Education Journal
2025, Vol. 14, No. 3, pp. 111-125
https://doi.org/10.11648/j.edu.20251403.14

Research Article

Practical Applications of Generative AI in Educational Support

Zhihua Zhang*

Department of Data Science Education, Kansai University of International Studies, Kobe, Japan

Abstract

This paper investigates the practical application and effectiveness of a custom generative AI teaching assistant, developed as a GPTs application named DS-ASST, within a "Data Science" course at Kansai University of International Studies. As generative AI technologies like Large Language Models rapidly advance, their potential to transform educational paradigms becomes increasingly evident, offering solutions to challenges inherent in traditional methodologies, particularly in areas requiring personalized learning and large-scale data interaction. This research utilizes the 'Reconsidering Education in the AI Era' framework, focusing on educational AI alignment and redesigning practices to support both instructors and students. The DS-ASST system was developed using Retrieval-Augmented Generation technology to integrate course-specific materials, including lecture notes and textbook content, ensuring responses are contextually relevant and minimizing factual inaccuracies or "hallucinations". We detail the system architecture, iterative prompt design experiments aimed at optimizing educational value, and strategies employed to mitigate technical challenges like hallucination. The system's effectiveness was evaluated through formative assessment across four key dimensions: enhancing teaching preparation efficiency, supporting active student learning, improving data analysis processes, and promoting advanced learning activities. Key findings indicate significant improvements, including a notable reduction in instructor preparation time (approximately 42%) and increased student engagement in discussions (38%) compared to control groups. The AI assistant effectively provided on-demand concept clarification, guided problem-solving, facilitated interaction with complex data, and supported advanced activities like critical evaluation and ethical reasoning. While demonstrating substantial benefits over traditional methods in scalability and flexibility, limitations related to domain specificity, assessment capabilities, and technical requirements were noted. This research assessed the wider ramifications of generative AI for educational reform, based on its practical implementation in education. It specifically considered the changing roles of teachers, developments in assessment techniques, and the necessity of ethical literacy. The study also outlined potential future research, emphasizing hybrid teaching approaches and the formulation of sound ethical guidelines.

Keywords

LLMs, Generative AI, GPTs, Educational Applications of GenAI, Data Science Education

1. Introduction

Generative Artificial Intelligence (GenAI) technologies have advanced rapidly, significantly impacting various societal domains including education. GenAI adopts a data-driven approach, fundamentally differing from conventional AI. At

*Corresponding author: z-zhang@kuins.ac.jp (Zhihua Zhang)

Received: 25 March 2025; **Accepted:** 2 April 2025; **Published:** 26 May 2025

its core are foundation models – general-purpose neural networks pre-trained on massive datasets, adaptable to diverse tasks like image recognition, language understanding, and text generation. This generation of AI exhibits exceptional performance in classification, generation, and prediction, with a strong capability for acquiring knowledge from data. Particularly influential are Large Language Models (LLMs) such as ChatGPT and Deep Seek. Within educational settings, these tools offer transformative potential by enhancing intellectual productivity, automating tasks, and expanding possibilities for personalized learning support and classroom assistance.

Traditional educational methodologies often face limitations in processing vast amounts of information and facilitating immediate, individualized student interactions. GenAI presents potential solutions to these challenges. Specifically, it can enable highly personalized learning pathways by adapting content dynamically to individual student progress and comprehension levels. Furthermore, GenAI facilitates interactive, dialogue-based learning environments for real-time questioning and flexible explanatory approaches using multiple modalities (text, diagrams, analogies), fostering a more learner-centered paradigm previously difficult to achieve. Capabilities inherent in LLMs, such as automated question-answering, document summarization, teaching material creation, and even advanced reasoning, can significantly improve educator efficiency while promoting tailored learning experiences.

LLMs tools like OpenAI's ChatGPT and Google's Gemini offer diverse functionalities beyond text generation, including question-answering, dialogue maintenance, document summarization, and advanced reasoning. These capabilities enhance educational settings by supporting automated question-answering, teaching material creation, and interactive learning, improving educators' efficiency and promoting personalized learning. Custom GenAI-based tools like GPTs and NotebookLM, leveraging Retrieval-Augmented Generation (RAG), are poised to revolutionize education. They can provide personalized learning paths, offer real-time support, and enable innovative teaching methods such as role-playing and interactive problem-solving. These technologies have the potential to fundamentally transform educational approaches, creating more efficient and adaptive learning environments that maximize individual learner potential.

Recognizing this potential, significant research and policy efforts are underway globally. Studies have investigated Data Science curriculum development for liberal arts universities, addressing aspects like knowledge domain-based models, implementation barriers, and future directions, thereby significantly contributing to the field's growth and dissemination [1] and research on how GenAI should be utilized in education analyzed optimal GenAI utilization strategies considering guidelines from bodies like UNESCO and national ministries [2, 3].

Governmental reports, such as those from the U.S. Department of Education [4, 5] and Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) [6], emphasize both the opportunities (improved quality, access, efficiency) and the critical need for responsible implementation. Key considerations include ethical usage, fairness, transparency, privacy protection, alignment with educational values, evidence-based practices, and security [4, 5, 6]. Curricula are also evolving, with frameworks like the 2024 MDASH revision in Japan explicitly integrating GenAI education at both literacy and applied levels [13, 14].

Despite this progress, a crucial challenge remains, particularly for institutions like liberal arts universities: effectively integrating GenAI's capabilities with established pedagogical approaches to demonstrably improve learning outcomes and streamline educators' workflows. The central question is how to synergistically combine GenAI technology with educational objectives, task design, and traditional teaching practices to construct novel, effective educational models for the AI era.

To address this challenge, this research employs the 'Reconsidering Education in the AI Era' framework (Figure 1). This framework advocates for a holistic approach centered on 'educational AI alignment'—ensuring truthful, fair, accessible, and secure AI integration—and operationalized through four key concepts: 'generative AI engagement' (practical understanding of AI), 'reconsidering learning objectives' (redefining goals incorporating AI literacy), 'revise your assignments' (designing assessments mindful of GenAI), and 'redesigning educational practices' (developing AI-enhanced personalized and collaborative methods).

As a practical application and empirical validation of this framework, this study details the construction and evaluation of a dedicated "Data Science" AI teaching assistant system. This system was implemented and tested within the "Data Science" core course of the Data Science Minor program (MDASH literacy level) at Kansai University of International Studies [7, 8]. Through this empirical investigation, we aim to illuminate the practical application of the theoretical framework, evaluate the transformative potential of GenAI in a real-world educational context, and identify both the benefits and challenges encountered.

