INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 10 ISSUE: 04 | APRIL - 2026 SJIF RATING: 8.659 ISSN: 2582-3930

AI-Powered Research Assistant System Using Retrieval-Augmented Generation

Shekhar Pandey[1], Janvi Panchal[2], Vishal Gupta[3], Armaan Vala[4], Sachin More[5]

[1,2,3,4] Student, Department of Artificial Intelligence and Machine Learning
Universal College of Engineering, Mumbai, India
[5]Assistant Professor, Department of Artificial Intelligence and Machine Learning
Universal College of Engineering, Mumbai, India

Abstract*—The rapid growth of scientific publications has made it increasingly difficult for researchers to efficiently discover and analyse relevant academic content. Traditional information retrieval systems focus primarily on keyword-based search, while modern AI systems emphasize text generation, often lacking reliable grounding in scholarly sources. To address these limitations, this paper proposes an AI-powered research assistant system that integrates semantic information retrieval techniques with Retrieval-Augmented Generation (RAG). The proposed system enables users to input natural language queries and retrieves relevant research papers using similarity-based retrieval methods. The retrieved content is then processed using generative AI models to produce concise summaries, extract key insights, and provide citation-aware responses. By combining retrieval mechanisms with generation models, the system reduces hallucination and improves the relevance and reliability of outputs. Furthermore, the system incorporates concepts from recommendation systems and natural language processing to enhance search efficiency and user interaction. Experimental observations indicate that the proposed approach improves research accessibility, reduces the time required for literature review, and provides a unified platform for academic knowledge discovery. This work demonstrates the potential of integrating retrieval and generation techniques for building intelligent research support systems.

***Keywords*— AI Research Assistant, Retrieval-Augmented Generation (RAG), Natural Language Processing, Information Retrieval, Academic Knowledge Discovery**

I. INTRODUCTION

The rapid growth of digital academic publications has created significant challenges for researchers in efficiently discovering and analysing relevant information. Traditional information retrieval systems primarily rely on keyword-based techniques, which often result in large volumes of data that require manual filtering and analysis [24], [25], [21]. Early research in information retrieval laid the foundation for ranking and retrieving relevant documents, but these approaches lacked contextual understanding and semantic interpretation.

Recent advancements in Natural Language Processing (NLP) have significantly improved the ability of systems to understand and process textual information. Transformer-based models such as BERT enable contextual understanding of queries and documents, improving semantic search and retrieval performance [13], [19], [20]. These developments have enhanced the accuracy of information retrieval systems and enabled more efficient access to scholarly content.

In addition, research paper recommendation systems have been developed to improve literature discovery by analysing user behaviour and document similarity. The base paper [1] introduces a recommendation approach that considers users' information-seeking behaviour, allowing systems to provide more relevant and personalized research suggestions.

Other studies further explore classification and recommendation techniques to enhance research paper discovery and organization [4], [7].

More recently, Retrieval-Augmented Generation (RAG) has emerged as a powerful technique that combines document retrieval with generative AI models. By grounding generated responses in retrieved documents, RAG improves accuracy and reduces hallucination in AI-generated outputs [5], [10], [11]. These techniques enable systems to not only retrieve relevant documents but also generate meaningful summaries and insights.

Furthermore, AI-powered research assistant systems have been proposed to integrate retrieval, recommendation, and content generation into a unified platform. These systems assist researchers in literature review, knowledge extraction, and scientific discovery, improving research efficiency and reducing manual effort [2], [3], [6].

Therefore, this paper proposes an AI-powered research assistant system that integrates semantic information retrieval, recommendation techniques, and Retrieval-Augmented Generation. The system enables users to search for research papers, generate summaries, extract citations, and interact with academic content in an efficient and intuitive manner, thereby improving research productivity and knowledge discovery.

