Lu, X. (2025). Key issues in promoting the classified development of universities. *China Higher Education*, (1), 4–8.

[7] Hou, H., & Wang, Q. (2025). Analysis of modern academic system of pedagogy with Chinese characteristics. *Academic Exploration*, (2), 142–149.

[8] Cai, S., Gao, X., & Yao, H. (2024). Digital empowerment of graduate education: A student-centered path to high-quality development. *Education Research Monthly*, (11), 3–9.

[9] Zhu, Z., Dai, L., & Hu, J. (2023). AIGC technology enables new ideas for digital transformation of higher education. *China Higher Education Research*, (6), 12–19.

[10] Yang, Q., & Tang, Q. (2025). Generative AI empowering postgraduate education: Theoretical logic, legal risks, and governance approaches. *Journal of Graduate Education*, (2), 26–33.

[11] Li, J., & Zhong, L. (2024). Application of a CAID-based artificial intelligence interactive teaching method in graduate education reform. *Packaging Engineering*, *45*(S1), 489–493.

[12] Wu, Q., & Liu, Y. (2023). Higher education response in the era of ChatGPT: Prohibition or reform. *Journal of Higher Education Management*, *17*(3), 32–41.

[13] Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. *Behavioral Sciences*, *15*(4), 467.

[14] Salih, S., Husain, O., Hamdan, M., Abdelsalam, S., Elshafie, H., & Motwakel, A. (2025). Transforming education with AI: A systematic review of ChatGPT's role in learning, academic practices, and institutional adoption. *Results in Engineering*, *25*, 103837.

[15] Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, *15*(1), 17458.

[16] Zhou, Y., & Yu, P. (2025). Study on the exhibition design pedagogy based on Constructivist Learning Theory: A twelve-step curriculum model. *Zhuangshi*, (6), 101–105.

[17] Pang, W. (2011). Cognitive load theory and its pedagogical implications. *Contemporary Education Sciences*, (12), 23–28.

[18] Zhipu AI. (2025, April 17). *ChatGLM*.

[19] Kozlowski, S. W. J., & Chao, G. T. (2018). Unpacking team process dynamics and emergent phenomena: Challenges, conceptual advances, and innovative methods. *American Psychologist*, *73*(4), 576–592.

LAF | https://doi.org/10.15302/J-LAF-2026-0031
第14卷 / 第3期 / 260031 / 2026
收稿日期 | 2025-06-03　录用日期 | 2026-01-12
中图分类号 | TU986, TP18　文献标识码 |B

基于大语言模型的城市规划人工智能助教构建与学科引擎研究——以《新城市科学》课程为例

郑湉，夏俊豪，马悦，黄潇斓，龙瀛*

清华大学建筑学院，北京 100084

*通信作者
地址：北京市海淀区双清路30号
邮编：100084
邮箱：ylong@tsinghua.edu.cn

摘要

人工智能（AI）的快速发展正推动高等教育由传统教学模式向智能化教学形态深刻转型。本文围绕AI助教系统开发、教学实践应用与学科引擎构建3个维度展开探讨。研究结果表明，通过融合领域知识注入、检索增强生成机制与文生图能力优化，AI助教不仅显著提升了专业知识问答的准确性，还在城市空间场景的图像生成任务中展现出较高的表达精度及对中国本土城市场景的适配能力。基于此，本文进一步提出以大语言模型为基底、面向学科知识体系深度定制的学科引擎建设路径，构建融合了文本生成、知识图谱与文生图协同驱动的系统架构，以实现从通用型语言模型向学科智能引擎的演进。研究为高校教育的数字化转型提供了可复制、可推广的理论框架与技术路径，同时拓展了AI与教育交叉研究的应用场景与发展方向。

关键词

人工智能；AI助教；大语言模型；检索增强生成；文生图；教学实践；学科引擎

文章亮点

- 构建了融合检索增强生成与文生图能力的课程AI助教系统
- 提出了面向城市规划教育的学科引擎构建框架与技术路径
- 实现了大语言模型在教学问答、空间表达与课程实践中的协同应用