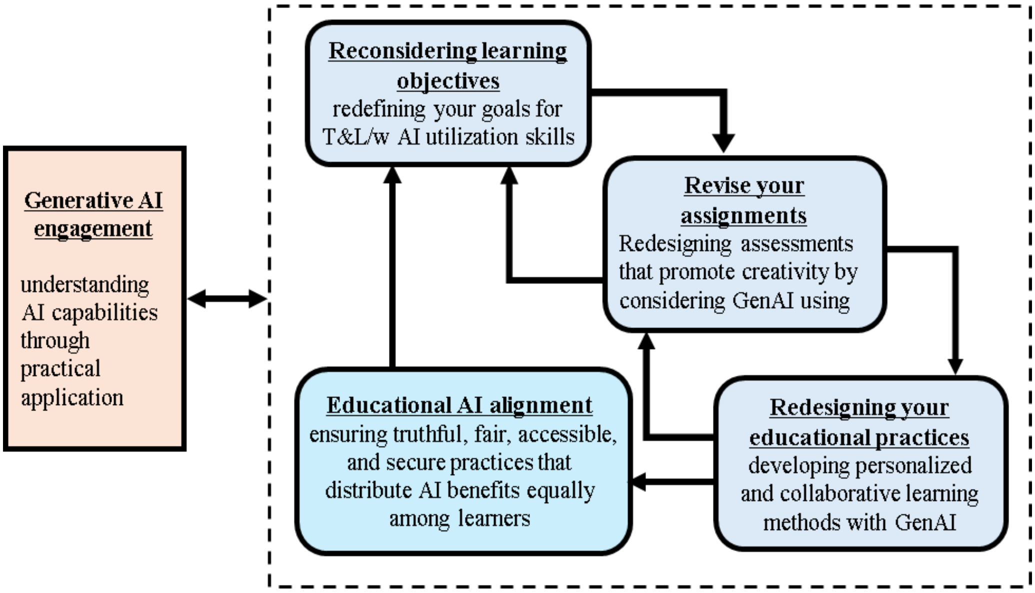

Figure 1. *Reconsidering Education in the AI Era.*

2. Representative LLMs, Prompt and Hallucinations

LLMs represent a significant advancement in natural language processing, developed leveraging extensive text corpora sourced primarily from the internet and employing sophisticated deep learning methodologies. The capabilities of these models in language understanding and generation are largely underpinned by innovations in deep learning, particularly the transformer architecture [15]. Major models such as ChatGPT (OpenAI), Gemini (Google), Claude (Anthropic), Copilot (Microsoft), and DeepSeek possess unique architectures and strengths.

2.1. LLMs: Overview, Characteristics and Ability to Handle Intelligent Tasks

The landscape of GenAI includes several prominent models relevant to various applications, such as ChatGPT, Gemini, Claude, Microsoft Copilot, and Deep Seek. These tools collectively offer diverse capabilities spanning natural language processing, multimodal input handling, and specialized functions, although each possesses unique strengths. Furthermore, the performance characteristics and features of these GenAI tools are evolving rapidly.

Examining their specific attributes reveals distinct advantages relevant to educational contexts:

1. ChatGPT demonstrates versatile conversational abilities and flexible text generation, making it highly adaptable for various educational dialogue and content creation scenarios.
2. Gemini excels with its strong multimodal processing capabilities and integration with Google Search, allowing it to effectively handle diverse forms of educational content and information queries.
3. Claude is noted for its advanced reasoning capabilities coupled with a particular emphasis on ethical considerations, a crucial aspect for responsible AI deployment in education.
4. Microsoft Copilot, leveraging its integration with Microsoft 365, shows strength in productivity and business-oriented tasks, including programming support, offering specialized functionality valuable in professional or vocational educational settings.
5. DeepSeek features open-source models with publicly available weights, allowing free use and customization. It offers high, cost-effective performance while being efficient with GPU resources. This is achieved using cutting-edge technologies like MoE, RLHF, knowledge distillation, and FP8 training. DeepSeek is versatile and multilingual, supporting applications such as text generation, summarization, translation, data analysis, and programming assistance.

A practical consideration for adoption is that these models are typically available in tiered versions. Premium or subscription options generally provide enhanced features, performance, and usage limits compared to free tiers. This tiered availability presents important strategic considerations for educational institutions needing to balance performance requirements, desired functionalities, and budget constraints when selecting and deploying these tools.

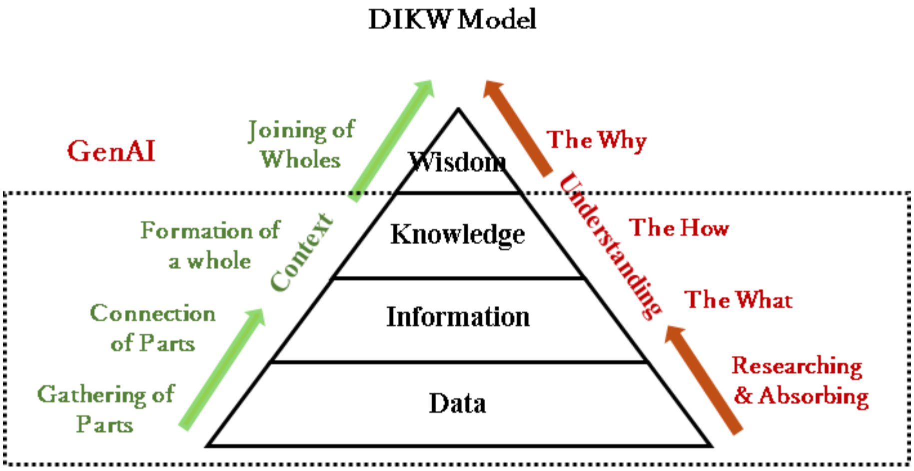

Figure 2. The Position of GenAI in the DIKW Model.

The DIKW (Data-Information-Knowledge-Wisdom) hierarchy, illustrated in Figure 2, provides a framework for understanding the progression from raw data to actionable insight. In this model, Data represents discreet, context-free facts (e.g., characters, numbers); Information emerges when relationships between data points are established (answering 'what' and 'how'); Knowledge involves synthesizing information into generalized understanding and formalized theories (addressing 'why'); and Wisdom signifies the ability to discern and apply knowledge effectively within broader contexts. This hierarchy helps analyze how GenAI technologies engage with different levels of abstraction.

1. GenAI significantly advances beyond traditional data processing by operating across multiple layers of the DIKW model.
2. At the Data layer, GenAI demonstrates sophisticated pattern recognition capabilities, excelling in processing large-scale, unstructured data with notable flexibility and scalability, offering potential benefits for analyzing diverse educational inputs.
3. Moving to the Information and Knowledge layers, GenAI shows superior ability in extracting context-dependent meaning, synthesizing disparate information, and reconstructing or generating domain-specific knowledge.
4. However, GenAI's current capabilities at the Wisdom layer are limited. While highly effective in processing and presenting information and knowledge, it primarily complements human judgment and decision-making rather than possessing independent wisdom.

Therefore, in educational contexts, while GenAI tools offer substantial support, their deployment necessitates careful human oversight. Educators remain crucial for verifying the accuracy, contextual relevance, and pedagogical alignment of AI-generated outputs. The optimal approach involves utilizing GenAI as a collaborative tool that augments and enhances—rather than replaces—human critical thinking, creativity, and the application of wisdom.

