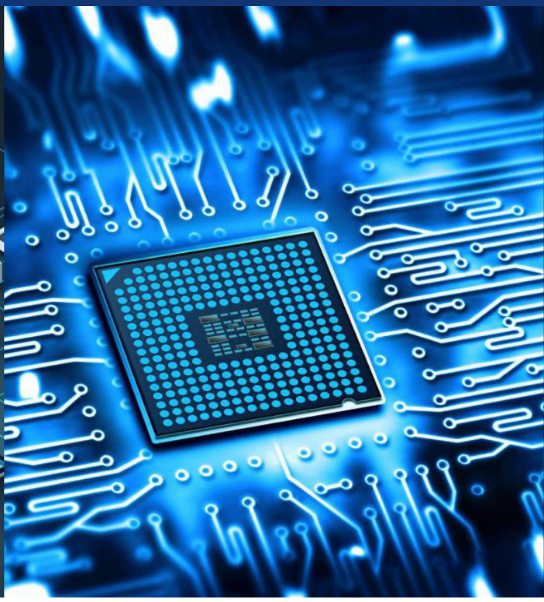
IJIRCCE

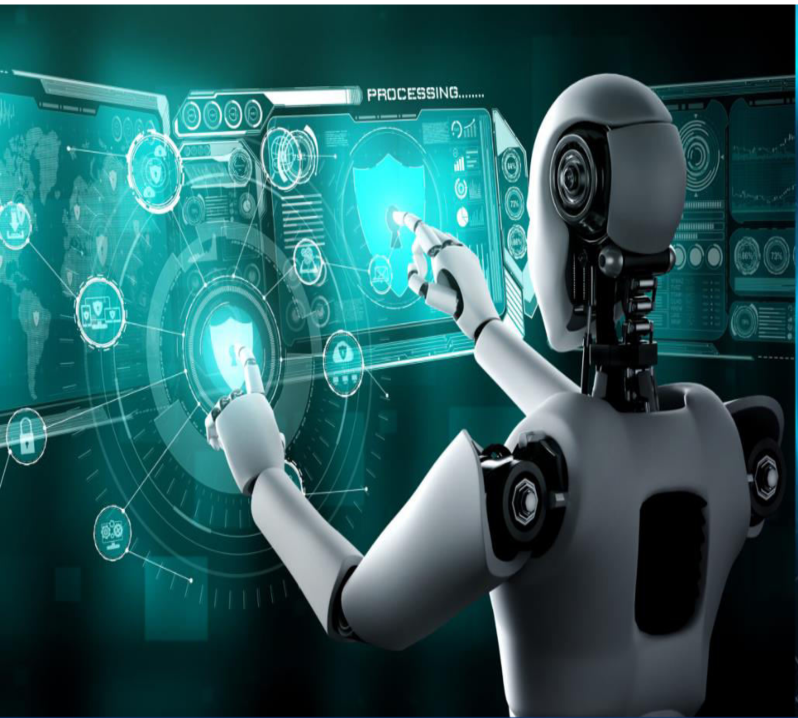
ISSN(O): 2320-9801

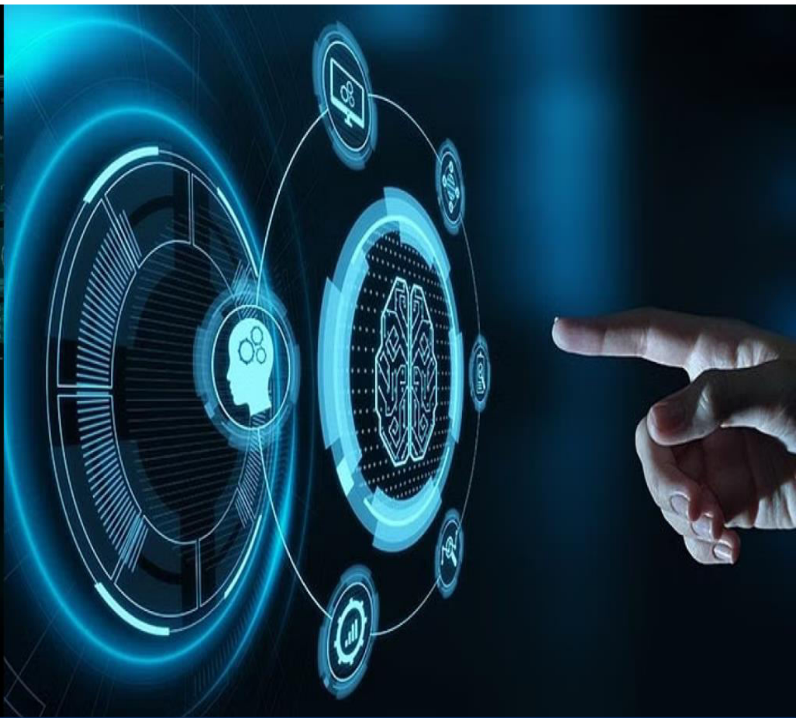
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 7, July 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A Retrieval-Augmented Generation (RAG) Framework for Clinical Decision Support, Drug Interaction Detection & Prescription Safety