基金项目

- 清华大学本科教育教学改革项目“基于教育部虚拟教研室的城乡规划设计课全流程数智赋能案例库建设”（编号：DX02_34）
- 清华大学本科教育教学改革项目“《新城市科学》人工智能校内教育教学应用探索项目”（编号：DX08_05）

1 引言

第四次工业革命带来了以人工智能（AI）为代表的颠覆性技术，推动了科研范式和教学模式的持续转型。随着技术的快速发展，研究者得以整合文本、图像与地理空间等多源异构数据，以更精细、动态、可视的方式解析城市问题[1–4]，同时高等教育也正经历深层次的结构性调整与系统重构。

当前，高等教育面临优质教育资源在区域、学科和院校间分布不均，以及“教师中心—统一教材—标准进度”的传统教学模式难以回应学生多样化与个性化学习需求[5–7]等挑战。然而，在科技驱动下实现深层转型的路径机制、关键节点与作用逻辑，仍有待系统梳理与实证验证。研究显示，AI在教学内容生成、知识图谱构建、人机协同学习等方面具有显著潜力，有望重塑知识获取方式，推动教学生态向智能化、开放化与适应性方向转变[8–12]。近期研究已在虚拟助教、写作辅导、教案编制与作业批改等方面验证了AI的有效性[13–15]，但针对AI助教的系统架构、功能设计及其与教学场景融合的研究尚处于起步阶段，缺乏成熟的理论框架与实践范式。从教育理论视角看，AI助教若能与建构主义学习理论[16]及认知负荷理论[17]相结合，将有助于解释其促进个性化学习和深度认知的作用机制，为在高等教育中的推广提供理论依据。

本文以清华大学本科通识课程《新城市科学》为典型案例，旨在揭示AI助教从通用型语言模型向学科化智能引擎演化的关键机制，探索AI在教学中的创新整合路径，为高校推动AI技术与教学实践的深度融合提供可复制、可推广的理论框架与操作范式。

2 AI助教的构建方法与技术实现

目前，AI教学系统在文本交互方面多依赖关键词匹配或通用大模型的生成结果，其回答往往缺乏对专业知识和最新研究进展的深入理解。

此外，多数系统尚不具备图像生成能力，或仅通过基础模型生成图像，难以满足专业教学中的表达与认知需求。《新城市科学》AI助教在现有技术基础上，针对课程教学需求进行学科适配，开发了文本生成与图像生成两项核心功能。

文本生成通过构建高质量专业语料库并引入检索增强生成（Retrieval-Augmented Generation，RAG）机制，增强模型在课程知识问答中的教学适配性与学科契合度；图像生成则依托大规模图文配对数据集与模型微调，提升模型在未来城市教学语境下的视觉表达与认知支持能力。在此基础上，通过将两大核心功能整合至智谱清言网站及荷塘雨课堂，并引入功能卡片板块，基于典型任务场景为学生提供提示词模板支持，从而提升使用便捷性与教学适用性。与通用模型（如ChatGPT、Midjourney等）相比，定制化AI助教虽在语料整理、模型微调与平台集成方面投入更高，但在教学任务适配度、回答准确性与文化契合度方面具有显著优势。

2.1 文本生成功能

在文本生成方面，团队首先聚焦高质量领域数据的系统整理与构建，整体工作流程如图1所示，旨在为课程教学提供具有学科特征与教育应用价值的知识支撑。所使用的文本数据包括教材、在线课程资料、经典论文与案例文献等。通过构建内容丰富、知识跨度广泛的新城市科学与规划学科语料库，使模型在课程教学与学科研究中具备更强的知识理解与情境适配能力。

在模型能力优化方面，AI助教选用ChatGLM-4[18]作为底层大模型，并结合RAG技术实现对教学场景的专业化增强。与传统微调方式相比，RAG无需更新模型参数，具有迭代速度快、知识更新便捷等优势，尤其适用于领域知识快速扩展或更新频繁的场景。在确保专业性的同时，有效缓解了大模型的“幻觉”问题。在回答生成环节，通过提示词工程明确要求大语言模型采用课程特定分析视角与框架来组织论述，引导AI助教更充分、更有条理地利用检索到的语料进行回答。