Recent GenAI models represent a significant leap beyond the performance boundaries of traditional machine learning, showcasing remarkable proficiency in advanced cognitive tasks. Particularly since 2023, LLMs have demonstrated capabilities that rival or even surpass human expert levels in domains demanding complex reasoning and specialized knowledge. Illustrative examples include GPT-4 achieving human-level performance across various professional and academic examinations [10], and a Google DeepMind AI model solving International Mathematical Olympiad (IMO) level problems with high accuracy (4 out of 6 correct) [12].

2.2. Prompt Engineering and Hallucination Challenges

Effectively utilizing LLMs in educational settings critically depends on prompt engineering: the systematic design of inputs to elicit desired outputs from the AI. This practice necessitates understanding the model's capabilities and limitations in relation to specific pedagogical objectives. Carefully constructed prompts are essential for guiding LLMs to produce responses that are not only accurate and relevant but also educationally sound.

The capacity for complex reasoning in LLMs can be significantly improved by generating thought chains, which are sequences of intermediate reasoning steps. Chain-of-Thought (CoT) prompting represents a simple yet highly effective technique for achieving this [11]. Its effectiveness has been empirically validated across diverse challenges, including arithmetic, common sense, and symbolic reasoning problems, indicating considerable utility for educational applications where explicit reasoning is beneficial.

Despite these capabilities, a significant challenge in deploying LLMs for education is the phenomenon of hallucination—the generation of plausible-sounding but factually incorrect or nonsensical information. This issue is particularly detrimental in educational contexts where accuracy and reliability are paramount. Potential causes include limitations or biases in the training data, architectural constraints, or ambiguity in the input prompts.

Addressing hallucination requires multifaceted mitigation strategies. From a prompt engineering perspective, providing clear and sufficient context is vital. System-level approaches include incorporating fact-checking mechanisms and designing models to acknowledge uncertainty rather than presenting fabricated information as fact. Vigilance against hallucinations is crucial in educational applications to prevent the propagation of misinformation that could undermine learning objectives and mislead students, emphasizing the need to treat LLMs as collaborative tools requiring critical oversight.

2.3. Custom GPTs and Educational Support

Custom GPTs, also known as GPTs or GPT Builder, represent specialized versions of ChatGPT applications that have been fine-tuned for specific domains or purposes. A significant advantage of these systems is their accessibility—they can be customized without extensive programming knowledge or skills, enabling many users to create personalized ChatGPT applications tailored to their specific needs. This accessibility has facilitated widespread adoption across various business applications and educational contexts.

In educational settings, custom GPTs can be tailored to specific datasets and requirements, providing more specialized question-answering capabilities and functions than general models. This customization allows for the development of AI assistants that address the needs of specific courses, subjects, or educational approaches.

2.4. RAG Technology

GPTs represent an innovative approach to enhancing LLMs through integration with RAG technology, allowing them to incorporate domain-specific knowledge. Similar tools include Google's NotebookLM. Unlike standard ChatGPT, GPTs offer retrieval functionality and the ability to call external APIs.

The retrieval capability allows ChatGPT to be pre-loaded with knowledge about specific topics or fields, enabling more accurate and faster response generation. Additionally, the integration with external APIs enables the incorporation of services like Google Maps or weather forecasts, producing higher-quality responses that leverage external functionality.

RAG technology integration enables GenAI systems to extract relevant information from external databases or knowledge repositories during response construction. This approach allows AI models to generate more accurate and contextually appropriate responses, proving particularly effective when answers to questions about specific specialized fields or specified data content are required.

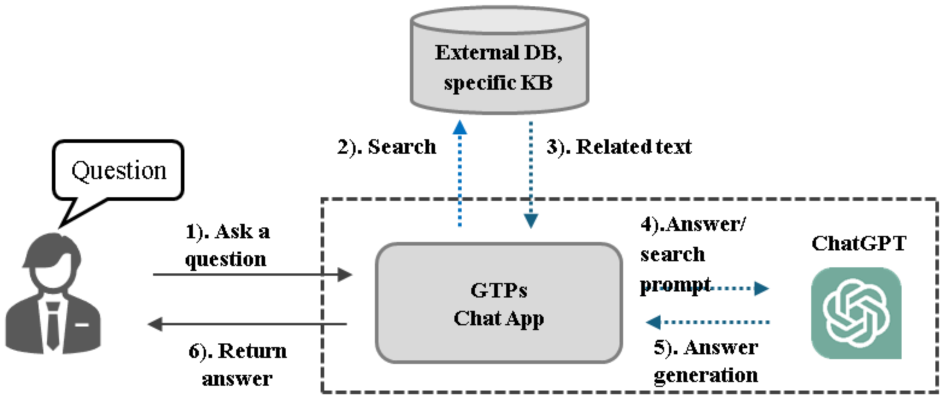

Figure 3. *GPTs mechanism incorporating external data sources using RAG technology.*

Figure 3 illustrates the question-answering mechanism of GPTs incorporating external data sources through RAG technology. The process follows these main steps:

1). Question reception: The user asks a question to the chat application.

2). Question processing: The application processes the question using vector data for search.

3). Retrieval of relevant text: Search results provide related text based on question keywords.

4). LLM query: Search results and the question are incorporated into a prompt template and input to ChatGPT.

5). Response generation: LLMs like ChatGPT generate response text.

6). Response transmission: After necessary post-processing, the response is sent to the questioner.

RAG technology allows GPTs to retrieve information from external sources and knowledge bases, creating more sophisticated outputs. In education, this enables curriculum-specific learning tools, contributing to more reliable educational support systems. Applications include specialized teaching materials, personalized learning support, and real-time responses to classroom questions. These capabilities help educators implement individualized teaching methods, improving the quality of students' educational experiences.

2.5. Guidelines for Use of GenAI in Education

The U.S. Department of Education's 2023 report recognizes AI's potential to improve education quality, expand opportunities, and enhance efficiency, while emphasizing it's not a universal solution and requires addressing ethical considerations, fairness, transparency, and privacy protection [4].

Their 2024 "AI Design for Education" guide outlines five key recommendations: understanding educational values, incorporating evidence-based practices, promoting fairness, ensuring safety, and fostering transparency and trust [5].

Additionally, the Consortium for Mathematics, Data Science, and AI Education Enhancement's 2024 model curriculum addresses GenAI in two stages: at the literacy level as an example of utilizing latest AI technology, and at the applied basic level as foundational theory and applied technology, providing educational institutions with a roadmap for AI integration [13, 14].

2.6. Ethical Considerations in Educational AI Implementation

Implementing AI in education requires addressing several ethical considerations: ensuring data privacy and security, preventing bias and discrimination, maintaining transparency about AI's capabilities and limitations, and preserving human agency in education. Educational AI systems must protect student data and avoid perpetuating biases. Clear communication about AI's role and limitations is essential when students and educators interact with these systems.

Ethical implementation also demands attention to accessibility and equity, ensuring AI-enhanced education is available to all students regardless of socioeconomic status or location. This includes addressing the "digital divide" to prevent AI from worsening educational inequalities.