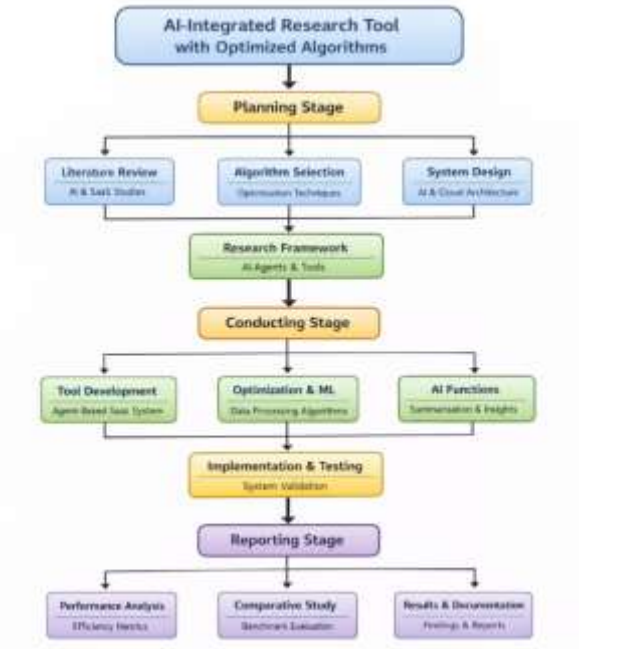

Fig. 1. Systematic Review Steps

II. METHOD

A. Planning Stage

The planning stage focuses on identifying the limitations of existing research systems and defining the scope of an AI-powered research assistant. A review of prior studies in information retrieval, recommendation systems, and AI-based research tools was conducted to understand current approaches and challenges. Traditional retrieval systems provide access to large datasets but often lack contextual understanding and intelligent filtering [21], [14].

Research on recommendation systems highlights the importance of user behaviour and document similarity in improving literature discovery [7], [17]. The base paper [1] further enhances this by considering users' information-seeking behaviour for more relevant recommendations. Additionally, advancements in NLP and transformer-based models have improved semantic understanding of queries and documents [13], [19].

Based on these insights, the proposed system is designed to integrate semantic search, recommendation techniques, and generative AI to support efficient research and knowledge discovery.

B. Conducting Stage

During the conducting stage, the system is developed by integrating multiple components such as data retrieval, semantic analysis, recommendation mechanisms, and content generation. Information retrieval techniques are used to identify relevant research papers based on user queries and similarity measures [24], [25].

Semantic understanding is achieved using embedding-based approaches and transformer models, improving the relevance of search results [13], [16]. Recommendation techniques are incorporated to suggest related papers based on content similarity and user interaction [1], [7].

Furthermore, Retrieval-Augmented Generation (RAG) is applied to generate summaries and extract insights from retrieved documents, enhancing the accuracy of responses [8], [10], [11]. These components work together to provide an efficient and integrated research platform.

TABLE I. LITERATURE SELECTION

Database	**Result**
ScienceDirect (Elsevier Journals)	8
Springer (Books & Conference Proceedings)	6
IEEE Xplore	4
Google Scholar / Other Indexed Journals	8

C. Reporting Stage

In the reporting stage, the system is evaluated based on its ability to support research activities effectively. The evaluation focuses on retrieval accuracy, recommendation relevance, usability, and overall system performance. Comparative analysis with existing systems is conducted to highlight improvements in efficiency and knowledge discovery [4], [12].

The results indicate that the system reduces the time required for literature review and improves access to relevant research information. AI-driven approaches enable efficient data processing and support better decision-making in academic research [18], [22].

The findings are presented through system analysis and performance observations, demonstrating the effectiveness of integrating retrieval, recommendation, and generative AI in a unified research assistant system.

