International Journal of Research in Advent Technology, Vol.14, No.1, March 2026
E-ISSN: 2321-9637
Available online at www.ijrat.org

From Large Language Models to Retrieval-Augmented Generation: A Systematic Review of AI-Powered Educational Systems

Sumedha Ayachit and Rajneeshkaur Sachdeo

Abstract. Artificial Intelligence (AI) is rapidly restructuring education by enabling personalized learning, automated assessment, and intelligent tutoring systems. Recent developments in Large Language Models (LLMs) are demonstrating strong capabilities in natural language understanding and content generation, enabling applications such as conversational tutoring, question generation, and feedback delivery. However, LLM-based educational systems suffer from critical limitations such as hallucinations, outdated knowledge, lack of curriculum alignment, and limited transparency. To address these challenges, Retrieval-Augmented Generation (RAG) is a promising paradigm that integrates external knowledge retrieval with generative models, enabling context-aware, accurate, and domain-specific responses. This paper presents an organized and inclusive review of AI-powered educational systems, tracking the evolution from traditional AI and standalone LLMs to RAG-enhanced architectures. The review analyzes recent research on RAG-based educational applications, including intelligent tutoring systems, virtual assistants, content generation, assessment, and feedback mechanisms. Finally, the study identifies critical research gaps and outlines future directions for developing curriculum-aware, adaptive, and trustworthy RAG-based educational systems.

Keywords: Retrieval-Augmented Generation (RAG); AI in Education (AIED); Large Language Models (LLM); Intelligent Tutoring Systems; Personalized Learning; Curriculum-Aware AI

Manuscript received February 22, 2026; revised March 12, 2026 and published on March 30, 2026
Sumedha Ayachit, Department of Computer Science and Engineering, MIT School of Computing, MIT ADT University, Pune, India. sumedha.ayachit@mituniversity.edu.in
Rajneeshkaur Sachdeo, Department of Computer Science and Engineering, MIT School of Computing, MIT ADT University, Pune, India. rajni.sachdeo@mituniversity.edu.in

I. INTRODUCTION

The education system has evolved rapidly in recent years as there is an arrival of AI. The use of AI in the education system enhances the traditional education system by providing personalized learning, adaptability, real-time feedback, efficient assessment, effective content generation, etc. Paper [1] states that traditional education methods provide structured, uniform learning experiences, while AI methods offer customization and flexibility, adapting to the needs of individual learners. While traditional education relies on standardized curricula and teacher-led instruction, AI-driven approaches offer personalized, adaptive learning experiences that cater to individual student needs. Large Language Models (LLMs) are a type of generative AI that excels at processing, understanding, and generating human language. LLM use machine learning algorithms to work on a large dataset. Large Language Models (LLM) such as GPT-4, PaLM, and BERT have shown significant potential to enhance the teaching-learning process through automated tutoring, content generation, question answering, and personalized feedback. However, despite their fluency and adaptability, LLMs often suffer from significant limitations: they may generate factually incorrect information (hallucination), lack transparency, and are typically disconnected from up-to-date or curriculum-specific knowledge. The integration of AI-generated education, particularly through Retrieval-Augmented Generation (RAG), presents a transformative alternative to traditional educational methods.

In recent days, the interest in RAG for education has been increasing rapidly. This review focuses on the paradigm shift from standalone Large Language Model–based educational systems to Retrieval-Augmented Generation frameworks, which ground generative outputs in external knowledge sources to improve reliability and pedagogical relevance. This paper is structured as follows. Section 2 covers a brief overview of AI-powered education system. Section 3 introduces the RAG framework. Section 4 focuses on a survey of

10.32622/ijrat.1401202613

RAG in education for various purposes to enhance student learning and engagement. Section V focuses on challenges with RAG and future direction.

II. AI-POWERED EDUCATION SYSTEM

The traditional education system is enhanced with Artificial Intelligence (AI) in education (AIED). [4] explores the systematic literature review of AI in education and find the key areas where AI is applied in education as follows.