Bhumi Kishor Makde[1], Ketki Narendra Thakre[2], Abhyudaya Aware[3], Md. Shahanwaz[4], Dr. Poorva Sabnis[5]

Department of Computer Engineering, St. Vincent Pallotti College of Engineering and Technology, Nagpur, India

Assistant Professor, Department of Computer Engineering, St. Vincent Pallotti College of Engineering and Technology, Nagpur, India

ABSTRACT: This project proposes a Retrieval-Augmented Generation (RAG) based healthcare assistant that intelligently combines trusted medical knowledge sources with advanced language models to provide accurate, reliable, and context-aware healthcare responses [11][13][16]. Traditional large language models (LLMs) often generate hallucinated or outdated medical information due to lack of real-time domain-specific grounding, which can be risky in healthcare applications [4][5]. To address this limitation, the proposed system introduces a retrieval mechanism that searches verified healthcare sources such as PubMed, DrugBank, clinical guidelines, and medical databases before generating responses [10][11]. User queries are processed using Natural Language Processing techniques such as intent detection, entity recognition, and semantic embedding search to retrieve the most relevant medical information [1][4][20]. The retrieved knowledge is then combined with a Large Language Model to generate structured, evidence-based, and explainable responses [11][12]. A validation layer is incorporated to verify drug interactions, dosage risks, outdated information, and safety concerns, thereby reducing hallucinations and improving trustworthiness [14][15]. The system also supports multilingual interaction in languages such as English, Hindi, Marathi, and Gujarati to improve accessibility [2] [8]. The system is implemented using Python, FastAPI/Flask, LangChain, vector databases such as FAISS or Pinecone, and backend databases such as MongoDB or PostgreSQL. It will be evaluated using healthcare question-answering benchmarks and retrieval performance metrics, demonstrating improved accuracy, reliability, and usability compared to traditional healthcare chatbot systems [6][11][17].

KEYWORDS: Retrieval-Augmented Generation (RAG), Healthcare Chatbot, Large Language Models (LLMs), Medical Knowledge Retrieval, Semantic Search, LangChain, FAISS, Evidence-Based Healthcare.

I. INTRODUCTION

The healthcare sector has witnessed remarkable technological advancements over the past decade, with Artificial Intelligence (AI) playing a significant role in improving diagnosis, clinical decision-making, hospital management, and patient care [3], [17]. Modern healthcare systems generate massive volumes of data, including Electronic Health Records (EHRs), laboratory reports, prescriptions, clinical guidelines, and medical research publications [3]. However, retrieving relevant and trustworthy medical information during time-sensitive situations remains a major challenge for healthcare professionals.

Recent developments in Large Language Models (LLMs) have demonstrated exceptional capabilities in natural language understanding, question answering, and conversational interaction [4][5]. Although these models can interpret complex medical queries and generate human-like responses, standalone LLMs often produce hallucinated or outdated information, which may lead to incorrect medical recommendations in critical healthcare environments [6][11]. Retrieval-Augmented Generation (RAG) has emerged as an effective approach to overcome these limitations by combining external knowledge retrieval with language generation [11][13][16]. Instead of relying solely on pre-trained knowledge, RAG retrieves relevant information from trusted medical sources before generating responses. This approach improves factual accuracy, reliability, and explainability while significantly reducing hallucinations [11][13]. Medication safety is another critical concern in healthcare. Prescription errors, drug-drug interactions, dosage conflicts, duplicate therapies, and adverse drug reactions remain major causes of preventable patient harm [10][14][15].