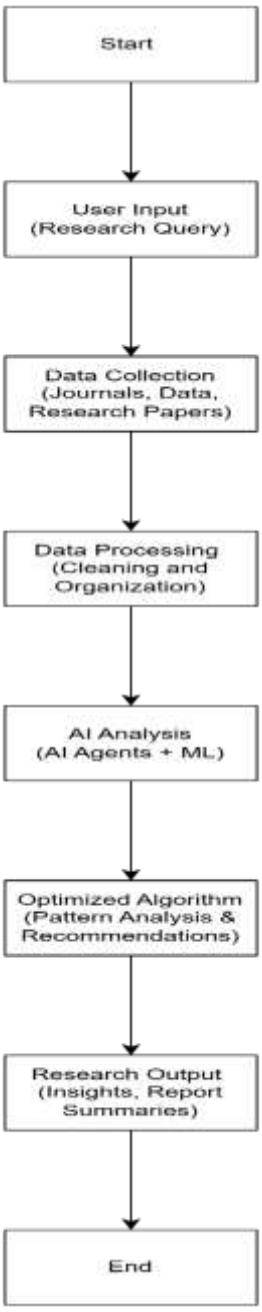

Fig. 2. Basic Working of the AI Integrated Tool with Optimized Algorithm for Research Purpose

III. RESULT AND DISSCUSSION

The proposed AI-powered research assistant system demonstrates strong potential in improving the efficiency of academic research and knowledge discovery. The system integrates information retrieval, recommendation techniques, and generative AI to assist users in collecting, analysing, and summarizing research data. The implementation of intelligent modules enables faster retrieval of relevant research papers and automated generation of summaries and insights. Previous studies indicate that AI-based systems improve research productivity and accessibility by enabling efficient processing of large datasets and providing structured outputs [3], [17].

Furthermore, the use of semantic retrieval and data-driven techniques enhances the accuracy of identifying relevant information from large research collections. Machine learning and NLP-based approaches contribute to improved knowledge extraction and decision support in research environments [18], [13]. The integration of generative AI also supports automated tasks such as literature summarization, insight generation, and recommendation of related research works, which are essential for modern research workflows [9], [10].

Overall, the results indicate that the proposed system reduces the time required for literature review and improves access to relevant academic content. The findings highlight that AI-integrated research platforms can significantly support students, researchers, and academic institutions by enhancing research efficiency, accessibility, and data-driven decision-making.

A. Individual Perspective

The individual perspective focuses on the factors that influence the adoption and usage of AI-powered research assistant systems by students, researchers, and educators. Perceived ease of use and perceived usefulness are key factors that determine whether users find the system effective for improving research productivity and knowledge discovery [1], [3], [17]. Additionally, performance expectancy and reliability influence users' confidence in the system's ability to provide accurate recommendations and meaningful insights [18], [13].

Trust, security, and perceived risk also play an important role in the acceptance of AI-based research tools, as users must be assured that their data is protected and the system produces reliable results [10], [14]. Furthermore, features such as personalization, collaboration, and quality of service enhance the overall user experience by enabling better interaction and more relevant research recommendations [16], [21].

These factors collectively contribute to the successful adoption of AI-driven research systems. AI-powered tools provide advantages such as automated analysis, intelligent recommendations, and faster access to academic information, making them more efficient compared to traditional research methods [3], [17].