1.Adaptive Learning & Personalized Tutoring (40% of studies): Includes intelligent tutoring systems (ITS) and adaptive hypermedia learning systems (AHLS) for personalized feedback, exercises, and learning paths.

2.Intelligent Assessment & Management (24.8%): Covers automatic grading, teaching evaluation, and learning management systems (LMS) for collaborative learning and resource management.

3.Profiling & Prediction (20%): Focuses on predicting academic outcomes, drop-out risks, and learner modelling using data mining and analytics.

4.Emerging Products (15.2%): Includes educational robots (e.g., chatbots) and immersive technologies like virtual reality (VR) and augmented reality (AR).

A comparative analysis of AI based education and traditional education from [1] states that AI-based education surpasses in personalization, accessibility, scalability, and efficiency, while traditional methods focus on teacher-student interaction, cultural integration, and critical thinking. Challenges for AI include privacy concerns and reliance on technology. [1] also propose a “Balanced Approach” that combines AI-driven tools with traditional methods to create a more inclusive and effective learning environment, leveraging the strengths of both approaches.

A. Traditional AI in Education

Initial AI based education, we can call it “Traditional AI” focused on rule-based expert systems and simpler information-retrieval methods for answering student questions. Intelligent tutoring systems (ITS) have been an active research domain since the 1990s, with systems like AutoTutor, Cognitive Tutor, and ALEKS demonstrating how AI can simulate human tutoring by adapting to student knowledge levels. These systems are rule-based systems.

B. Generative AI in Education

The arrival of large language models (LLMs) has expanded these possibilities, allows more natural interactions and personalized support. With the development of large-scale generative models (e.g., GPT-3, GPT-4, PaLM), educational researchers began exploring how these models could support tutoring, essay feedback, question generation, and conversation-based learning. Studies have demonstrated their potential to improve engagement and responsiveness. However, generative models are prone to hallucinations and fail to consistently ground their answers in verifiable content, raising concerns about factual correctness and pedagogical alignment. [5] provides survey on the applications, challenges, and future developments of Large Language Models (LLMs) in the field of education, highlighting their potential to enhance the teaching and learning process. Table 1. shows application and case study of LLM in education.

Table 1. Application and Case study of LLM in education

Sr. No.	Applications of LLM in Education	Sr. No.	LLM generated educational content	Sr. No.	Case Study
1	Dynamic Content Generation	1	Writing Prompts	1	Khan Academy's AI Tutor
2	Interactive Tutoring	2	Explanatory Diagram	2	Socratic by Google
3	Personalized Learning Paths	3	Summaries and Study Guides	3	Carnegie Learning's MATHia
4	Adaptive Assessment	4	Quiz Questions and Answers		

III. FROM LARGE LANGUAGE MODELS TO RETRIEVAL AUGMENTED GENERATION

In an introduction, we discussed that the “traditional educational system” is now shifted to “AI-powered

10.32622/ijrat.1401202613

education system" with the inclusion of generative AI models like LLM. To improve the performance of AI-powered education system in future, LLM paradigm is shifted to RAG framework. A survey of Retrieval Augmented Generation for natural language processing and introduces each component of RAG, including details about the retriever from building to querying [3]. Combining retrieval and generation in a single framework addresses many limitations of stand-alone LLMs. RAG architectures provide the ability to condition generative outputs on retrieved, trusted sources, improving factual accuracy and curriculum alignment. Integrating RAG framework with LLM enhance the capability of LLM by resolving challenges with LLM. Following are the challenges with LLM:

1. Hallucinations and Factual Inaccuracies- Generating false information or misinterpreting historical events

2. Context Limitations- limited context window. Lack of domain specific knowledge

3. Lack of Real-Time Information- LLMs are trained on a static dataset. Out-dated knowledge

4. Ambiguity and Vague Queries- Difficult for LLM to understand and re-spond accurately to ambiguous or vague user queries.