2.2 文生图功能

在图像生成能力方面，为实现AI助教对城市场景的辅助表达，团队通过在自建高质量图文配对数据集上进行微调，以提升模型的文生图能力（图2）。数据来源包括人工筛选与自动化收集两部分。人工部分聚焦未来城市空间设计典型案例，筛选了1 100张具有代表性的图像，并由具备设计背景的研究人员进行标注。在自动化方面，从国际建筑设计平台采集35万张高分辨率真实场景图像，同时融合5万张高质量AI生成图像，并通过CogVLM[18]视觉语言模型自动生成图像描述，最终形成40万组图文配对数据，涵盖居住空间、办公空间与公共空间等多类场景，为模型训练提供了丰富样本。

基于上述数据集，团队以CogView3[18]为基础模型进行微调，采用低秩自适应（Low-Rank Adaptation，LoRA）技术对关键参数进行优化，并完成约2 000步的训练。为增强模型对提示词的理解能力，在图像生成前，系统调用大语言模型对用户输入的提示词进行改写与扩展，将含有专业知识的简要关键词转化为空间构成明确、场景描述完整的语句，从而提升图像生成的视觉质量与语义表达精度。

在测试评估阶段，围绕30类典型城市空间场景构建240条测试样本（每类场景配备8条描述），覆盖经典与未来、室内与室外等不同类型。结果显示，最新版CogView3-plus模型[18]在测试样本上的用户满意度约60%，而基于CogView3进行微调的模型在画面内容准确性、视觉风格契合度及文化语境表达方面均表现更优（图3），尤其在中国本土城市场景表达方面具有明显提升，测试满意度达75%，证实了该模型在辅助创意表达上的实用价值。

2.3 界面研发

团队将文本生成与图像生成两项核心功能集成至智谱清言平台与荷塘雨课堂，实现AI能力在真实教学场景中的应用。前者作为学生测试平台，在功能完善后接入后者，以更好地链接现有课程资源，提升整体学习体验。

为提高学生操作效率与提问针对性，团队在智谱清言平台设计并部署了涵盖实验计划、观察报告、改造方案、修改评价，以及出题共5类应用场景的25个功能卡片模块（图4）。功能卡片获得了学生的积极反馈，随后迁移至荷塘雨课堂系统，以进一步拓展在实际教学场景中的应用价值。团队基于教学经验梳理了课程重点与难点，定制提问模板，并以功能卡片形式嵌入系统，使学生在完成章节学习后可直接调用对应卡片获取结构清晰的智能回答（图5）。同时，通过整合知识图谱、MOOC资源与课程资料，AI助教系统实现了多源学习资源的联通，可促进学生深入理解课程内容（图6）。

3 AI助教在教学实践中的应用

3.1 课堂智能交互

在课堂教学中，教师可利用AI助教根据课程进度生成兼具学术性与启发性的导入内容，激发学生的学习兴趣。在讲授过程中，AI助教作为“第三教学主体”参与师生互动，为学生提供启发性回应，辅助教师引导批判性思考。在讨论环节，学生可借助AI助教明确讨论方向、补充知识与案例，从而提升讨论质量与广度（图7）。

3.2 课后辅导支持

在课后阶段，AI助教围绕知识深化和作业支持提供多维辅助。针

对课堂上未能完全理解的内容，学生可通过功能卡片快速提问，或调用24小时智能学伴获取解答（图8）。对于拓展性学习需求，可即时调取知识库中的相关内容进行延伸学习。此外，系统支持语音互动、截图问答及语音生成思维导图等多种交互方式。

AI助教贯穿课程项目的各个阶段，可协助构建研究框架、梳理思路并推荐研究方法；结合学术资源提供分析建议，帮助优化逻辑结构；高效筛选文献与案例，提升信息整合能力。在可视化表达方面，其图像生成功能可辅助呈现未来场景，增强成果表现力。最终，系统还可自动润色文本，提升语言规范性与学术表达水平。