By proactively addressing these considerations, institutions can maximize AI's benefits while minimizing risks and aligning with core educational values.

3. Methodology

3.1. Overview of the "Data Science" Course

This study focuses on the "Data Science" course, a core component of the "Data Science Minor" educational program (MDASH literacy level) implemented at Kansai University of International Studies since 2022. This program is available to all students at the university, providing foundational knowledge in data science regardless of their major field of study [7, 8].

The "Data Science" course serves as an introductory-level subject covering mathematics, data science, and artificial intelligence. The syllabus is structured according to the "Cultivation of Data Thinking" model curriculum for mathematics, data science, and AI (literacy level) provided by the "Consortium for Mathematics and Data Science Education Enhancement." The course utilizes "Data Science as Liberal Arts" [9] as its primary textbook, which is recommended by the Consortium for Mathematics, Data Science, and AI Education Enhancement.

Table 1. *Course Schedule and Assessments.*

Session	Content	Assessment
Session 1	Introduction	Pre-course Survey
Session 2	Changes in Society (1)	Quiz-1, Assignment-1
Session 3	Changes in Society (2)	Quiz-2, Assignment-2
Session 4	Data Utilization in Society	Quiz-3, Assignment-3
Session 5	Applications of Data and AI	Quiz-4, Assignment-4
Session 6	Data & AI Utilization Techniques (1)	Quiz-5, Assignment-5
Session 7	Data & AI Utilization Techniques (2)	Quiz-6, Assignment-6
Session 8	Practical Applications of Data & AI	Mid-term Exam
Session 9	Emerging Trends in Data & AI Utilization	Quiz-7, Assignment-7
Session 10	Reading Data	Quiz-8, Assignment-8
Session 11	Explaining Data	Practical Exercise 1
Session 12	Handling Data	Practical Exercise 2
Session 13	Considerations in Handling Data & AI	Quiz-9, Assignment-9

Session	Content	Assessment
Session 14	Considerations in Data Protection	Quiz-10, Assignment-10
Session 15	Review and Summary	Post-course Survey

This course mixes learning new concepts, which take up about two-thirds of the time, with practical, hands-on exercises using data for the remaining one-third. The course is divided into three main sections taught sequentially throughout the semester. The first section, "Introduction," covers changes in society, current data usage, and AI technologies. Following this, the "Fundamentals" section focuses on understanding and working with data (data literacy) and basic data science methods. The final section, "Principles," addresses important issues such as ethics, laws, and societal impacts (often called ELSI), data privacy rules like GDPR, considerations for using AI, and data security. For a detailed schedule of topics for each class, please refer to Table 1.

Each session includes:

1. PDF slide materials
2. Video content with AI narration
3. 10-question quizzes before and after videos
4. Outside-class tasks are due before the next lecture.
5. Q&A corner for knowledge sharing.

3.2. Development of the Custom GPT for Data Science Education

3.2.1. System Architecture and Design

We created the GPT-based AI teaching assistant for Data Science education (DS-ASST app), using RAG to integrate course materials. This accessibility was particularly valuable in the educational context, enabling the creation of a specialized tool tailored to the specific needs of the data science curriculum. This approach allowed the AI assistant to access and utilize course-specific materials, textbook content, and supplementary resources when generating responses to student queries.

The system was designed to support multiple aspects of the educational process:

1. Instructor support for course preparation and material development.
2. Student support for understanding course concepts and completing assignments.
3. Data analysis support for both instructors and students.
4. Advanced learning activity facilitation.

3.2.2. Prompt Design Experiments

A critical aspect of developing the AI teaching assistant system was the design and optimization of prompts. We conducted systematic experiments with different prompt structures to determine the most effective approaches for educational support. These experiments focused on several key areas:

1. Clarity and Specificity: Testing various levels of detail and specificity in prompts to determine the optimal balance for educational contexts
2. Context Provision: Experimenting with different methods of providing course-specific context to the AI system
3. Instructional Design Integration: Incorporating principles of instructional design into prompt structures
4. Error Reduction Strategies: Developing prompt techniques to minimize hallucinations and factual errors
5. Pedagogical Alignment: Ensuring that prompts elicited responses aligned with the pedagogical goals of the course

The prompt design process was iterative, with each version being tested and refined based on the quality and educational value of the responses generated. This process resulted in a set of optimized prompt templates for different educational scenarios, such as concept explanation, problem-solving guidance, and feedback provision.

As a response to rapid technological evolution, it is possible to implement related teaching materials in external data sources and knowledge bases in PDF or text format in the GPTs application. In addition, the DS-ASST app can automatically generate supplementary explanations using multifaceted perspectives and concrete examples to promote understanding of complex concepts or learning items that teachers feel are problematic. As an example, in response to a request for a simple and easy-to-understand explanation of GenAI, the DS-ASST app has the function of providing an explanation as shown in Table 2. The research demonstrated that AI can automatically generate customized quizzes and assignments aligned with specific curriculum requirements. The system can create questions with varied formats (multiple choice, written, coding) and difficulties, and can even fulfill specific student requests like randomly selecting a question with its solution. Verification confirmed that the generated content precisely matched the existing test materials of the 6th test-problem5.txt in the knowledge base.

Table 2. Example of lesson review prompting.

Prompt experiment of reviewing test questions
(Question prompt) Randomly select one question from Quiz Question 5 and provide the correct answers and explanations. (DS-ASST app) Question 1: Choose the most inappropriate reason that makes prediction difficult. Because there are phenomena that are inherently unpredictable, like the roll of a die. Because the future situation may differ from the past, and rules derived from past data may not apply to the future. Because the latest prediction theories are too mathematically advanced. Because it can be difficult to gather the necessary data for prediction. Correct Answer: 3 Explanation: The high complexity of the latest prediction theories is not a cause of prediction difficulties. Even with advanced prediction theories, factors such as unpredictable phenomena (randomness), future situational changes, and lack of data affect the accuracy of predictions, making prediction itself not easy. [54:1 † 6th_test-problem5.txt].

3.2.3. Avoiding Hallucinations and Using in Assessments

The analysis of past use case experiments has confirmed that the frequency of hallucination problems discussed above is extremely low on the DS-ASST app. Specifically, the system appropriately indicates a lack of knowledge in response to questions about content outside the scope of the teaching materials (see Table 3). However, in response to intentionally leading questions such as "Please create a CSV file containing the number of visitors to the Kansai University of International Studies' Sakura Festival," the generation of pseudo-data such as that shown in Figure 4 was observed. Here, CSV data was generated even though the festival is not an event that exists.

Table 3. Example of hallucination verification by DS-ASST.