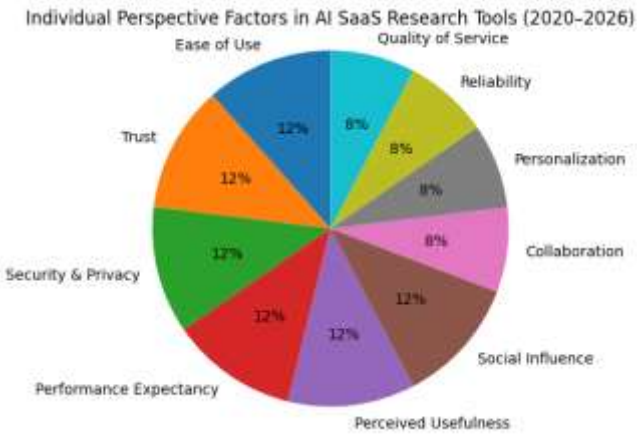

Fig. 3. Individual Perspective Factors Influencing AI-Based Research Systems

TABLE II. INDIVIDUAL PERSPECTIVE FACTOR

Number	Factors	References
1	Perceived Ease of Use	[1], [2], [6]
2	Social Influence	[7], [17]
3	Perceived Usefulness	[2], [3], [15]
4	Security and Privacy	[10], [14]
5	Effort Expectancy	[1], [6]
6	Trust	[3], [10]
7	Performance Expectancy	[18], [13]
8	Attitude Toward AI Technology	[3], [6]
9	Perceived Benefit	[2], [15]
10	Collaboration	[2], [9]
11	Perceived Risk	[10], [14]
12	Perceived Cost	[6], [17]
13	Task Technology Fit	[13], [18]
14	Self-Efficacy	[6], [23]

15	Mobility / Accessibility	[24], [14]
16	Personalization	[7], [8]
17	Perceived Convenience	[2], [24]
18	Strategic Value	[3], [11]
19	Marketing Effort	[14], [21]
20	Reliability	[13], [21]
21	Responsiveness	[8], [16]
22	Facilitating Condition	[17], [6]
23	Quality of Service	[10], [11]
24	Compatibility	[13], [18]
25	Relative Advantage	[1], [7]

B. Organizational Perspective

The organizational perspective focuses on how academic institutions and research organizations adopt AI-powered research assistant systems to improve research efficiency and knowledge management. Factors such as strategic value, collaboration, reliability, and quality of information play an important role in determining the effectiveness of such systems. Organizations increasingly utilize AI-based tools to manage large volumes of research data, support literature analysis, and enhance decision-making processes. Prior studies indicate that AI-driven systems can improve research workflows, enable better data processing, and support collaborative academic environments [2], [17].

In addition, organizational readiness plays a key role in successful implementation. Factors such as facilitating conditions, system compatibility, and trust in AI technologies influence the adoption of research platforms. Institutions must ensure that the system integrates well with existing infrastructure and provides reliable and secure access to academic data [10], [14]. Research also highlights that effective use of AI technologies can improve knowledge discovery, streamline research activities, and enhance the overall quality of academic outputs [3], [11].

By adopting AI-powered research assistant systems, organizations can improve research productivity, enable faster access to relevant information, and support data-driven decision-making in academic and scientific environments.

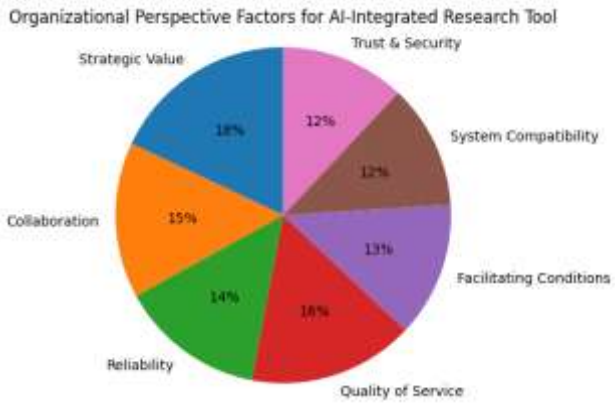

Fig. 4. Organizational Perspective

1) Technological Factors

Technological factors refer to the characteristics of the system that influence the adoption and effectiveness of AI-powered research assistant tools. Key aspects include system compatibility, reliability, security, and quality of information processing. Modern research systems leverage advancements in information retrieval, natural language processing, and embedding techniques to improve the relevance and accuracy of retrieved content [21], [13], [16].

The integration of semantic search and transformer-based models enables better understanding of user queries and research documents, resulting in more precise retrieval and analysis [13], [19]. Additionally, Retrieval-Augmented Generation (RAG) enhances system performance by combining document retrieval with generative models, allowing accurate summarization and insight generation [5], [10], [11].