5. Bias and Fairness- Reflects biases present in training data

6. Scalability and Cost- Expensive to use LLM

RAG is an AI framework that combines both retrieval and generation to improve the generated output quality as well as accuracy. [8] white paper describes RAG architecture as below which shows RAG work flow from user prompt to user response. Figure 1 shows the architecture of RAG framework [8]

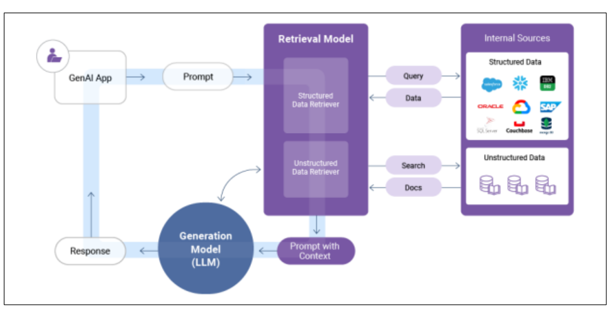

Figure1. Architecture of RAG [8]

From Figure1 working of RAG framework is given below:

1. When the user enters a prompt with a query, data retrieval model is triggered to access the company's internal sources, as required.
2. The retrieval model queries the company's internal structured data (enterprise systems) and unstructured data (knowledge bases)
3. The retrieval model enhances the user's original query with relevant context and passes this enriched prompt to LLM API for generation.
4.The LLM uses this improved prompt to generate a better accurate and relevant response, which is then provided to the user.

RAG framework is combined with an existing LLM to improve its performance. Retrieval-Augmented Generation (RAG) enhances LLM accuracy by dynamically [15] fetching data from external, trusted knowledge bases during inference. By grounding responses in retrieved, up-to-date evidence, RAG significantly reduces hallucinations, provides source attribution, and improves contextual relevance without needing to retrain models [7].

Survey of RAG in Education:

Retrieval-Augmented Generation (RAG) merges the generative power of LLMs with precise, external document retrieval to deliver accurate, context-aware, and up-to-date responses. It enhances AI reliability by grounding model outputs in trusted, non-parametric knowledge sources, reducing hallucinations and mitigating the need for frequent retraining. In educational contexts, this integration allows AI systems to generate contextually accurate, up-to-date, and domain-specific responses, overcoming the limitations of static, pre-trained models. This survey systematically reviews recent peer-reviewed studies that apply Retrieval-Augmented Generation to educational systems, focusing on methodology, application domains, performance outcomes, and limitations. Table 2 provides a literature survey of RAG in education.

Table 2. Survey of RAG in Education

10.32622/ijrat.1401202613

Sr. No.	Paper and author details	Methodology	Application Area	Key Findings	Limitations
1	[9]	KG-RAG + LLM	Intelligent Tutoring Systems (ITS)	Personalized answers improved learning outcomes	The scope is limited to science and math subjects.
2	[10]	RAG Transformer architecture-based LLM models Meta-llama/Llama-2-7b-chat-hf and Mistralai/Mistral-7B-Instruct-v0.2 [10]	Virtual Assistant	**1.** Addressing the information access challenge 2. Improving accuracy with RAG performance.	1. Computational constraints: The research relied on Google Colab's free GPU 2. Knowledge base dependency: The accuracy of the virtual assistant is heavily dependent on an extensive and up-to-date knowledge base.
3.	[11]	1. RAG+ open-source tools. 2. Evaluated Three leading 7B-parameter open-source Chinese Large Language Models (LLMs) from the Model Scope community [11]	Teacher education and training	1.Model accurately grasp and reproduce subject matter content.	1. Limited data set 2. In-depth analysis generated by model is relatively low
4.	[12]	RAG+ VERTEX AI Palm-2 mode	Enhancing information retrieval and support in an academic environment. [12]	1.Method significantly improve the accuracy and reduce hallucinations in responses from a virtual assistant. [12] 2. Virtual assistant's response accuracy is increased to 87.73% as relevant contextual information is retrieved.	1. Generalization issue: Affect accuracy of virtual assistant when input context is specific. 2. Bias in specialized topics: Virtual assistant exhibited biases in its response on some specialized topics where training data lacked neutral or diverse perspectives
5.	[13]	1.RAG+ Langchain 2.Video Generation Pipeline -	Interactive Virtual Assistant (IVA) supports students and educators in	1. Transform traditional classroom teaching to a	Delays due to network latency and high bandwidth usage for video streaming, which limit