Prompt experiment of hallucination problem verification
(Question prompt) What is the content of Quiz 15? * (DS-ASST app) No information on the content of Quiz 15 was found in the current materials. If you think a specific file is related to Quiz 15, please upload it to confirm. Also, if you have any other hints about the content you are looking for, we may be able to provide more accurate information. (Question prompt) The GAN model uses the transformer method. (DS-ASST app) The GAN (Generative Adversarial Network) model does not basically use the transformer method. Both GAN and transformer are attracting attention as generative models, but each has a different architecture and mechanism. About the GAN model Structure: GAN is a mechanism in which two networks, a generative network (Generator) and a discriminator network (Discriminator), compete to learn. Operating principle: The Generator generates new data, and the Discriminator determines whether it is real or fake. This opposing relationship allows the Generator to learn to generate data that is closer to reality. (...omitted)

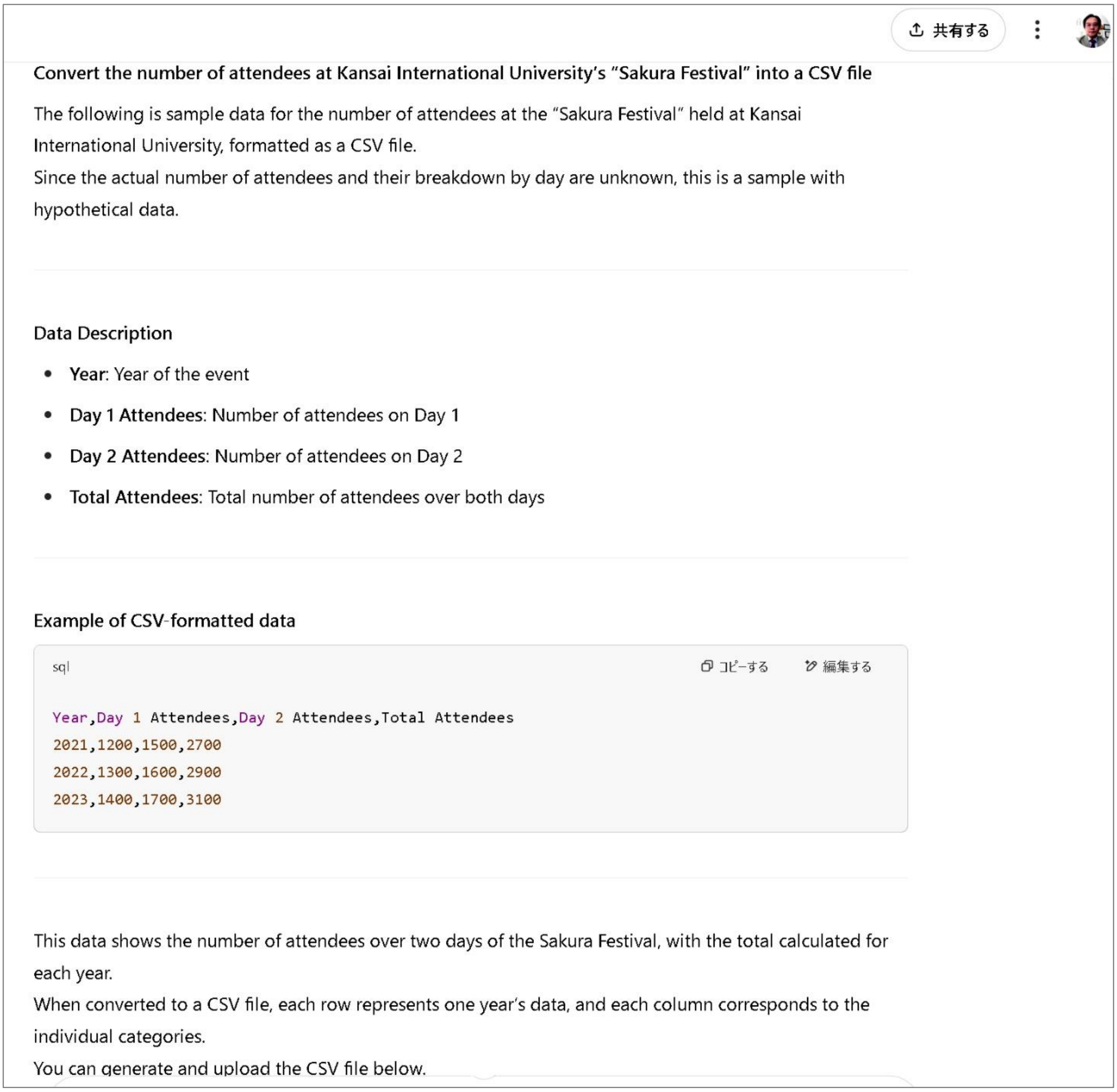

Figure 4. *Verification of the hallucination problem.*

3.2.4. Implementation in the Educational Environment

The AI teaching assistant system was implemented in the "Data Science" course as a supplementary tool to support both instructors and students. For instructors, the system aided with course preparation, material development, and assessment design. For students, it offered on-demand explanations, guided practice opportunities, and personalized feedback. The implementation process included:

1). Integration of course-specific materials into the system's knowledge base; 2). Training of instructors on effective system utilization; 3). Introduction of the system to students with clear guidelines on its capabilities and limitations; 4). Establishment of protocols for monitoring system performance and addressing any issues that arose.

3.3. Evaluation Methods and Criteria

The effectiveness of the AI teaching assistant system was evaluated through formative assessment, focusing on four key perspectives:

1. Teaching Preparation Efficiency: Assessing the system's impact on instructors' workflow and preparation time, including material development, assessment creation, and feedback provision

2. Active Learning Support: Evaluating the system's ability to facilitate student engagement, self-directed learning, and knowledge construction

3. Data Analysis Process Enhancement: Measuring the

system's effectiveness in supporting data analysis tasks, including data interpretation, visualization, and statistical analysis

4. Advanced Learning Activity Promotion: Assessing the system's contribution to higher-order learning activities, such as critical thinking, problem-solving, and creative application of concepts

By combining quantitative metrics and qualitative feedback, the evaluation provided a nuanced understanding of the system's educational impact, revealing its strengths and potential areas for improvement across teaching and learning processes. This comprehensive evaluation framework allowed for a nuanced understanding of the system's educational impact across different dimensions of the teaching and learning process, providing insights into both its strengths and areas for improvement.

4. Results and Discussion

4.1. Effectiveness of the AI Teaching Assistant System

The implementation of the "Data Science" AI teaching assistant system as a GPTs application demonstrated significant effectiveness across multiple dimensions of educational support. Through formative assessment, we evaluated the system's impact from four key perspectives: teaching preparation efficiency, active learning support, data analysis process enhancement, and advanced learning activity promotion. The following sections detail the findings in each of these areas.

4.1.1. Improving Teaching Preparation Efficiency

The DS-ASST developed in this research was confirmed to significantly enhance the efficiency of instructors' processes for course preparation and material development. Specifically, the following benefits were observed:

Content Generation and Refinement: The DS-ASST significantly streamlined content creation, helping instructors generate initial drafts of lecture materials and assessments. Instructors noted about 40%-time reduction in basic content development, enabling them to concentrate more on customization and quality improvement.

Diverse Explanation Generation: The DS-ASST generated multiple explanatory approaches for core concepts, offering diverse analogies, examples, and visualization suggestions. This versatility allowed instructors to select and refine explanations best suited to their students' needs, effectively making complex data science concepts more accessible.