These technological capabilities enable efficient processing of large-scale academic data and support automated research assistance tasks such as literature analysis, recommendation, and knowledge extraction. As a result, advanced AI-based architectures contribute to the development of scalable and intelligent research systems for academic and scientific applications.

2) Organizational Factors

Organizational factors focus on the internal capabilities and strategies that influence the adoption of AI-powered research systems. Key elements include strategic value, collaboration, facilitating conditions, and organizational readiness. Institutions adopt AI-based tools to improve research productivity, streamline workflows, and enhance collaboration among researchers and academic departments [2], [17].

AI-driven systems enable organizations to efficiently manage large volumes of research data, automate repetitive tasks, and support decision-making processes. Successful implementation depends on the availability of proper infrastructure, skilled personnel, and institutional support for digital transformation. Studies indicate that organizations that actively adopt AI technologies are better positioned to improve research outcomes and innovation [3], [11].

Additionally, trust in AI systems and compatibility with existing infrastructure play a crucial role in ensuring effective adoption. Organizations must ensure that research platforms are reliable, secure, and aligned with their academic and technological requirements [10], [14].

3) Environmental Factors

Environmental factors refer to external influences that impact the adoption and development of AI-powered research systems. These include technological advancements, research trends, collaboration opportunities, and the availability of digital resources. The rapid progress in AI, NLP, and generative models has encouraged institutions to adopt intelligent research tools to remain competitive in the evolving academic landscape [5], [11].

Collaboration among universities, research communities, and technology providers further supports the development of advanced research platforms. Shared knowledge, open-access resources, and interdisciplinary research contribute to innovation in AI-based systems. Additionally, the availability of large-scale datasets and digital libraries enables more effective implementation of intelligent retrieval and recommendation systems [24], [14].

These environmental factors create a supportive ecosystem for the adoption of AI-driven research tools, enabling improved knowledge discovery, faster research processes, and enhanced academic productivity across institutions [17], [3].

IV. CONCLUSION

The proposed AI-powered research assistant system presents an effective approach for improving the efficiency and quality of academic research. By integrating information retrieval, recommendation techniques, and generative AI, the system enables users to efficiently access, analyze, and utilize large volumes of scholarly information. These capabilities support automated literature analysis, knowledge extraction, and research assistance, thereby reducing manual effort and improving decision-making processes.

Advancements in Natural Language Processing and retrieval techniques have significantly enhanced the ability of systems to understand user queries and provide relevant academic content [13], [21]. In addition, the incorporation of recommendation approaches based on user behavior improves the relevance and personalization of research results [1], [7]. The use of Retrieval-Augmented Generation further enhances system performance by combining retrieval with generative models, enabling accurate summarization and context-aware responses [5], [10], [11].

Furthermore, AI-powered research systems contribute to improved research workflows by enabling efficient data processing, automated insights, and better knowledge discovery. Studies indicate that AI-based tools enhance research productivity and accessibility, allowing researchers to focus more on analysis and innovation [3], [17]. The integration of intelligent technologies also supports collaboration and effective utilization of academic resources in research environments.

In conclusion, the proposed system demonstrates the potential of combining retrieval, recommendation, and generative AI into a unified platform for academic research. Such systems provide a scalable and efficient solution for modern research challenges, enabling students, researchers, and institutions to enhance productivity, improve knowledge discovery, and contribute to the advancement of AI-driven scientific research.

REFERENCES

[1] Xu, Zhelin, Shuhei Yamamoto, and Hideo Joho. "Research paper recommender system by considering users' information seeking behaviors." 2025 International Joint Conference on Neural Networks (IJCNN). IEEE, 2025.

[2] Chaudhari, Devchand, et al. "ARA: Artificial Intelligence-Powered Research Assistant." 2024.