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Furthermore, patients frequently experience difficulty understanding medication instructions and post-discharge care, particularly in multilingual countries such as India, where language barriers negatively affect treatment adherence. Multilingual language models can help bridge this communication gap by providing healthcare information in regional languages [2][7][8][18]. To address these challenges, this research proposes a Retrieval-Augmented Generation (RAG)-based multilingual healthcare assistant that integrates trusted medical knowledge sources with advanced language models to provide accurate, context-aware, and evidence-based healthcare assistance. The proposed system employs Natural Language Processing (NLP) techniques, including biomedical language understanding, semantic embedding, and intelligent information retrieval, to process user queries efficiently [1][4][5][20].

The platform is designed to support both healthcare professionals and patients. It assists clinicians by providing evidence-based clinical decision support and automated prescription safety analysis while enabling patients to access multilingual chatbot services, medication guidance, and post-discharge assistance [11][12][13][15]. Secure communication and privacy-preserving mechanisms are incorporated to ensure safe handling of sensitive healthcare data [19]. Overall, the proposed RAG-based healthcare assistant aims to enhance the accuracy, reliability, accessibility, and safety of AI-assisted healthcare by integrating trusted medical knowledge with advanced language models. The system seeks to reduce misinformation, improve clinical decision-making, strengthen prescription safety, and provide an intelligent multilingual platform for healthcare support [11][12][13][15].

II. DATASETS AND DATA FOUNDATION

The proposed RAG-based healthcare assistant utilizes trusted medical datasets and knowledge sources to generate accurate and evidence-based responses. The primary dataset is MedQuAD, which provides medical question-answer pairs for healthcare applications [6]. Additional medical knowledge is obtained from PubMed, MedlinePlus, WHO guidelines, DrugBank, SIDER, and the openFDA API to support clinical decision-making, drug interaction detection, and medication safety [10][11].

The collected data undergoes preprocessing, including text cleaning, tokenization, document chunking, and medical terminology normalization [1][5]. Biomedical text is converted into vector embeddings using BERT, BioBERT, and Sentence-BERT, which are stored in FAISS or Pinecone for efficient semantic retrieval [4][5][20]. Multilingual support is provided through MuRIL and mT5, enabling healthcare assistance in English, Hindi, Marathi, and other regional languages [2][8].

III. RESEARCH GAPS IN RAG-BASED HEALTHCARE ASSISTANT SYSTEMS

Recent advancements in Artificial Intelligence (AI), Large Language Models (LLMs), and Retrieval-Augmented Generation (RAG) have significantly improved healthcare applications such as medical question answering, clinical decision support, and healthcare chatbots. However, several research gaps still limit the effectiveness and reliability of these systems [11][12][13].

One of the primary limitations of standalone LLMs is their dependence on pre-trained knowledge, which can result in hallucinated, outdated, or inaccurate medical information. In healthcare, such errors may lead to incorrect diagnoses, medication recommendations, or treatment advice, making reliability a critical concern [11][12][16]. Although RAG improves factual accuracy by retrieving information from trusted medical sources, many existing systems focus mainly on medical question answering and lack integrated features such as prescription safety validation, drug-drug interaction detection, dosage verification, and personalized clinical recommendations [13][14][15]. Additionally, most healthcare AI systems provide limited multilingual support, reducing accessibility for patients who communicate in regional languages [2][8].

Another significant challenge is the lack of patient-specific context and continuous post-discharge support. Existing solutions often generate generalized responses without considering individual medical history, ongoing medications, or follow-up care requirements [11][13]. Furthermore, many proposed systems have been evaluated only on benchmark datasets, with limited real-world clinical validation and deployment [3][19].