Assessment Development Support: The system demonstrated utility in generating varied assessment items at different cognitive levels. It could produce not only knowledge-checking questions but also application-oriented problems and analytical scenarios, helping instructors develop more comprehensive assessment strategies that aligned with course learning objectives.

Feedback Template Creation: The AI system developed structured feedback templates for common student misconceptions, enabling more consistent and comprehensive responses. This approach allowed instructors to maintain high-quality feedback while significantly reducing the time spent on individualized assessments.

These efficiency gains did not merely reduce workload but fundamentally transformed how instructors allocated their time, shifting focus from routine content creation to higher-value pedagogical activities such as personalization, refinement, and strategic instructional design.

4.1.2. Supporting Active Learning

The AI teaching assistant system proved effective in supporting students' active engagement with course content and promoting self-directed learning behaviors:

On-Demand Concept Clarification: Students utilized the system to obtain immediate clarification on concepts they found challenging, with the system providing explanations tailored to their specific questions. This on-demand support reduced barriers to understanding and prevented the accumulation of knowledge gaps that might otherwise impede progress.

Guided Problem-Solving: The AI system guides problem-solving with prompts and hints, fostering independent thinking, rather than just providing answers.

Learning Path Personalization: The system demonstrated capability in suggesting personalized learning paths based on students' interactions, recommending supplementary resources or alternative explanatory approaches when students struggled with concepts. This adaptive support helped address the diverse learning needs within the student cohort.

Engagement with Complex Data: Students reported greater confidence in engaging with complex datasets when supported by the AI assistant, which could provide contextual information, suggest analytical approaches, and help interpret patterns or anomalies in the data. This support was particularly valuable for students with limited prior experience in data analysis.

The active learning support provided by the system aligned well with the course's objective of developing students' data thinking capabilities, encouraging exploration and inquiry rather than passive information reception.

4.1.3. Enhancing Data Analysis Processes

In the context of data science education, the system's ability to support data analysis processes was particularly significant:

Analysis Workflow Guidance: The system effectively guided students through structured data analysis workflows, from initial data exploration and cleaning to more sophisticated analytical techniques. This guidance helped students internalize systematic approaches to data analysis that reflect professional practice.

Code Assistance and Debugging: For programming components of the course, the system provided valuable assistance with code development and debugging. It could suggest code improvements, identify errors, and explain the logic behind different analytical approaches, supporting students' development of technical skills.

Interpretation Support: Beyond technical execution, the system assisted with the interpretation of analysis results, helping students connect statistical outputs to meaningful insights. This interpretive support was crucial for developing students' ability to derive actionable knowledge from data.

Visualization Recommendations: The system provided contextually appropriate visualization recommendations based on data characteristics and analytical objectives. These recommendations helped students select effective visual representations that accurately communicated their findings.

The enhancement of data analysis processes contributed significantly to students' development of practical data science skills, bridging the gap between theoretical understanding and applied capability.

4.1.4. Promoting Advanced Learning Activities

The AI teaching assistant system demonstrated value in supporting higher-order learning activities that extend beyond basic knowledge acquisition:

Critical Evaluation of AI Outputs: Students were encouraged to critically evaluate the system's responses, identifying potential inaccuracies or limitations. This meta-learning activity developed students' critical thinking skills while also fostering a nuanced understanding of AI capabilities and constraints.

Interdisciplinary Connection Facilitation: The system effectively helped students connect data science concepts to their primary fields of study, suggesting relevant applications and examples that bridged disciplinary boundaries. This interdisciplinary perspective was especially valuable in the liberal arts university context.

Research Question Formulation: The system supported students in developing sophisticated research questions that could be addressed through data analysis. By suggesting refinements and highlighting analytical possibilities, it helped students conceptualize meaningful inquiry projects.

Ethical Reasoning Development: When addressing topics related to data ethics and responsible AI use, the system facilitated nuanced discussions of ethical dilemmas, helping students develop reasoned positions on complex issues such as privacy, bias, and algorithmic transparency.

These advanced learning activities contributed to the development of higher-order thinking skills that transcend specific course content, preparing students for the complex challenges they will encounter in professional contexts.

4.1.5. Brief Statistical Summary

Small group observations revealed the following statistical results compared to the groups that did not use AI-assisted tools:

1. Instructor preparation time was reduced by 42% (from an average of 5.2 hours to 3.0 hours per week).
2. Student participation in discussions increased by 38% compared to the control group.
3. 80% of students reported improved confidence in applying data analysis techniques.
4. The complexity of questions increased by 25% in the AI-assisted group.
5. The completion rate of course evaluations increased from 70% to 85%.

However, at this stage, we are unable to quantitatively measure the learning outcomes of students across all student populations. We plan to design evaluation items and conduct rigorous measurements through future educational activities.

4.2. Technical Challenges and Solutions

The implementation of the AI teaching assistant system was not without challenges, particularly in the areas of prompt design and hallucination management:

Prompt Design Optimization: Initial prompt designs often yielded responses that were either too general or excessively technical for the target audience. Through iterative refinement, we developed structured prompts that incorporated specific pedagogical goals, student background information, and course context. This contextual enrichment significantly improved the educational relevance of the system's responses.

Hallucination Mitigation: To address the challenge of GenAI hallucinations, we implemented several strategies: 1). Knowledge base integration through RAG technology, anchoring responses in verified course materials; 2). Explicit uncertainty acknowledgment in system responses when venturing beyond its knowledge base; 3). Fact-checking protocols for instructors reviewing system-generated content; 4). Student education on critical evaluation of AI-generated information.

Technical Vocabulary Calibration: The system initially struggled with appropriately calibrating technical vocabulary to students' knowledge level. We addressed this by incorporating vocabulary guidelines into prompts and developing a progressive disclosure approach that introduced technical terms with accompanying explanations.

Response Length and Structure: Finding the optimal balance between comprehensive explanation and cognitive manageability required significant experimentation. We ultimately developed templates for different response types (concept explanations, procedural guidance, feedback) with appropriate structural elements such as summaries, step-by-step breakdowns, and visual suggestions.

These technical challenges and their solutions provided valuable insights into the effective design of AI educational tools, highlighting the importance of pedagogical principles in guiding technical implementation.

4.3. Comparison with Traditional Educational Approaches

The AI teaching assistant system demonstrated several distinctive advantages when compared to traditional educational approaches:

Temporal Flexibility: Unlike traditional office hours or synchronous support, the AI system provided 24/7 availability, allowing students to receive assistance at their optimal learning times and addressing questions as they arose rather than requiring them to be accumulated for scheduled interactions.

Scaling of Individualized Support: The system enabled a degree of individualized support that would be logistically challenging to provide through human instructors alone, especially in larger course sections. This scaling of personalization represents a significant advancement over traditional approaches to educational support.

Reduced Judgment Barriers: Some students reported greater comfort in asking "basic" questions to the AI system than to human instructors, fearing less judgment or perception of inadequacy. This reduced psychological barrier facilitated more comprehensive support for foundational understanding.