10.32622/ijrat.1401202613

		consists of a Wav2Lip model for lip-syncing video and a GFPGAN model for face-specific restoration (enhancement	bridging the shortcomings of teacher-student interactions in an e-learning environment	virtual learning medium.	its feasibility for widespread internet-based deployment
6	[16]	RAG, Technology Acceptance Model (TAM) to assess system acceptance, examined learning outcomes, and explored the influence of general self-efficacy on system acceptance and OwlMentor use. [16]	University students	OwlMentor is an AI-powered learning environment designed to assist university students in comprehending scientific texts. [16] Implementation and evaluation of OwlMentor [16]	1.Sample size is small only 16 participants 2. not include specific measures of scientific text comprehension
7	[17]	Automatic Short Answer Grading (ASAG): ASAG is the task of automatically assigning appropriate scores to open-ended student responses. [17] Three-Dimensional Learning RAG and Domain-specific knowledge	Science Education new frameworks for K-12 science education. [17]	1.Assessing Three-Dimensional Learning [17] Increase emphasis on analyzing students' multidimensional comprehension of scientific understanding, including Disciplinary core ideas (DCIs), science and engineering practices (SEPs), and crosscutting concepts (CCCs). [17] 2. Providing Timely Feedback to Students 3. Enhancing Grading Efficiency:	1. Scope Limitation: Only for science education, not for other subjects 2. Lack of Empirical Evaluation: empirical evaluation is limited to a single dataset of student responses. Testing on diverse datasets would be needed to validate the generalizability of the proposed approach. 3.Incomplete Analysis of In-Context Learning 4. Lack of Comparison to Human Grading
8	[18]	1.ASAS-F-Z: Zero-shot ASAS-F 2. ASAS-F-Opt: Automatic Few-shot Optimization with DSPy	Can be applied across various educational levels, from K-12 to higher education. The ability to automate the grading process	1.Scoring Performance Comparison 2. Feedback Quality Evaluation	1. Lack of Context-specific Evaluation The paper does not provide any information about the specific educational contexts or student demographics the ASAS-F system is intended for.

10.32622/ijrat.1401202613

		3. ASAS-F-RAG: Few-shot ASAS-F using Retrieval-Augmented Generation (RAG) ColBERT retrieval model	and generate elaborated feedback can be beneficial for students of all ages who are required to answer short-form questions as part of their assessments.		2. Subjectivity in Feedback Evaluation 3. Difficulty in Calibrating Large Language Models 4. Limitations of Automatic Optimization
	[19]	1.Direct Automated Feedback Delivery (DAFeeD) [19] 2. open-source reference implementation called Athena, connected to the learning platform Artemis [19]	University Students from computer science, information systems (Software Engineering Course)	1. Comfort with Feedback Source 2. Perceived Effectiveness 3. Usability and Helpfulness	Internal Validity External Validity Construct validity

A survey paper on RAG Chatboats for educational purposes. The author identified 47 research papers from bibliographic data access platforms: Scopus, Web of Science, and Google Scholar [14]. [15] represents a detailed survey of RAG's application in the educational domain. A summary of RAG application in education is extracted from the paper [15]. Figure 2 is the pictorial representation of different uses of RAG in education such as Interactive learning systems, content generation, student assessment, and feedback etc.