The proposed work addresses these research gaps by developing a unified RAG-based multilingual healthcare assistant that combines trusted medical knowledge retrieval, clinical decision support, prescription safety analysis, drug

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

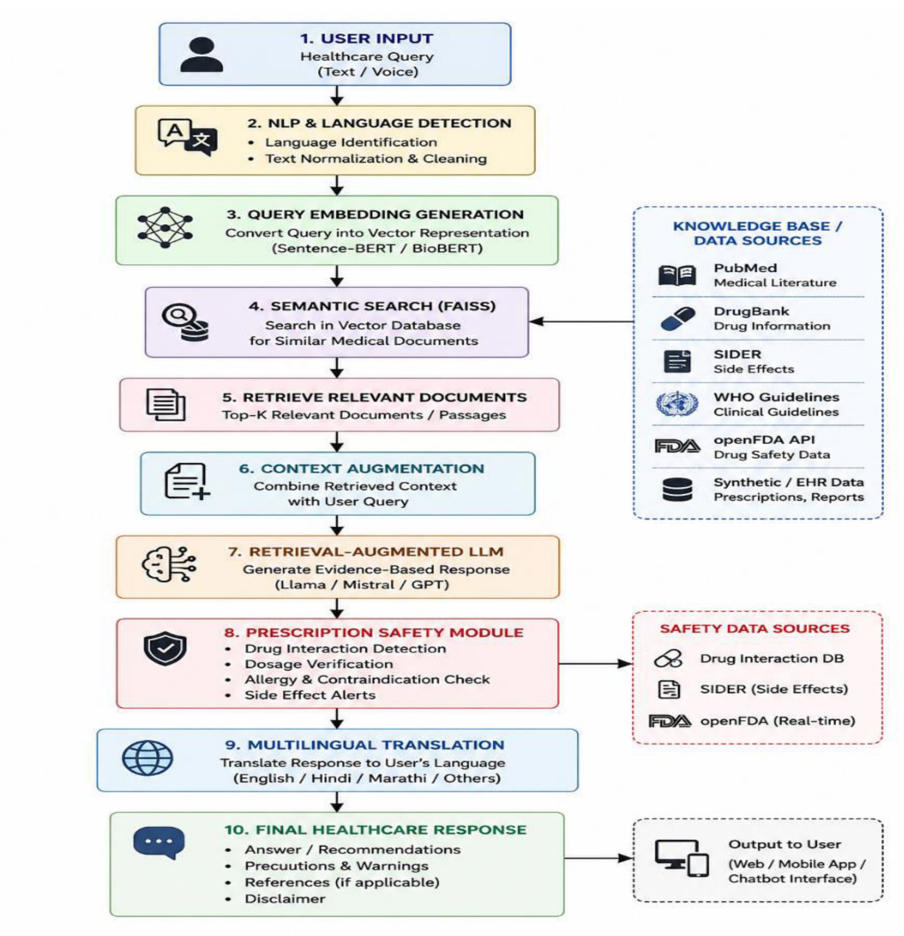
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

interaction detection, multilingual communication, and post-discharge patient assistance. This integrated approach aims to provide accurate, evidence-based, and context-aware healthcare support while improving accessibility, reliability, and patient safety [11][13][16].

IV. REVIEW OF DEEP LEARNING METHODOLOGIES

Retrieval-Augmented Generation (RAG) has emerged as an effective approach for improving the reliability of Large Language Models (LLMs) in healthcare applications. Unlike conventional LLMs that rely only on pre-trained knowledge, RAG first retrieves relevant information from trusted external sources and then uses this information to generate evidence-based responses, thereby reducing hallucinations and improving factual accuracy [11][13][16]. The RAG framework consists of three major components: document retrieval, context augmentation, and response generation. User queries are converted into vector embeddings using models such as Sentence-BERT, BERT, or BioBERT, and semantically similar documents are retrieved from vector databases like FAISS or Pinecone [4][5][20]. The retrieved medical context is then provided to a Large Language Model, enabling it to generate accurate and context-aware responses.

Recent studies have demonstrated the effectiveness of RAG in clinical decision support, prescription safety, and medical question answering. Advanced frameworks such as PACE-RAG further enhance response quality by incorporating patient specific context and evidence-based reasoning [11][12][13]. Additionally, multilingual models such as MuRIL and mT5 improve accessibility by supporting healthcare communication in multiple languages [2][8]. Overall, RAG provides a scalable and trustworthy methodology for healthcare AI by combining semantic retrieval, biomedical language models, and validated medical knowledge, making it well suited for clinical decision support, drug interaction detection, and multilingual patient assistance [11][13][16].