Consistency with Adaptability: The system maintained consistent quality and comprehensiveness in its responses while adapting to individual student needs—a balance that can be challenging for human instructors to maintain, particularly when managing multiple courses and responsibilities.

However, the comparison also revealed areas where traditional approaches retain advantages:

Emotional Intelligence and Motivation: Human instructors demonstrated superior ability to recognize emotional states, provide motivational support, and build rapport with students—elements that remained beyond the capabilities of the AI system.

Spontaneous Connection-Making: Experienced instructors excelled at making spontaneous connections to current events, student interests, or emerging research that the AI system could not match without specific programming.

Judgment in Intervention Timing: Human instructors demonstrated better judgment regarding when to intervene in student learning processes versus when to allow productive struggle, a nuanced capability that proved difficult to replicate in the AI system.

These comparative insights suggest that optimal educational outcomes may be achieved through thoughtful integration of AI systems and human instruction rather than replacement of one with the other.

4.4. Limitations of the Current Implementation

Despite its demonstrated effectiveness, the current implementation of the AI teaching assistant system has several limitations that warrant acknowledgment:

Domain Specificity: The system's knowledge base was optimized for the specific data science curriculum used in the course. Its effectiveness would likely be reduced if applied to other courses without substantial reconfiguration and knowledge base enhancement.

Assessment Limitations: While the system could generate assessment items and provide feedback on responses, it lacked sophisticated capabilities for evaluating open-ended work or creative problem-solving approaches that deviated from expected patterns.

Technical Requirements: The system's implementation required certain technical infrastructure and connectivity, potentially creating access barriers for students with limited technological resources or in areas with connectivity challenges.

Adaptation Lag: The system required periodic updates to incorporate new course materials, emerging research, or changes in the field of data science. This updating process introduced a lag in the system's ability to reflect the most current developments.

These limitations highlight areas for future development and refinement of AI teaching assistant systems, particularly as the underlying technologies continue to evolve.

4.5. Implications for Educational Paradigm Transformation

The findings from this implementation suggest several broader implications for educational paradigm transformation in the AI era:

Redefinition of Instructor Roles: The effective integration of AI teaching assistants may lead to a redefinition of instructor roles, with greater emphasis on higher-order instructional design, relationship building, and complex judgment tasks rather than routine information delivery or basic question answering.

Evolving Assessment Approaches: As AI systems become more capable of supporting learning processes, assessment approaches may need to evolve to emphasize uniquely human capabilities such as creative problem formulation, ethical reasoning, and collaborative innovation rather than information recall or procedural execution.

Blending of Formal and Informal Learning: The temporal flexibility and on-demand nature of AI support may accelerate the blending of formal and informal learning experiences, with students moving fluidly between structured course activities and self-directed exploration supported by AI guidance.

Metacognitive Skill Prioritization: The presence of powerful AI tools may increase the importance of metacognitive skills—the ability to plan, monitor, and evaluate one's own learning processes—as students navigate information-rich environments with sophisticated AI support.

Ethical Literacy Development: As AI systems become more integrated into educational experiences, the development of ethical literacy regarding AI use, including understanding of limitations, potential biases, and appropriate reliance, becomes an essential component of educational programs.

These implications suggest that the integration of AI teaching assistants represents not merely a technological enhancement of existing educational approaches but a potential catalyst for fundamental reconsideration of educational structures, practices, and priorities.

5. Conclusions

The rapid advancement of GenAI technology, particularly Large Language Models, has presented transformative possibilities for the field of education. This study developed and evaluated an AI teaching assistant system as a GPTs application specifically designed to support data science education. Through formative assessment, we demonstrated the system's effectiveness across multiple dimensions of educational support, from enhancing instructor efficiency to facilitating active learning among students.

5.1. Summary of Research Findings

Our research findings confirm that GenAI technology, when thoughtfully implemented in educational settings, can provide significant benefits across multiple aspects of the teaching and learning process. The DS-ASST demonstrated effectiveness from four key perspectives:

First, the system substantially improved teaching preparation efficiency, enabling instructors to allocate more time to high-value pedagogical activities rather than routine content creation. The AI assistant's ability to generate diverse explanations, assessment items, and feedback templates transformed instructors' workflow, allowing for greater focus on personalization and quality enhancement.

Second, the system effectively supported active learning by providing on-demand concept clarification, guided problem-solving assistance, personalized learning path suggestions, and support for engaging with complex data. These capabilities fostered self-directed learning behaviors and helped address the diverse learning needs within the student cohort.

Third, the system enhanced data analysis processes through structured workflow guidance, code assistance, interpretation support, and visualization recommendations. These features contributed significantly to students' development of practical data science skills, bridging the gap between theoretical understanding and applied capability.

Fourth, the system promoted advanced learning activities, including critical evaluation of AI outputs, interdisciplinary connection facilitation, research question formulation, and ethical reasoning development. These higher-order learning activities prepare students for the complex challenges they will encounter in professional contexts.

These findings suggest that GenAI technology has the potential to fundamentally transform educational paradigms, not merely as a technological enhancement of existing approaches but as a catalyst for reconsidering educational structures, practices, and priorities.

5.2. Educational Significance of GenAI Applications

The educational significance of GenAI applications extends beyond the specific context of data science education examined in this study. Our findings suggest several broader implications for educational practice:

The integration of AI teaching assistants enables a more personalized and adaptive learning experience, addressing the long-standing challenge of providing individualized support on a scale. This capability is particularly valuable in educational contexts with diverse student populations and varying levels of prior knowledge.

AI's flexible, on-demand support aligns with modern learning, extending beyond classrooms and integrating formal and informal education. The ability of AI systems to present multiple explanatory approaches for complex concepts addresses the challenge of cognitive diversity in educational settings. By offering various analogies, examples, and visualization options, AI assistants can help more students find entry points to challenging material that resonate with their existing knowledge structures and learning preferences.

Furthermore, the implementation of AI teaching assistants creates opportunities for meta-learning about AI itself, helping students develop critical evaluation skills and nuanced understanding of AI capabilities and limitations. This meta-learning becomes increasingly important as AI systems become more prevalent in professional and civic contexts.

5.3. Challenges Requiring Careful Consideration

Notwithstanding the promising outcomes demonstrated in this study, the integration of GenAI into educational environments necessitates careful consideration of several inherent challenges.

A primary challenge pertains to ensuring the veracity of AI-generated content, particularly within specialized or rapidly evolving knowledge domains. The phenomenon of AI 'hallucinations' mandates the implementation of robust mitigation strategies. These should include integration with verified knowledge bases, mechanisms for explicit uncertainty signaling, and rigorous fact-checking protocols.

Furthermore, the protection of data privacy emerges as a critical concern, especially given the potential involvement of sensitive student information in educational applications. Educational institutions are obligated to formulate explicit policies governing data collection, storage, and utilization, ensuring strict adherence to pertinent regulations and the preservation of student trust.