10.32622/ijrat.1401202613

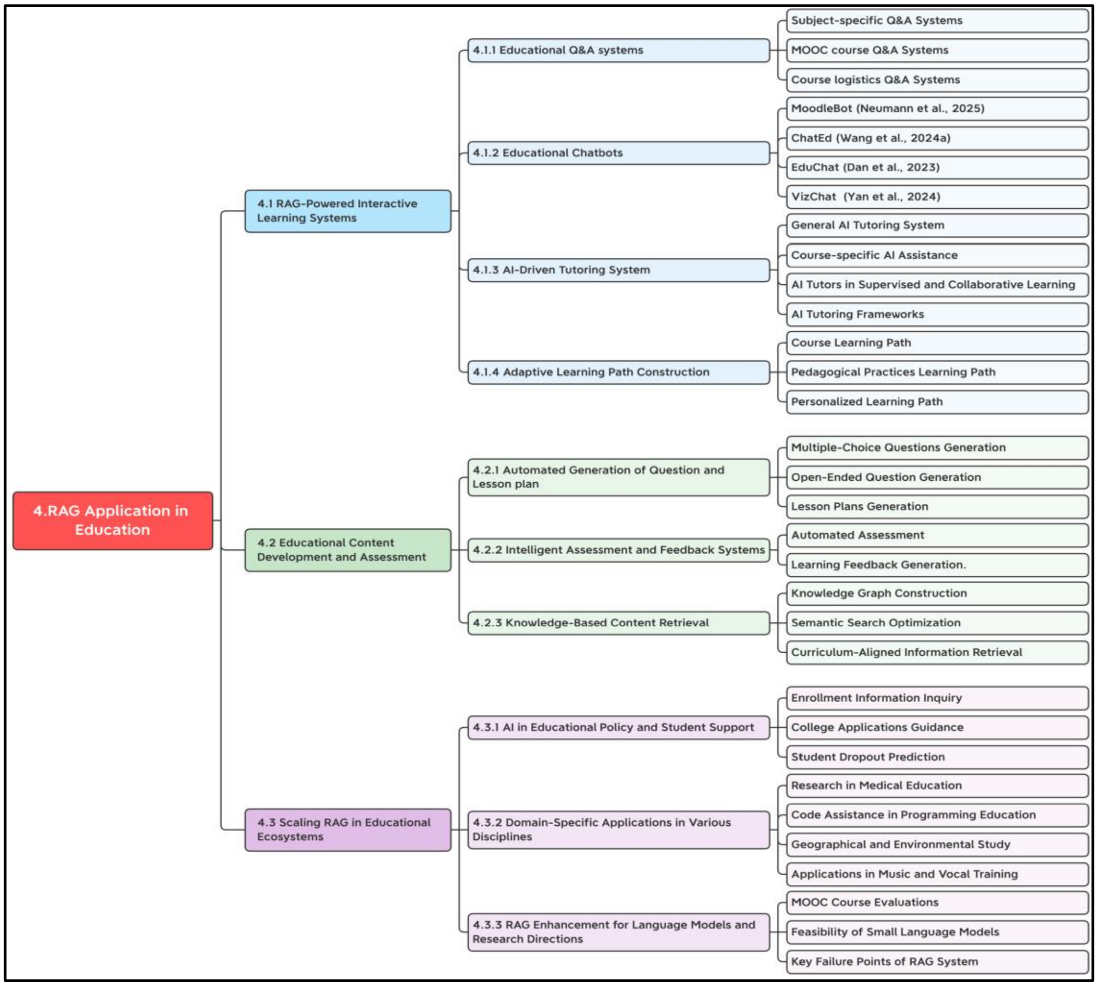

Fig. 2. RAG application in Education [15]

IV. DISCUSSION

RAG has demonstrated significant potential in educational applications by improving the accuracy and contextual relevance of AI-generated responses [15]. This section examines the challenges that occurs during use of RAG in education [15] and provides future direction for improvements.

A. Challenges in applying RAG in education

1. Curriculum Alignment and Contextualization-

One of the most significant challenges is aligning retrieved content with national or regional curricula. General-purpose RAG systems may pull accurate but contextually irrelevant or grade-inappropriate material. Educational applications require fine-grained control over topic scope, age- age-appropriateness, and pedagogical intent, which remains difficult with generic retrieval strategies.