3.3 学生及教师评价

为评估AI助教的实际应用效果，团队对2023年和2024年秋季学期的学生使用数据进行分析，并通过问卷方式收集了学生反馈。

在使用情况方面，2023年秋季学期共有30名学生参与反馈，约70%的学生在信息检索、概念理解、研究框架搭建及文字润色等环节使用了AI助教。2024年秋季学期，17名学生全部使用AI助教的底层大模型ChatGLM进行辅助，总使用次数达432次，最大提问次数达108次，最长使用时间达24 h 11 min（图9）。不同学习风格的学生在AI助教使用模式上呈现出显著差异：倾向于独立思考的学生多以批判性和验证性使用为主，聚焦于信息检索与快速核验，因而使用时长相对较短；而偏好互动式学习的学生则更倾向于将其作为对话性认知工具，通过持续交流拓展认知边界并深化理解，因而使用时间相对较长。此外，使用时长排名前20%的学生，其课程成绩多为“优秀”，在一定程度上反映出AI助教使用程度与学习表现之间的正相关趋势，但该关系可能同时受到学习投入等因素的影响，尚需进一步验证。

同时，学生反馈结果表明，AI助教在优化研究思路与表达质量方面具有明显作用，但仍存在回答深度不足、图像生成质量不稳定、响应速度较慢、文献查找功能不完善等问题。部分学生建议增强AI助教对城市空间和校园环境等场景的理解力、提升上下文记忆与更新能力、减少模板化表达，以及优化跨领域适应性。此外，已有研究表明，过度依赖智能工具可能削弱学习者的自主学习与问题解决能力[19]。因此，在教学实践中需明确AI助教的辅助角色，通过教学引导维持学生的批判性思维与反思能力。未来AI助教的研发不仅需从技术层面提升回答的深度与灵活性，还应在设计层面强化可解释性与透明性，从而引导学生掌握与AI对话的能力，而非依赖AI。

从教师视角来看，此类实践涵盖了AI助教的建设与应用两个层面。教师既是AI助教的设计者和搭建者，也是教学过程中的使用者和观察者。在建设过程中，关键难点主要包括教学需求的精准提炼、知识库的科学构建，以及系统与课程内容的深度契合。使用过程中的操作要点主要包括以下方面：一是在课堂教学中适时引入AI助教，突出其辅助而非替代作用，以有效激发学生兴趣、提升课堂参与度；二是在课后充分发挥AI助教的知识性解答功能，将常规性、重复性问题交由系统处理，从而使教师能够更加专注于引导学生开展系统性思考；三是针对学生的个性化问题，结合AI助教的反馈结果，提供更具专业深度的分析与指导。

4 AI赋能学科引擎建设的探索

4.1 从AI助教迈向学科垂直大模型

清华大学率先提出基于在线教育布局和AI赋能教学等实践，创新开展AI赋能学科引擎建设，以推动高等教育教学范式转型。《新城市科学》AI助教可作为起点，逐步迈向城市规划领域的学科引擎建设。学科引擎的技术架构主要由4个层次构成：知识图谱层围绕城市规划领域的核心概念、方法与案例，将多源知识以图谱形式组织，实现逻辑推理和跨领域检索；多模态融合层引入文本、图像、表格等多模态信息，通过跨模态对齐与表示学习，使模型能够协同理解和处理图像、统计数据与文本知识；模型演进层以通用大语言模型为基础，结合RAG机制，通过领域语料微调、知识注入与人机反馈，逐步演进为面向城市规划学科的垂直大模型；应用接口层则支持个性化理论学习、规划设计指导等多样化应用场景。由此，AI助教不再局限于教学辅助工具，而进一步具备面向城市规划学科的泛化与推理能力。

具体而言，团队进一步拓展了AI助教的输入内容和应用范围（图10）。在原有AI助教模型的基础上，依据城市规划本科生培养方案，确定20个城市规划子领域方向，构建内容丰富、知识跨度广泛的新城市科学与规划学科语料库，为大语言模型提供扎实的领域知识支撑。同时，团队收集了49份注册城乡规划师真题和模拟题用于模型测试。结果显示，该学科垂直大模型在相关测试中的表现优于其他通用模型（表1）。