Ethical considerations extend beyond privacy, encompassing issues of equity, transparency, and appropriate technological reliance. Implementations must rigorously ensure that AI systems do not perpetuate or amplify extant biases. Moreover, the capabilities and limitations of these systems must be communicated transparently, and their role must be positioned to augment, rather than supplant, the development of essential human cognitive abilities.

Addressing practical aspects, significant challenges also arise concerning technical infrastructure, requisite teacher training, sustainable implementation models, and the prevention of academic dishonesty, such as students utilizing GenAI inappropriately for assignments like project reports. Educational institutions must comprehensively evaluate the resources necessary not only for initial deployment but also for ongoing maintenance, system updates, and continuous evaluation to ensure effective and responsible AI integration.

5.4. Future Directions

Based on the insights gained from this research, we identify several promising directions for future development and investigation:

First, there is a need for expanded instruction on appropriate LLMs usage, helping both educators and students develop the skills necessary to effectively leverage these tools while maintaining critical awareness of their limitations. This instruction should include practical guidance on prompt engineering, output evaluation, and appropriate task selection.

Second, the development of comprehensive privacy protection guidelines specifically tailored to educational AI implementations is essential. These guidelines should address not only technical security measures but also pedagogical considerations regarding what types of student data should be accessible to AI systems and how that access should be governed.

Third, deeper exploration of ethical considerations is needed, including preventing misuse of AI systems, eliminating bias in educational applications, and acknowledging system limitations transparently. This exploration should involve diverse stakeholders, including educators, students, administrators, and ethical specialists.

Fourth, longitudinal studies examining the long-term impact of AI teaching assistants on learning outcomes, skill development, and educational trajectories would provide valuable insights beyond the relatively short-term evaluation conducted in this study. Such research could help identify both immediate benefits and potential unintended consequences of AI integration in education.

Finally, investigation into hybrid models that optimally combine human and AI instruction represents a particularly promising direction. Rather than viewing AI as either a replacement for or supplement to human teaching, research should explore how the distinctive strengths of human educators and AI systems can be integrated to create educational experiences that exceed what either could provide independently.

5.5. Concluding Remarks

The integration of GenAI technology into educational settings represents a significant opportunity to enhance teaching and learning processes across multiple dimensions. Our development and evaluation of an AI teaching assistant system for data science education demonstrates the potential of such applications to improve efficiency, support active learning, enhance analytical processes, and promote advanced learning activities.

However, realizing this potential requires thoughtful implementation that addresses challenges related to accuracy, privacy, AI ethics education, and practical integration. By approaching these challenges proactively and continuing to refine our understanding of effective AI-enhanced educational practices, we can work toward educational environments that harness the capabilities of GenAI while remaining firmly grounded in core educational values and objectives.

As we navigate this transformative period in educational technology, ongoing dialogue between technologists, educators, students, and policymakers will be essential to ensure that AI implementation serves the fundamental goals of education: fostering understanding, developing capabilities, and preparing learners for meaningful participation in an increasingly complex world.

Abbreviations

LLMs	Large Language Models
GenAI	Generative AI
RAG	Retrieval-Augmented Generation
CoT	Chain of Thought
MoE	Mixture of Experts
RLHF	Reinforcement Learning from Human Feedback
FP8	8-bit Floating-Point number
DS-ASST	GPT-based AI Teaching Assistant App for Data Science Education

Funding

This work was supported by MEXT/JSPS KAKENHI Grant Numbers JP (22K02930).

Author Contributions

Zhihua Zhang is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Zhang, Z., et al. Development of Education Curriculum in the Data Science Area for a Liberal Arts University. *Towards a Collaborative Society Through Creative Learning,* In Springer Nature, 2023, Keane, T. Lewin, C., Brinda, T., & Bottino, R. (Eds.). https://doi.org/10.1007/978-3-031-43393-1_32

[2] Misa Tei. (2023). How should generative AI be used in education? Possibilities and challenges considered from various guidelines, Compass for SDGs & Society 5.0, Institute Business Environment Report, pp. 1-11. Available from: https://www.dlri.co.jp/files/ld/285589.pdf [Accessed 19 August 2024].

[3] UNESCO, 「AI and education: guidance for policy-makers」, Published in 2021 by the United Nations Educational, Scientific and Cultural Organization, 2021/12/26, https://doi.org/10.54675/PCSP7350 [Accessed 2 May 2024].

[4] The U.S. Department of Education Office of Educational Technology, Artificial Intelligence and the Future of Teaching and Learning. Insights and Recommendations. Available from: https://www2.ed.gov/documents/ai-report/ai-report.pdf [Accessed 15 August 2024].

[5] The U.S. Department of Education Office of Educational Technology, Designing for Education with Artificial Intelligence: An Essential Guide for Developers. Available from: https://tech.ed.gov/files/2024/07/Designing-for-Education-with-Artificial-Intelligence-An-Essential-Guide-for-Developers.pdf [Accessed 12 December 2024].

[6] Ministry of Education, Culture, Sports, Science and Technology of Japan, Provisional Guidelines for the Use of Generative AI in Primary and Secondary Education. Available from: https://www.mext.go.jp/content/20230718-mtx_syoto02-000031167_011.pdf 1-24, 2023. [Accessed 26 December 2024].

[7] Zhang, Z. Practical application of on-demand lessons in data science education for liberal arts students at private universities, Proceedings of the 85th National Conference of the Information Processing Society of Japan, Vol. 4, 97-112, 2023.

[8] Zhang, Z. et al. The Course Design of Basic Data Science Taking into Account Both Face-To-Face and On-Demand Teaching and Effect Analysis, the 2023 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC). https://doi.org/10.23919/PNC58718.2023.10314978

[9] Genshiro Kitagawa, Akimichi Takemura. Data Science as Liberal Arts, 2nd. Tokyo: Kodansha; 2022, 1-229.

[10] OpenAI. GPT-4 Technical Report, Available from: https://cdn.openai.com/papers/gpt-4.pdf 2023, 1-100. [Accessed 20 April 2024].

[11] A Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, Denny Zhou, Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, 2022, 1-43, https://arxiv.org/pdf/2201.11903

[12] AI achieves silver-medal standard solving International Mathematical Olympiad problems, Available from: https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/ 2024. [Accessed 19 October 2024].

[13] Japan Inter-University Consortium for Mathematics, Data Science and AI Education, Mathematics, Data Science, AI (Literacy Level) Model Curriculum - Cultivating Data Thinking - (Revised February 22, 2024), http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy_20240222.pdf, 2024 [Accessed 20 September 2024].

[14] Japan Inter-University Consortium for Mathematics, Data Science and AI Education, Mathematics, Data Science, AI (Applied Basic Level) Model Curriculum - Cultivating Data Thinking - (Revised February 22, 2024), http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_ouyoukiso_20240222.pdf 2024. [Accessed 25 December 2024].

[15] Vaswani et al., Attention Is All You Need [Submitted on 12 Jun 2017 (v1), last revised 2 Aug 2023], arXiv:1706.03762v7, https://doi.org/10.48550/arXiv.1706.03762 [Accessed 15 April 2024].