2. Factual Accuracy and Hallucination Risks

RAG models can also hallucinate the contents if they generate the information that is not present in the retrieved documents. Such a situation is critical in education , where misinformation can deceive learners.

3. Multilingual and Multimodal Support

Educational content is often found in multiple languages. Contents also take various formats, not only restricted to text, but also include diagrams (images), videos, and equations. Most current RAG systems are text-centric and trained in English language. Supporting regional languages, mathematical notation, or subject-specific visual

10.32622/ijrat.1401202613

content (e.g., chemical structures or physics diagrams) is an ongoing challenge.

4. Student Personalization and Adaptivity

   Personalized learning is a cornerstone of modern education, but RAG systems struggle to adapt answers based on a learner's prior knowledge, learning style, or misconceptions.

B. Future Directions

RAG-based education can focus on curriculum-aware and domain-specific systems that align with established educational standards such as NCERT, CBSE, and IB, that ensures subject- and grade-level relevance. The development of multilingual and culturally inclusive RAG models is equally critical to support learners in linguistically diverse and multicultural contexts. Finally, advancing personalized and adaptive RAG through learner profiling, feedback loops, and ethical user modeling can promote tailored, safe, and effective learning experiences.

V. CONCLUSION

The paper provides a systematic review of AI-powered educational systems with its evolution from traditional artificial intelligence and standalone large language models to Retrieval-Augmented Generation–based architectures. LLMs have significantly enhanced educational content generation and interaction, but its limitations in factual accuracy, curriculum relevance, and transparency restrict their reliability in formal learning environments. By integrating external knowledge retrieval, RAG frameworks address these limitations and enable more accurate, personalized, and context-aware educational applications, including intelligent tutoring systems, virtual assistants, assessment tools, and feedback mechanisms. The paper also identified persistent challenges related to curriculum alignment, multilingual and multimodal support, personalization, explainability, and pedagogical evaluation.

Future research should focus on curriculum-aware retrievers, learner-adaptive RAG pipelines, and robust evaluation frameworks grounded in educational outcomes rather than traditional NLP metrics. Advancements in these areas will be critical for developing trustworthy, scalable, and inclusive AI-powered educational systems. Overall, this review explores how the shift from LLM to Retrieval-Augmented Generation is developing the next generation of intelligent and reliable educational technologies.