[3]Bui-Thanh, Tan. "The AI Research Assistant: Promise, Peril, and a Proof of Concept." arXiv preprint arXiv:2602.22842 (2026).

[4] Rahman, Shadikur, et al. "Automated Research Article Classification and Recommendation Using NLP and ML." arXiv preprint arXiv:2510.05495 (2025).

[5] Brown, Andrew, Muhammad Roman, and Barry Devereux. "A systematic literature review of retrieval-augmented generation: Techniques, metrics, and challenges." arXiv preprint arXiv:2508.06401 (2025).
[6]Poikonen, Topias. Examining the applicability and attitudes towards artificial intelligence as a research assistant. MS thesis. T. Poikonen, 2025.

[7] Xia, Ziyuan, et al. "Contemporary recommendation systems on big data and their applications: A survey." IEEE access 12 (2024): 196914-196928.

[8]Chen, Jiawei, et al. "Benchmarking large language models in retrieval-augmented generation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 16. 2024.

[9] Schopf, Tim, and Florian Matthes. "NLP-KG: A system for exploratory search of scientific literature in natural language processing." Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations). 2024.

[10] Yu, Hao, et al. "Evaluation of retrieval-augmented generation: A survey." CCF Conference on Big Data. Singapore: Springer Nature Singapore, 2024.

[11] Wu, Shangyu, et al. "Retrieval-augmented generation for natural language processing: A survey." arXiv preprint arXiv:2407.13193 (2024).

[12] Kotstein, Sebastian, and Christian Decker. "RESTBERTa: a Transformer-based question answering approach for semantic search in Web API documentation." Cluster Computing 27.4 (2024): 4035-4061.

[13] Wang, Jiajia, et al. "Utilizing bert for information retrieval: Survey, applications, resources, and challenges." ACM Computing Surveys 56.7 (2024): 1-33.

[14] Hambarde, Kailash A., and Hugo Proenca. "Information retrieval: recent advances and beyond." IEEE Access 11 (2023): 76581-76604.

[15] Lála, Jakub, et al. "Paperqa: Retrieval-augmented generative agent for scientific research." arXiv preprint arXiv:2312.07559 (2023).

[16] Muennighoff, Niklas. "Sgpt: Gpt sentence embeddings for semantic search." arXiv preprint arXiv:2202.08904 (2022).

[17] Urdaneta-Ponte, María Cora, Amaia Mendez-Zorrilla, and Ibon Oleagordia-Ruiz. "Recommendation systems for education: Systematic review." Electronics 10.14 (2021): 1611.

[18] Zhou, Ming, et al. "Progress in neural NLP: modeling, learning, and reasoning." Engineering 6.3 (2020): 275-290.

[19] Zhang, Zhuosheng, et al. "Semantics-aware BERT for language understanding." Proceedings of the AAAI

conference on artificial intelligence. Vol. 34. No. 05. 2020.

[20]Patel, Manish. "TinySearch--semantics based search engine using bert embeddings." arXiv preprint arXiv:1908.02451 (2019).

[21] Ponte, Jay M., and W. Bruce Croft. "A language modeling approach to information retrieval." ACM SIGIR Forum. Vol. 51. No. 2. New York, NY, USA: ACM, 2017.

[22]Cambria, Erik, and Bebo White. "Jumping NLP curves: A review of natural language processing research." IEEE Computational intelligence magazine 9.2 (2014): 48-57.

[23]Alhawiti, Khaled M. "Natural language processing and its use in education." International Journal of Advanced Computer Science and Applications 5.12 (2014): 72-76.

[24]Kobayashi, Mei, and Koichi Takeda. "Information retrieval on the web." ACM computing surveys (CSUR) 32.2 (2000): 144-173.

[25] Song, Fei, and W. Bruce Croft. "A general language model for information retrieval." Proceedings of the eighth international conference on Information and knowledge management. 1999.

[26]Saracevic, Tefko. "Evaluation of evaluation in information retrieval." Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval. 1995.