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

TABLE I. SUMMARY OF RELEVANT LITERATURE (ordered by year, most recent first)

Year	Author(s)	Contribution	Gap / Limitation
2026	Chaeyoung Huh et al.	Proposed PACE-RAG, integrating patient context, similar case retrieval, and prescribing pattern analysis for personalized drug recommendation.	High dependency on historical patient data; risk of bias; computational complexity of multistage pipeline.
2026	Huh et al. (Framework Enhancement)	Introduced policy-driven prescription refinement and explainable clinical summaries for better interpretability.	Limited real-world clinical deployment; explainability still dependent on retrieved data quality.
2026	Fatima Saidu & Julie Wall	Proposed KG-enhanced LLM for clinical decision support; improved diagnostic reasoning, confidence, and reduced diagnosis time in rare diseases	Focuses more on performance & clinician interaction, but lacks large-scale real-world deployment and evaluation across diverse diseases
2025	Bin Yang et al.	Proposed DGAT (Dual Graph Attention Network) for predicting drug-drug interactions in Traditional Chinese Medicine using deep learning and graph structures.	Focuses only on molecular interaction; lacks patient-specific context; not integrated into clinical decision systems.
2024	Daniel Shu Wei Ting et al.	Developed a RAG-based LLM Clinical Decision Support System (CDSS) for medication safety. Showed improved accuracy, recall, and F1-score, especially in co-pilot mode with pharmacists.	Lower precision; limited real-world validation; dependency on institutional datasets reduces generalizability.
2023	Aftab Alam et al.	Reviewed AI/ML approaches for DDI prediction; highlighted improvements in accuracy, real-time detection, and reduced alert fatigue	Mostly review-based, lacks implementation framework and experimental validation
2021	Linting Xue, Noah Constant, Adam Roberts, et al.	Developed mT5, a multilingual transformer trained on 101 languages for text-to-text NLP tasks	Language data imbalance, accidental translation in zero shot settings, high training cost.
2021	Z. Borsos, M. Mutný, M. Tagliasacchi, and A. Krause	Proposed a generic coreset framework via bilevel optimization for efficient data summarization across models.	Computational complexity, scalability issues, limited to differentiable models.
2021	Simran Khanuja, Diksha Bansal, Sarvesh Mehtani, et al.	Proposed MuRIL model for Indian languages; uses translated & transliterated data; handles code-mixed text; better than mBERT	Supports limited languages (17); depends on parallel data; not tested on global low-resource languages
2021	Yu Gu, Robert Tinn, Hao Cheng, et al.	Developed PubMedBERT using domain-specific pretraining from scratch; introduced BLURB benchmark; achieved state of-the-art results in biomedical NLP	Requires large domain-specific data; limited to biomedical field; lacks cross domain adaptability
2020	Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Edouard Grave, et al.	Introduced **XLM-R (Cross-lingual Language Model)** for unsupervised multilingual representation learning across 100+ languages, significantly improving cross-lingual transfer learning and multilingual NLP performance.	Requires extensive computational resources and massive multilingual corpora; less effective for highly specialized domains such as biomedical text without domain-specific fine-tuning.