REFERENCES

1. Susan Rochelle ,Dr. Sushith, "Exploring the AI Era: A Comparative Analysis of AI-Driven Education and Traditional Teaching Methods", International Journal for Multidisciplinary Research (IJFMR) E-ISSN: 2582-2160, Volume 6, Issue 4, July-August 2024
2. Shervin Minaee, Tomas Mikolov, Narjes Nikzad, Meysam Chenaghlu, Richard Socher, Xavier Amatriain, Jianfeng Gao, "Large Language Models: A Survey", arXiv:2402.06196v2 [cs.CL] 20 Feb 2024
3. Shangyu Wu City University of Hong Kong, MBZUAI Ying Xiong MBZUAI Yufei Cui McGill University, Mila Haolun Wu McGill University, Mila Can Chen McGill University, Mila, Ye Yuan McGill University, Mila Lianming Huang City University of Hong Kong, Xue Liu McGill University, MilaTei-Wei Kuo National Taiwan University, Nan Guan City University of Hong Kong Chun Jason Xue MBZUA, "Retrieval-Augmented Generation for Natural Language Processing: A Survey"
4. Shan Wang , Fang Wang , Zhen Zhu , Jingxuan Wang , Tam Tran , Zhao Du," Artificial intelligence in education: A systematic literature review", https://www.elsevier.com/locate/eswa , https://doi.org/10.1016/j.eswa.2024.124167
5. Hanyi Xua, Wensheng Gana, Zhenlian Qib, Jiayang Wua and Philip S. Yuc, "Large Language Models for Education: A Survey", arXiv:2405.13001v1 [cs.CL] 12 May 2024
6. Hao Qin, "Transforming Education with Large Language Models: Opportunities, Challenges, and Ethical Considerations", http://dx.doi.org/10.13140/RG.2.2.16976.52488 ,August 2024
7. Muhammad Arslana, Hussam Ghanema, Saba Munawarb and Christophe Cruza ," A Survey on RAG with LLMs", Procedia Computer Science 246 (2024) 3781–3790
8. 2024 K2view White Paper. K2view and Micro-Database are trademarks of K2view. "What is Retrieval Augmented Generation? The Complete Guide"
9. Chenxi Dong, Yimin Yuan, Kan Chen, Shupei Cheng, Chujie Wen, "How to Build an Adaptive AI Tutor for Any Course Using Knowledge Graph-Enhanced Retrieval-Augmented Generation (KG-RAG)", https://doi.org/10.48550/arXiv.2311.17696 (2023)
10. Umair Hasan Khana , Muneeb Hasan Khana , Rashid Ali, "Large Language Model based Educational Virtual Assistant using RAG Framework", Science Direct, Procedia Computer Science 252 (2025) 905–911
11. Ke Fang, Ci Tang & Jing Wang, "Evaluating simulated teaching audio for teacher trainees using RAG and local LLMs", 1 | https://doi.org/10.1038/s41598-025-87898-5, 1 | www.nature.com/scientificreports
12. Zlatan Mori´c , Leo Mrši´c, , Mario Filjak , Goran Dambi´c , "Integrating a Virtual Assistant by Using the RAG Method and VERTEX AI Framework at Algebra University", Appl. Sci. 2024, 14, 10748. https://doi.org/10.3390/app142210748. https://www.mdpi.com/journal/applsci
13. Shawn Abraham, Vinodh Ewards, Sebastian Terence, "Interactive Video Virtual Assistant Framework with Retrieval Augmented Generation for E-Learning", DOI: 10.1109/ICAAIC60222.2024.10575255
14. Jakub Swacha, Michał Gracel, "Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications", https://doi.org/10.3390/app15084234, Appl. Sci. 2025, 15, 4234
15. Zongxi Li ,Zijian Wang , Weiming Wang , Kevin Hung , Haoran Xie ,Fu Lee Wang , "Retrieval-augmented generation for educational application: A systematic survey", https://doi.org/10.1016/j.caeai.2025.100417,

10.32622/ijrat.1401202613

www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

16. Dominik Thüs, Sarah Malone and Roland Brünken, "Exploring generative AI in higher education: a RAG system to enhance student engagement with scientific literature.", Front. Psychol Frontiers in Psychology 15:1474892. doi: 10.3389/fpsyg.2024.1474892
17. Yucheng Chu, Peng He, Hang Li, Haoyu Han, Kaiqi Yang, Yu Xue, Tingting Li, Joseph Krajcik, and Jiliang Tang. "Enhancing LLM-Based Short Answer Grading with Retrieval-Augmented Generation", Proceedings of the 18th International Conference on Educational Data Mining, Palermo, Italy, July, 2025, pp. 396–402. International Educational Data Mining Society (2025). https://doi.org/10.5281/zenodo.15870304
18. MENNA FATEEN, BO WANG, AND TSUNENORI MINE 2, "Beyond Scores: A Modular RAG-Based System for Automatic Short Answer Scoring With Feedback". received 30 September 2024, accepted 21 November 2024, date of publication 29 November 2024, date of current version 17 December 2024. Digital Object Identifier 10.1109/ACCESS.2024.3508747
19. Maximilian Sölch, Felix T.J. Dietrich, and Stephan Krusche. 2025. "Direct Automated Feedback Delivery for Student Submissions based on LLMs." In 33rd ACM International Conference on the Foundations of Software Engineering (FSE Companion '25), June 23–28, 2025, Trondheim, Norway. ACM, New York . https://doi.org/10.1145/3696630.3727247

10.32622/ijrat.1401202613