4.2 学科引擎建设的未来展望

以AI助教为起点建设城市规划学科引擎，将在知识、功能与应用等多个维度重塑学科发展格局，为培养适应未来城市发展需求的创新型规划人才、解决复杂城市规划问题及推动学科前沿研究开辟新的路径（图11）。

在知识来源方面，AI助教可定期整合教材与课程资源、规范标准，以及最新文献与案例，并按主题与章节将新知识嵌入既有知识图谱，实现从权威来源到具体条款、指标与典型案例的稳定映射。保证模型能够在持续学习与动态更新的基础上，对学科概念与规则形成更精准的理解与逻辑推理能力。

在知识融合方面，首先对既有教材与规范条文、最新研究成果与实践案例进行切片、去重、脱敏与版本标注，并根据课程大纲与主题标签

表 1：学科垂直大模型与通用大语言模型的性能对比

大语言模型	单选题（933 道）	判断题（217 道）
ChatGLM-4 + 专业知识	77.4%	72.6%
ChatGLM-4	75.0%	71.3%
GPT-4o	70.7%	72.2%
Claude 3.5 Sonnet	70.2%	70.0%
Gemini-Pro	54.7%	58.8%

注

1. 百分比表示各模型回答结果与标准答案之间的匹配准确率。
2. 本研究所选用的底层大模型 ChatGLM-4 发布于 2024 年 1 月，而 DeepSeek-R1、Qwen2.5-72B 等大模型发布于 2024 年下半年，晚于本研究中学科垂直大模型的构建阶段，因此未被纳入性能对比范围。

构建索引体系。其次，利用术语表与概念网络约束同义词与近义概念的使用边界。最后，在不改变现有知识体系用法的前提下，实现新知识的精准嵌入与动态更新。

在推理机制方面，根据任务目标动态调整检索顺序与范围：在研究支持场景中，加强近年文献与案例检索；在实践指导场景中，同步检索规范条文与相关案例。证据链生成模块则在每一步检索后，将条文摘录、参数或条件、适用范围与引用位置进行串联，并在必要时开展对照与冲突消解，使模型的中间推理过程始终受到学科证据约束。最终结果采用“结论—关键依据—适用边界—可再查路径”的统一结构进行表述。

应用场景的拓展也是未来城市规划学科引擎的重要发展方向。城市管理者可利用该专业模型预测不同规划政策下的城市发展趋势，为科学决策提供支持；规划设计师可借助模型的创意启发与规范审核功能，提高设计方案的创新性与合规性；公众则可通过开放接口获取个性化空间规划建议。

5 结论与展望

本研究以《新城市科学》课程为案例，系统梳理了AI助教的构建路径与教学实践，围绕文本生成、图像生成与平台集成，构建了面向城市规划课程的智能化教学支持体系。通过构建高质量领域语料库，并引入RAG机制与图文模型微调，该系统显著提升了AI在知识检索与视觉表达方面的能力。同时，平台化集成进一步优化了教学互动与学生学习体验。

在此基础上，研究提出“学科引擎”概念，构建了以大语言模型为核心，融合专业语料、知识图谱与生成模型调优的多模态教学系统。该系统不仅能够支持课堂教学，还可拓展至科研训练、能力评估与跨学科协同等场景，成为推动学术创新与教育转型的重要智能基础设施。

然而，AI教学的推广仍面临教学语境理解不足、学生依赖性强，以及系统透明性欠缺等挑战。因此，AI助教的构建与应用不能仅停留于技术“可用”层面，更需在教育目标与知识建构逻辑上持续反思与优化。未来应在建构主义学习理论与认知负荷理论指导下，进一步优化学习路径与教学互动机制，形成技术与教育理论相互支撑的复合范式，推动AI从辅助工具向深度嵌入的教学伙伴转变。

此外，伦理治理是AI教育可持续发展的关键议题。未来需在数据隐私保护、算法公平性与系统可解释性等方面建立更加完善的规范体系。只有在技术创新与伦理规范的双重保障下，AI助教才能真正成为推动高等教育智慧化与可持续化转型的重要力量。

致谢

本研究得到清华大学计算机系王宏宁教授团队及学堂在线的技术支持，特此致谢。