2020	Jingwen Xu, Benjamin S. Glicksberg, Chao Su, Peter Walker, Jiang Bian, and Fei Wang	Reviewed **Federated Learning** techniques for healthcare, enabling collaborative model training across hospitals while preserving patient privacy by keeping data decentralized. Highlighted applications in medical imaging, EHR analysis, and clinical prediction.	Faces challenges related to communication overhead, heterogeneous data distributions, security vulnerabilities, and implementation complexity across healthcare organizations.
2020	Di Jin, Eileen Pan, Nassim Oufattole, Wei-Hung Weng, Hanyi Fang, Peter Szolovits	Introduced MEDQA, a large-scale multilingual medical QA dataset requiring deep reasoning and domain knowledge.	Low performance of current QA systems, weak multi-hop reasoning, difficulty in handling complex clinical decision making.
2019	Jacob Devlin, MingWei Chang, Kenton Lee, Kristina Toutanova	Introduced BERT; proposed MLM & NSP techniques; foundation for modern NLP models.	General-domain only; needs fine-tuning; weak in domain specific tasks
2019	Nils Reimers and Iryna Gurevych	Developed **Sentence-BERT (SBERT)**, a Siamese BERT architecture that generates semantically meaningful sentence embeddings for efficient semantic similarity, clustering, and information retrieval. It significantly improves retrieval speed and accuracy, making it well-suited for semantic search and Retrieval-Augmented Generation (RAG) systems.	Requires domain-specific fine-tuning for optimal performance; inherits BERT's computational complexity; embedding quality depends on the training data and may be less effective for highly specialized biomedical tasks without additional adaptation.
2019	Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, et al.	Adapted BERT for biomedical domain; improved NER & QA tasks; early domain-specific NLP model	Uses continued training not from scratch; limited vocabulary adaptation; later outperformed
2018	Alvin Rajkomar, Eyal Oren, Kai Chen, et al	Uses full raw EHR data; applies deep learning with FHIR format; improves prediction of mortality, readmission, LOS	High computational cost; low interpretability; limited generalization across hospitals
2016	Edward Choi,Mohammad Taha Bahadori, Andy Schuetz, Walter F. Stewart, and Jimeng Sun	Proposed **Doctor AI**, an RNN-based framework for predicting future diagnoses, medications, and clinical events from longitudinal Electronic Health Records (EHRs).	Limited interpretability; depends on large, high-quality EHR datasets; may not generalize well across healthcare institutions.
2016	Yonghui Wu et al.	Proposed GNMT, a neural machine translation system using deep LSTMs, attention, and Word Piece units for improved translation quality.	High computational cost, challenges with rare words, limited robustness across diverse languages
2016	Michael Kuhn, Ivica Letunic, Lars Juhl Jensen, Peer Bork	Developed the SIDER database, a comprehensive resource of drugs, side effects (ADRs), and their frequencies from clinical and post-marketing data	Limited causal relationships, bias in reported ADR data, lacks real-time updates and deeper integration with genomic/clinical systems.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. RELIABILITY AND TRUST IN RAG-BASED HEALTHCARE ASSISTANT SYSTEM

High response accuracy alone is insufficient in healthcare; healthcare professionals and patients must also understand the basis of AI-generated recommendations. Retrieval-Augmented Generation (RAG) improves explainability by retrieving relevant information from trusted medical sources such as PubMed, DrugBank, clinical guidelines, and healthcare databases before generating responses [11][13][16]. By grounding answers in verified medical evidence, the system reduces hallucinations and enables users to trace recommendations back to reliable knowledge sources, thereby increasing confidence in AI-assisted healthcare.

In addition to evidence retrieval, incorporating safety validation mechanisms such as drug-drug interaction checks, dosage verification, and contraindication analysis further enhances the reliability of generated responses [12][14][15]. These validation layers help ensure that recommendations are clinically relevant and aligned with established medical practices. However, providing retrieved evidence alone is not sufficient to establish complete clinical trust. Healthcare professionals require AI systems to present concise explanations, relevant references, and clear reasoning behind recommendations while acknowledging uncertainty when sufficient evidence is unavailable. Therefore, the proposed RAG-based healthcare assistant combines evidence-based retrieval, multilingual communication, and safety validation to deliver transparent, explainable, and trustworthy healthcare assistance. Such an approach supports informed clinical decision-making while ensuring that final medical judgments remain under the supervision of qualified healthcare professionals [11][13][16].

VI. CHALLENGES AND REAL-WORLD DEPLOYMENT CONSIDERATIONS

Although Retrieval-Augmented Generation (RAG)-based healthcare assistants offer significant improvements in accuracy and reliability, several challenges remain for real-world deployment. One major challenge is ensuring the quality and freshness of medical knowledge sources. Since healthcare guidelines, drug information, and treatment protocols are continuously updated, the retrieval database must be regularly maintained to provide accurate and evidence-based recommendations [10][11][13].

Another challenge is integrating the system with existing hospital infrastructure, including Electronic Health Records (EHRs), laboratory systems, and pharmacy databases. Achieving interoperability while maintaining data consistency requires standardized healthcare protocols and secure APIs [3][19]. Protecting sensitive patient information is equally important, requiring robust authentication, encryption, and privacy-preserving mechanisms to ensure compliance with healthcare data security standards.

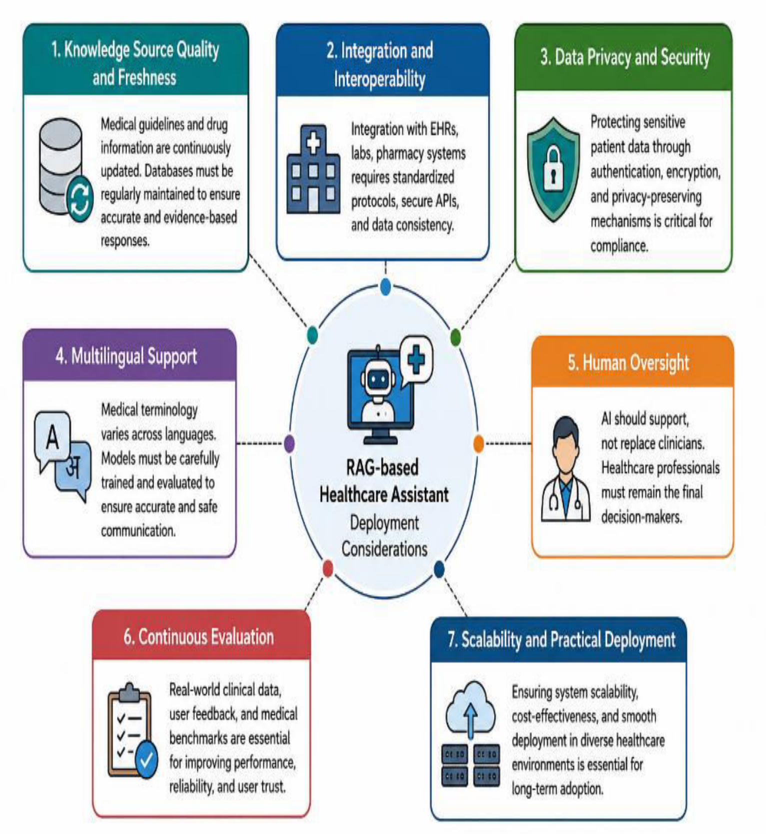
Multilingual support also presents practical challenges. Medical terminology varies across languages, and accurate translation is essential to avoid misunderstandings that could affect patient safety. Therefore, multilingual models must be carefully trained and evaluated for healthcare-specific communication [2][8].

In addition, AI-generated responses should support—not replace—clinical expertise. Healthcare professionals must remain the final decision-makers, particularly in complex or high-risk cases. Continuous evaluation using real-world clinical data, user feedback, and medical benchmarks is necessary to improve system performance, reliability, and user trust [11][12][13]. Addressing these challenges will enable the safe, scalable, and effective deployment of RAG based healthcare assistants in real healthcare environments.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



VII. FUTURE RESEARCH DIRECTIONS

Retrieval-Augmented Generation (RAG) has demonstrated significant potential in improving the accuracy and reliability of AI-assisted healthcare. However, several opportunities remain for enhancing its capabilities and expanding its real world applications. Future research can focus on integrating real-time Electronic Health Records (EHRs), wearable device data, and Internet of Things (IoT)-based health monitoring systems to provide more personalized and context-aware healthcare recommendations [3][17].

Another important direction is the development of advanced multimodal RAG systems that combine medical text, laboratory reports, medical images, and clinical records for comprehensive clinical decision support. Incorporating larger biomedical knowledge graphs and continuously updated medical databases can further improve the factual accuracy and explainability of generated responses [11][13][16].

Future work may also explore stronger multilingual capabilities by supporting additional regional languages and improving cross-lingual retrieval using advanced transformer models such as MuRIL and mT5 [2][8]. Enhancing explainability through evidence citation, confidence estimation, and personalized recommendations based on patient history will further increase clinician and patient trust.

Additionally, future research should investigate privacy-preserving techniques such as federated learning and secure datasharing mechanisms to protect sensitive patient information while enabling collaborative model improvement across healthcare institutions [19]. Large-scale clinical validation, integration with hospital information systems, and compliance with healthcare regulations will be essential for translating RAG-based healthcare assistants into routine clinical practice. These advancements will contribute to the development of intelligent, scalable, secure, and trustworthy AI-driven healthcare systems capable of supporting both healthcare professionals and patients [11][13][16].

VIII. CONCLUSION

Retrieval-Augmented Generation (RAG) has emerged as a promising approach for developing reliable and intelligent healthcare assistance systems by combining the reasoning capabilities of Large Language Models (LLMs) with trusted medical knowledge sources. Unlike conventional language models, RAG retrieves relevant information from verified databases before generating responses, thereby improving factual accuracy, reducing hallucinations, and enhancing clinical reliability [11][13][16].

This paper reviewed the datasets, research gaps, RAG methodology, explainability, deployment challenges, and future research directions for AI-driven healthcare assistants. Existing healthcare AI systems often face limitations such as fragmented medical information, limited multilingual support, prescription safety concerns, and insufficient patient specific recommendations. The proposed RAG-based multilingual healthcare assistant addresses these challenges by integrating semantic retrieval, evidence-based clinical decision support, drug interaction detection, prescription safety validation, multilingual communication, and post-discharge patient assistance into a unified platform [11][12][14][15]. By leveraging trusted medical databases, vector-based retrieval, biomedical language models, and rule-based safety validation, the proposed system aims to provide accurate, context-aware, and explainable healthcare guidance for both healthcare professionals and patients. Furthermore, its scalable architecture and multilingual capabilities make it suitable for deployment in diverse healthcare environments, particularly in multilingual regions.

Overall, RAG-based healthcare assistants have the potential to significantly improve clinical decision support, patient engagement, and healthcare accessibility while ensuring that AI-generated recommendations remain transparent, evidence based, and trustworthy. Continued research, real-world clinical validation, and secure integration with healthcare information systems will further strengthen the adoption of RAG-enabled intelligent healthcare solutions [11][13][16].

REFERENCES

[1] Yu Gu et al., "Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing, " 2021.
[2] Simran Khanuja et al., "MuRIL: Multilingual Representations for Indian Languages, " 2021.
[3] Alvin Rajkomar et al., "Scalable and Accurate Deep Learning with Electronic Health Records, " npj Digital Medicine, 2018.
[4] Jacob Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, " 2019.
[5] Jinhyuk Lee et al., "BioBERT: a pre-trained biomedical language representation model, " 2019.
[6] D. Jin, E. Pan, N. Oufattole, W.-H. Weng, H. Fang, and P. Szolovits, "What Disease Does this Patient Have? A Largescale Open Domain Question Answering Dataset from Medical Exams, " 2020.
[7] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, et al., "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, " 2016.
[8] L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, et al., "mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer, " 2021.
[9] Z. Borsos, M. Mutný, M. Tagliasacchi, and A. Krause, "Data Summarization via Bilevel Optimization, " 2021.
[10] M. Kuhn, I. Letunic, L. J. Jensen, and P. Bork, "The SIDER Database of Drugs and Side Effects, " Nucleic Acids Research, vol. 44, 2016.
[11] F. Saidu and J. Wall, "Retrieval-Augmented Large Language Model for Clinical Decision Support with a Medical Knowledge Graph, " Electronics, vol. 15, 2026.
[12] D. S. W. Ting, J. C. L. Ong, L. Jin, et al., "Development and Testing of a Novel Large Language Model-Based Clinical Decision Support Systems for Medication Safety in 12 Clinical Specialties, " 2024.
[13] C. Huh, H. Hwang, J. H. Shin, et al., "PACE-RAG: Patient-Aware Contextual and Evidence-based Policy RAG for Clinical Drug Recommendation, " arXiv preprint arXiv:2603.17356, 2026.
[14] B. Yang, D. Song, Y. Li, and J. Wang, "Drug-drug interaction prediction of traditional Chinese medicine based on graph attention networks, " Scientific Reports, vol. 15, 2025.
[15] A. Alam, A. Saran, S. B. Sridhar, et al., "Artificial Intelligence in Drug Interaction Management: A Transformative Approach in Clinical Decision Support Systems, " International Journal for Innovative Research in Multidisciplinary Field, vol. 11, no. 4, 2025
[16] Huh, C., et al., Policy-driven Prescription Refinement in RAG-based Clinical Systems, 2026.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[17] E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. Sun, "Doctor AI: Predicting clinical events via recurrent neural networks, " Proceedings of Machine Learning for Healthcare, vol. 56, 2016.
[18] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale, " arXiv preprint arXiv:1911.02116, 2020.
[19] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics, " arXiv preprint arXiv:1911.06270, 2020.
[20] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-networks, " arXiv preprint arXiv:1908.10084, 2019.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING