ISSN: 2473-4901
v11 n6337

Retrieval Augmented Generation (RAG) Based Question and Answer System

Emmanuel Balogun
eb20284@georgiasouthern.edu
School of Computing
Georgia Southern University
Statesboro, GA 30460, USA

Hayden Wimmer
hwimmer@georgiasouthern.edu
School of Computing
Georgia Southern University
Statesboro, GA 30460, USA

Carl M Rebman, Jr.
carlr@sandiego.edu
Knauss School of Business
Department of Supply Chain, Operations, and Information Systems
University of San Diego
San Diego, CA 92110, USA

Abstract

This study focuses on improving LLM contextual understanding through external context sourcing RAG. In this approach, documents containing text, images, and tables, are transformed into high-dimensional vectors via LLM-generated embeddings, then stored in a vector database. When a user submits a query, the LLM retrieves the most relevant documents from this vector store, enabling a deeper understanding of the query and improving response quality. A comparative analysis—supported by a t-test—indicates that this RAG-based model outperforms typical LLM, demonstrating greater efficiency and accuracy. These research studies are combined in this thesis to unlock the hidden potentials of LLMs by providing a comprehensive understanding of improving Artificial Intelligence in contextualization for accurate answer generation and engagement. This solution eradicates LLMs' hallucinations by providing them with the right context and putting machines' reasoning on the same page as humans. The thesis also creates a solution to the long challenges of AI and the battle with the spread.

Keywords: RAG, Retrieval Augmented Generation, LLM, Vector Store, Query, Prompt.

ISSN: 2473-4901
v11 n6337

Retrieval Augmented Generation (RAG) Based Question and Answer System

Emmanuel Balogun, Hayden Wimmer and Carl M. Redman Jr.

1. INTRODUCTION

The integration of artificial intelligence (AI) into healthcare has introduced transformative tools, with AI-powered chatbots emerging as a key technology for patient engagement and information dissemination (Laranjo, 2018). In a field where the precision and reliability of information can have profound consequences, the demand for accurate and contextually aware conversational agents is paramount. However, the efficacy of many current chatbot systems is often undermined by their reliance on generalist language models (Kung et al., 2023). These models, trained on vast but non-specialized datasets, are prone to significant shortcomings in high-stakes environments like healthcare, frequently producing responses plagued by factual inaccuracies, hallucinations, and a general lack of domain-specific knowledge, thereby eroding user trust and limiting their clinical utility (Jiang et al., 2023; Thirunavukarasu et al., 2023).

While various approaches have been developed to ground AI models in specific knowledge domains, a persistent challenge lies in effectively parsing the complex and varied document formats endemic to medical literature, such as PDFs containing a mix of unstructured text and structured tabular data (Lin et al., 2019). Existing systems often struggle to extract and synthesize information from these heterogeneous sources, failing to capture the complete context required for precise query resolution (Wang et al., 2018). This gap highlights the need for a sophisticated data processing pipeline capable of interpreting both textual and tabular content to build a comprehensive and reliable knowledge base for specialized applications.

To address these challenges, this paper introduces a novel Retrieval-Augmented Generation (RAG) framework specifically designed to enhance the accuracy of healthcare chatbots for answering frequently asked questions (FAQs). The model implements a specialized ingestion process that not only extracts textual content from PDF documents but also transforms embedded tabular data into a markdown text format, a method shown to improve multimodal data handling (Lewis et al., 2020). By converting both content types into a unified embedding representation, the system creates a robust, domain-specific knowledge base, building on prior work in the joint understanding of textual and tabular data (Yin et al., 2020).

This research argues that by systematically processing the entirety of specialized documents, RAG-based approach significantly mitigates the risks of misinformation and hallucination, leading to a more reliable and effective AI-powered conversational tool for the healthcare sector (Shuster et al., 2021).

2. LITERATURE REVIEW

The fast rising and adoption of Large Language Models (LLMs) has created new opportunities for patient engagement, with AI-powered chatbots being explored for tasks ranging from answering patient queries to medical education (Laranjo et al., 2018). However, the high-stakes nature of healthcare demands a high level of precision and reliability that general-purpose models often fail to provide. Studies evaluating models like ChatGPT on medical licensing exams have shown promising results, but also highlight significant limitations (Brucks & von Bayern, 2020). These models, trained on broad internet corpora, are likely to generating factually incorrect information, a phenomenon often termed "hallucination" [3] (Thirunavukarasu et al., 2023). In a clinical context, such inaccuracies can erode user trust and pose significant risks, underscoring the critical need for systems grounded in domain-specific, verified knowledge.

RAG has emerged as a leading architectural model to address the limitations of standalone LLMs. First proposed by (Lewis et al., 2020), the RAG framework enhances generative models by enabling them to retrieve relevant information from an external knowledge base before generating a response. This approach directly grounds the model's output in factual data, which has been shown to significantly reduce the incidence of hallucination in conversational AI (Shuster et al., 2021). In the medical field, RAG offers a pathway to create more reliable AI systems by connecting them to curated sources of medical literature and data. Recent work by (Yang et al., 2024) demonstrates the application of RAG in healthcare to mitigate biases and

ISSN: 2473-4901
v11 n6337

overcome the static knowledge limitations of generative models, thereby improving the reliability of AI-driven responses.

While RAG provides a robust framework, its effectiveness is fundamentally dependent on the quality of its knowledge base. In healthcare, a significant portion of specialized knowledge is encapsulated in PDF documents, which often feature complex layouts containing a mixture of unstructured text, tables, and figures. Traditional methods for clinical information extraction have long grappled with the complexities of parsing these varied formats (Lin et al., 2019; Wang et al., 2018). This challenge persists in modern RAG systems. (Lin, 2024) highlights that conventional PDF parsing methods often fail to accurately recognize document structures, leading to flawed information retrieval. To address this, the introduction of deep learning-based parsers like ChatDOC has been shown to significantly improve the chunking and extraction of content from complex layouts, including tables and multi-column text, thereby enhancing the accuracy of downstream RAG tasks.

Beyond simple text extraction, a key challenge is creating a unified understanding of both the textual and tabular data within medical documents. Foundational work (Yin et al., 2020) with TaBERT demonstrated the feasibility of pretraining models to jointly comprehend information from natural language text and structured tables. This is crucial for healthcare applications where clinical data, lab results, and treatment guidelines are often presented in tabular form.

Building on this, recent advancements in RAG have focused on optimizing the retrieval process itself. For instance, the Blended RAG framework proposed by (Sawarkar et al., 2024) combines keyword-based, dense vector, and semantic search to form hybrid queries, improving retrieval relevance. Similarly, (Ghali et al., 2024; Koo et al., 2024) introduced a query optimization method that refines user queries to enhance retrieval accuracy, particularly in scientific contexts.

These advanced retrieval strategies, when combined with sophisticated parsing techniques, provide a foundation for building a comprehensive and reliable knowledge base from the heterogeneous data found in medical literature. This review reveals a clear need for an integrated RAG framework that specifically addresses the end-to-end challenge of ingesting complex PDF documents, transforming both textual and tabular data into a unified representation, and leveraging this for accurate, domain-specific answer generation in healthcare.

1. METHODOLOGY

The methodology described herein outlines the development of a RAG based question-answering (QA) system. The system's primary goal is to utilize document retrieval to enhance the relevance and accuracy of LLM outputs when answering questions based on knowledge retrieved from PDF documents. While we use PDF documents as case study, modern RAG can accept a myriad of document types. This section delves into all RAG process stages, from data preprocessing, embedding, and indexing strategies, retrieval techniques, and the final question-answering mechanism, thoroughly exploring each component of the system.

We utilized a series of highly optimized Python libraries like **Unstructured**, **LangChain**, **PDFplumber**, **Natural Language Processing** (NLP), **Numpy**, and **pandas** to handle the various data types and formats present in PDF documents. By partnering with Institute for Health Logistics & Analytics (IHLA), case study documents which contains the One health joint plan of action (2022–2026) and The US investigating the origins of covid–19 hearing. These documents were parsed using UnstructuredPDFLoader, a component of langchain library built to process unstructured text data from diverse document formats, making it well-suited for extracting both textual and tabular data from complex PDF documents. Additionally, the PDFplumber was employed to extract tabular data from document. This choice was guided by the PDFplumber's ability to accurately detect table boundaries and parse them into structured data formats. Using multiple parsing libraries was critical to ensure data integrity and maximize the quality of the extracted content.

Document Preprocessing

Document processing involved reading the PDF input document using the UnstructuredPDFLoader. The text contained within the document was extracted and prepared for subsequent operations. This was done to ensure that the model received a clean and consistent input across all document sections.

One of the significant challenges when handling PDFs is the extraction of tabular data. Tables, which often contain structured numerical or categorical information, cannot be processed in their native format by LLMs. To address this, PDFplumber was used to extract tables, which

ISSN: 2473-4901
v11 n6337

were then converted into a Markdown format. Markdown allows for the structured representation of tabular data using simple text symbols, which the downstream LLM can easily understand during the question-answering phase.

Figures 1 and 2 as shown below illustrates how tabular data is converted into markdown format to help LLMs understand the concept of tables.

Column 1	Column 2	Column 3
Row 1, Col 1	Row 1, Col 2	Row 1, Col 3
Row 2, Col 1	Row 2, Col 2	Row 2, Col 3
Row 3, Col 1	Row 3, Col 2	Row 3, Col 3

Figure 1: Initial Table

```
| Column 1 | Column 2 | Column 3 |
|----------|----------|----------|
| Row 1, Col 1 | Row 1, Col 2 | Row 1, Col 3 |
| Row 2, Col 1 | Row 2, Col 2 | Row 2, Col 3 |
| Row 3, Col 1 | Row 3, Col 2 | Row 3, Col 3 |
```

Figure 2: Markdown text format for LLM

This process of converting tables into Markdown offers two key advantages:

- **Text-based Representation:** The LLM can now process the tables as text, enabling it to understand tabular relationships linearly, which is vital for generating answers that consider tabular data.
- **Seamless Integration:** Markdown tables integrate smoothly with other text, ensuring that the context of both text and tabular data is preserved during the chunking and embedding stages.

Once the text and tables were extracted and cleaned, the next step involved dividing the document into smaller chunks for embedding. The document was split into text blocks of approximately 4,000 characters each using the RecursiveCharacterTextSplitter, a LangChain Python library tool. This recursive text splitter was selected due to its ability to maintain context while chunking large volumes of text.

Chunking is an essential step in building retrieval-based systems because large documents must be broken down into more manageable pieces. In this context, a balance was struck between chunk size and computational efficiency:

- **Chunk Size:** A 4,000-word chunk was chosen because it fits well with the input token limits of 4,000 OpenAI's language models input token and ensures that the context within a chunk remains coherent.
- **Recursive Splitting:** The recursive splitting strategy allows the algorithm to adjust chunk boundaries intelligently, ensuring that the content isn't arbitrarily truncated in a way that may lose important context (e.g., between paragraphs or sections).

Text Embedding and Indexing

After chunking, each text block was passed through the OpenAI embedding model to generate a dense vector representation. Embeddings are critical in enabling the system to perform semantic search, as they convert textual information into numerical vectors that capture the meaning of the text.

Mathematical Representation

Let:

$\boldsymbol{T_i}$ represent the i-*th* text block in the set of text blocks.
$\mathbf{f(Ti)} = \boldsymbol{vi}$, Where vi is the dense vector representation (embedding) of dimension k generated by the embedding model $f(\cdot)$

Thus, the embedding generation process can be expressed as:

$$v_i = f(T_i), \quad \text{for each } i \in \{1,2,\ldots,n\}$$

Here, $f(Ti)$ maps the text block T_i to its corresponding embedding vector v_i which enables tasks like semantic search by comparing vectors in the embedding space.

We used OpenAI's embedding model to generate vector representations of each text chunk (*Figure 3*). Embeddings are a core component of most modern NLP models, as they encode the semantic content of text in a continuous vector space. The advantage of OpenAI embeddings lies in their pre-training on diverse datasets, enabling robust and generalizable representations of a wide variety of text types.

By encoding the text as vectors, we could transform the text retrieval problem into a vector similarity search problem. This transformation enables the model to retrieve chunks based on the semantic similarity between the user query and the embedded text.

Once the document chunks were embedded, they were stored in a vector index for fast and efficient retrieval. The FAISS (Facebook AI Similarity Search) library was chosen to build this vector store. FAISS is a widely used library for fast approximate nearest neighbor (ANN) search, particularly when working with large sets of dense vectors. FAISS supports both exact and approximate searches, with its approximate search algorithms being ideal for large-scale

ISSN: 2473-4901
v11 n6337

systems where retrieval speed is critical. In this system, FAISS provided the following benefits:

Efficiency: FAISS offers highly efficient indexing and search capabilities, even on datasets containing millions of vectors.

Scalability: The system can easily be scaled up to handle much larger or multiple documents without a significant increase in query latency.

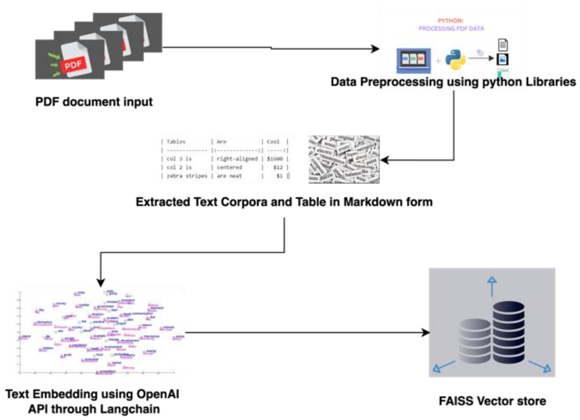

Figure 3: Document Embedding using FAISS Vector store.

Retrieval and Answer Generation
Each document chunk's embedding was added to the FAISS index. This allowed the system to perform a similarity search across the document when the user posed a query. The index was built using FAISS's IndexFlatL2 algorithm, which performs an exact search based on Euclidean distance between vectors. FAISS was chosen for its ability to handle large amounts of vector data in memory, providing a balance between accuracy and speed. By storing the embedded document chunks in this index, we enabled the model to perform quick lookups of relevant sections of the document based on a user's query.

The core of the retrieval process is based on cosine similarity, which measures the angular distance between vectors. Cosine similarity was chosen for this application because it has been shown to be highly effective in semantic search tasks, where the goal is to retrieve content based on meaning rather than exact keyword matches.

Given two vectors v_q and v_d representing the query embedding and document embedding, respectively, the cosine similarity $\cos(v_q, v_d)$ between the query and a document chunk can be computed using the following formula:

$$\mathbf{cos}(v_q, v_d) = \frac{\boldsymbol{v_q} \cdot \boldsymbol{v_d}}{|\boldsymbol{v_q}||\boldsymbol{v_d}|}$$

Where:

$\boldsymbol{v_q} \cdot \boldsymbol{v_d}$ is the dot product of the query vector and the document vector.
$|\boldsymbol{v_q}|$ is the Euclidean norm (magnitude) of the query vector.
$\boldsymbol{v_d}|$ **is** the Euclidean norm (magnitude) of the document vector.

The cosine similarity measures the cosine of the angle between these two vectors. A higher value (closer to 1) indicates higher semantic similarity between the query and the document chunk, while a value closer to 0 indicates less similarity.

Therefore, in the retrieval process:

- For each document chunk d in the FAISS index, compute its embedding v_d
- For a given query, compute its embedding v_q using the same embedding model.
- Cosine similarity between v_q and v_d .
- Rank the document chunks based on their cosine similarity scores and retrieve the most relevant chunks (those with the highest similarity).

When the user asks a question (query), the system first generates an embedding for the query using the OpenAI embedding. This query embedding will then be compared against the document embeddings stored in the FAISS index using cosine similarity to determine which chunks from the document are semantically closely related to the query.

Once the most relevant document chunks were retrieved, they were passed along with the user's question into the final question-answering stage. The retrieval process ensured that the model had access to only the most pertinent sections of the document, reducing noise and improving the quality of the generated response.

The system as shown in Figure 4 used LangChain, a framework designed to integrate document retrieval with LLMs to generate an answer from the retrieved document chunks. LangChain was selected due to its modular architecture, which facilitates easy integration of various components, such as document retrieval, question answering, and generation workflows.

 ISSN: 2473-4901
v11 n6337

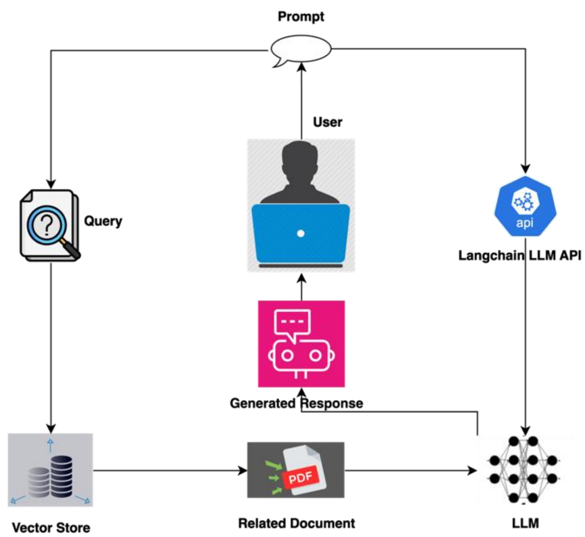

Figure 4: RAG Question and Answer working system.

Once LangChain combined the retrieved document chunks and the user's query, it was passed into OpenAI's language model for answer generation. The language model used the retrieved context to generate accurate, context-aware answers to the user's question. This hybrid approach, leveraging both retrieval and generation, ensured that the system could answer complex queries grounded in the document's content while benefiting from the broad linguistic capabilities of OpenAI's language models.

2. RESULTS

The evaluation of LangChain and LLM involves assessing the semantic accuracy of their answers compared to a ground truth answer. These answers were evaluated based on Bidirectional Encoder Representations from Transformers score (BERT) similarity, providing a detailed comparison of the effectiveness of each model in understanding and generating responses.

Model Performance on Various Questions

Various questions from the document were passed on to the models; their answers were evaluated against the ground truth using BERT. BERT scores quantify the semantic similarity between two pieces of text, and in this case, it was used to compare the answers generated by the models to the ground truth. A detailed breakdown of the questions, model answers, and their corresponding BERT scores is shown in the Appendix.

Each model demonstrated distinct strengths across different types of questions. The LLM was generally more consistent in providing concise and factual responses, particularly when the question directly aligned with specific document content. However, with more BERT score values, the RAG often outperformed the LLM in generating more precise and contextually richer responses that offered more detailed insights, especially when dealing with more complex or interpretive queries.

Research Result Hypothesis

Answers generated by RAG and LLM seem to show a significant difference in efficiency. To further demonstrate these differences, we conducted a human survey through Institutional Review Board (IRB) approved health professionals recruited from a prominent One Health conference. All subjects are verified experts in the OneHealth model. A total of 17 respondents participated, each rated 15 sets of questions and answers. There is a null hypothesis (H0) and an alternative hypothesis (H1).

H0 (null hypothesis): *No statistically significant difference exists between standard LLM and RAG.*

H1 (research hypothesis): *RAG significantly outperforms LLM.*

We used Qualtrics, an online survey platform, to conduct our survey. Out of 20 RAG questions, 14 were randomly shown to participants. The answers from the model and RAG were also randomized, ensuring that the order of the answer switches between the LLM and RAG responses. Participants were asked to rate the accuracy of the answers based on how closely they matched the ground truth, using a 1 to 7 scale. On this scale, a lower score indicated more substantial agreement with the answer, while a higher score indicated stronger disagreement. Data from the survey was collected and analyzed, and the hypothesis testing was followed by applying a t-test evaluation to assess the differences between the models. The results of the analysis are presented in the *Table 1.*

ISSN: 2473-4901
v11 n6337

	LLM	RAG
Mean	3.919	2.784
Variance	3.993	2.583
Observations	74	74
Pearson Correlation	0.225	
Hypothesized Mean Difference	0	
df	73	
t Stat	4.31	
P(T<=t) one-tail	0.0000	
t Critical one-tail	1.666	
P(T<=t) two-tail	0	
t Critical two-tail	1.993	

Table 1: T-test for Research Result Hypothesis.

The mean score for RAG (2.784) is lower than that for LLM (3.919), indicating that participants rated RAG answers as more accurate and closely aligned with the ground truth. The t-statistic of 4.31 exceeds the critical t-value for both one-tail (1.666) and two-tail (1.993) tests, indicating that the difference is statistically significant at the 5% level. The p-value for the one-tailed test is approximately 0.000000, which is well below the significance threshold (0.05). Therefore, we reject the null hypothesis and conclude that there is a statistically significant difference between LLM and RAG.

Since RAG has a lower mean score (indicating higher accuracy based on the rating scale), the results support the alternative hypothesis (H1) that RAG with BERT similarity improves robustness and significantly outperforms LLM model. Through human evaluation, findings shows that RAG significantly outperforms LLM regarding accuracy and alignment with the ground truth.

3. CONCLUSION

A comparative analysis was performed to evaluate the performance of a RAG model against a standard LLM in a question-and-answer system. The study aimed to assess the accuracy and alignment of responses generated by both models through a structured human survey. Participants rated the accuracy of the answers based on their agreement with the ground truth using a 1 to 7 scale, where lower scores indicated more substantial agreement and higher scores indicated stronger disagreement. A paired t-test was conducted to determine whether this difference was statistically significant. The t-statistic value of 4.31 exceeded the critical t-values for both one-tailed (1.666) and two-tailed (1.993) tests at the 5% significance level.

The p-value for the one-tailed test was approximately 0.000000, which is well below the significance threshold of 0.05. As a result, the null hypothesis (H_0) was rejected, and findings support the alternative hypothesis (H_1) that RAG significantly outperforms LLM in terms of accuracy based on human evaluation. The superior performance of RAG is attributed to its ability to combine external knowledge retrieval with language generation.

Unlike standard LLMs relying solely on pre-trained knowledge, RAG systems enhance accuracy by retrieving information from authoritative sources, ensuring that generated answers are contextually grounded and accurate. This hybrid approach enables RAG to provide more detailed and reliable answers, particularly for complex or fact-based queries.

These results underscore the potential of RAG-based systems to improve the accuracy and trustworthiness of AI-generated content, especially in real-time information validation and fact-checking applications. This research has several key limitations that will be addressed in future work, including a limited case study dataset, the exclusion of image content retrieval, and the use of a single LLM model for comparison.

In conclusion, the findings demonstrate that RAG significantly outperforms LLM regarding accuracy and alignment with the ground truth, highlighting the value of retrieval-augmented frameworks in delivering more reliable and context-aware responses.

ISSN: 2473-4901
v11 n6337

4. REFERENCES

Brucks, D., & von Bayern, A. M. P. (2020). Parrots Voluntarily Help Each Other to Obtain Food Rewards. *Current Biology*. https://doi.org/10.1016/j.cub.2019.11.030

Ghali, J. P. E., Shima, K., Moriyama, K., Mutoh, A., & Inuzuka, N. (2024). Enhancing Retrieval Processes for Language Generation with Augmented Queries. *arXiv preprint arXiv:2402.16874*.

Jiang, Z., Xu, F. F., Gao, L., Sun, Z., Liu, Q., Dwivedi-Yu, J., Yang, Y., Callan, J., & Neubig, G. (2023). Active retrieval augmented generation. *arXiv preprint arXiv:2305.06983*.

Koo, H., Kim, M., & Hwang, S. J. (2024). Optimizing Query Generation for Enhanced Document Retrieval in RAG. *arXiv preprint arXiv:2407.12325*.

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., & Maningo, J. (2023). Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS digital health*, *2*(2), e0000198.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., & Lau, A. Y. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), 1248-1258.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., & Rocktäschel, T. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, *33*, 9459-9474.

Lin, C., Miller, T., Dligach, D., Bethard, S., & Savova, G. (2019). A BERT-based universal model for both within-and cross-sentence clinical temporal relation extraction. Proceedings of the 2nd clinical natural language processing workshop,

Lin, D. (2024). Revolutionizing retrieval-augmented generation with enhanced PDF structure recognition. *arXiv preprint arXiv:2401.12599*.

Sawarkar, K., Mangal, A., & Solanki, S. R. (2024). Blended RAG: Improving RAG (Retriever-Augmented Generation) Accuracy with Semantic Search and Hybrid Query-Based Retrievers. *arXiv preprint arXiv:2404.07220*.

Shuster, K., Poff, S., Chen, M., Kiela, D., & Weston, J. (2021). Retrieval augmentation reduces hallucination in conversation. *arXiv preprint arXiv:2104.07567*.

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature medicine*, *29*(8), 1930-1940.

Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., Liu, S., Zeng, Y., Mehrabi, S., & Sohn, S. (2018). Clinical information extraction applications: a literature review. *Journal of biomedical informatics*, *77*, 34-49.

Yang, R., Ning, Y., Keppo, E., Liu, M., Hong, C., Bitterman, D. S., Ong, J. C. L., Ting, D. S. W., & Liu, N. (2024). Retrieval-Augmented Generation for Generative Artificial Intelligence in Medicine. *arXiv preprint arXiv:2406.12449*.

Yin, P., Neubig, G., Yih, W.-t., & Riedel, S. (2020). TaBERT: Pretraining for joint understanding of textual and tabular data. *arXiv preprint arXiv:2005.08314*.

ISSN: 2473-4901
v11 n6337

Appendix

Detailed breakdown of the questions, model answers, and their corresponding BERT scores

QUESTION 1 *What organizations make up the quadripartite?*

LLM ANSWER

The Quadripartite is composed of four international organizations: the World Health Organization (WHO), the Food and Agriculture Organization (FAO), the World Organisation for Animal Health (WOAH, formerly OIE), and the United Nations Environment Programme (UNEP). These organizations collaborate to address global health challenges at the human-animal-environment interface, focusing on areas like zoonotic diseases, food safety, and antimicrobial resistance. Their combined efforts aim to promote a unified approach to global health, known as the One Health initiative.

LANGCHAIN-ANSWER

The quadripartite is made up of the Food and Agriculture Organization of the United Nations (FAO), the United Nations Environment Programme (UNEP), the World Organisation for Animal Health (WOAH, founded as OIE), and the World Health Organization (WHO). These organizations collaborate to address health challenges at the human-animal-plant-environment interface. They have developed the One Health Joint Plan of Action (OH JPA) to promote sustainable health solutions.

LANGCHAIN-BERT-score 0.9366606

LLM-BERT-score 0.90872216

QUESTION 2 *Why is a joint Plan of Action needed?*

LLM ANSWER

A joint Plan of Action is needed to ensure coordinated efforts and alignment among stakeholders toward achieving a common goal. It helps in pooling resources, avoiding duplication of efforts, and maximizing the impact of collective actions. Moreover, it establishes clear roles and responsibilities, which enhances accountability and efficiency in addressing the targeted issue.

LANGCHAIN-ANSWER

A joint Plan of Action is needed to address the interconnected health threats to humans, animals, plants, and the environment in an integrated manner, promoting sustainable development. It aims to improve collaboration, coordination, and capacity-building across sectors to effectively prevent, predict, detect, and respond to health threats. Additionally, it seeks to leverage resources and expertise from various stakeholders to achieve comprehensive One Health outcomes.

LANGCHAIN-BERT-score 0.8954361

LLM-BERT-score 0.87486625

QUESTION 3 *What is the definition of One Health?*

LLM ANSWER

One Health is an integrated, unifying approach that aims to optimize the health of people, animals, and ecosystems by recognizing their interconnection. It involves collaboration across various disciplines, including human medicine, veterinary medicine, environmental science, and public health, to address complex health challenges. The One Health approach is critical for effectively managing zoonotic diseases, food safety, and environmental hazards that impact global health.

LANGCHAIN-ANSWER

One Health is an integrated, unifying approach that aims to sustainably balance and optimize the health of humans, animals, plants, and ecosystems. It recognizes that the health of humans, domestic and wild animals, plants, and the wider environment are closely linked and interdependent. The approach mobilizes multiple sectors, disciplines, and communities at varying levels of society to work together to foster well-being and tackle threats to health and ecosystems.

LANGCHAIN-BERT-score 0.9097894

LLM-BERT-score 0.91218096

QUESTION 4 *List the key objectives of the Joint Plan of Action.*

LLM ANSWER

The Joint Plan of Action (JPA) aimed to ensure that Iran's nuclear program would be exclusively peaceful by limiting its nuclear activities. It sought to provide Iran with limited, temporary, and reversible sanctions relief in exchange for verifiable actions that curtail its nuclear program.

 ISSN: 2473-4901
v11 n6337

Additionally, the JPA established a framework for future negotiations towards a comprehensive and long-term solution to Iran's nuclear issue.

LANGCHAIN-ANSWER

The key objectives of the One Health Joint Plan of Action (OH JPA) include providing a framework for collective and coordinated action to mainstream the One Health approach at various levels, offering policy and legislative advice and technical assistance to set national targets and priorities, and promoting collaboration, learning, and exchange among nations, sectors, and disciplines. Additionally, the OH JPA aims to take stock of existing cross-sectoral initiatives, identify synergies and overlaps, and support coordination. It also focuses on mobilizing and optimizing the use of resources across sectors, disciplines, and stakeholders.

LANGCHAIN-BERT-score	0.84955204
LLM-BERT-score	0.8311661

QUESTION 5 *How can the environmental sector contribute to the initiative?*

LLM ANSWER

The environmental sector can contribute to the initiative by promoting sustainable practices that reduce carbon emissions and conserve natural resources, ensuring long-term environmental health. By integrating eco-friendly technologies and supporting renewable energy solutions, it can drive innovation in reducing environmental impact. Additionally, collaboration with other sectors can help raise awareness and implement policies that prioritize sustainability in both local and global contexts.

LANGCHAIN-ANSWER

The environmental sector can contribute to the initiative by developing and implementing advocacy training and tools for decision-makers, conducting national environment sector needs assessments, and creating interoperable One Health training courses for professionals. Additionally, the sector can ensure the inclusion of environmental considerations in medical, public health, and veterinary training, and integrate systems thinking into core academic modules. These efforts will help influence One Health policies and ensure that environmental priorities are adequately represented.

LANGCHAIN-BERT-score	0.88010305
LLM-BERT-score	0.8791212

QUESTION 6 *What is the Joint Plan of Action?*

LLM ANSWER

The Joint Plan of Action (JPOA) is an interim agreement reached in 2013 between Iran and the P5+1 (the U.S., UK, France, Russia, China, and Germany) concerning Iran's nuclear program. It aimed to limit Iran's nuclear activities in exchange for temporary relief from economic sanctions while a more comprehensive agreement was negotiated. The JPOA served as the foundation for the subsequent 2015 Joint Comprehensive Plan of Action (JCPOA), commonly known as the Iran nuclear deal.

LANGCHAIN-ANSWER

The One Health Joint Plan of Action (OH JPA) is a strategic framework developed by the Quadripartite organizations (FAO, WHO, WOAH, and UNEP) to address health threats at the human-animal-environment interface. It aims to prevent, predict, detect, and respond to health threats while promoting sustainable development. The OH JPA outlines specific actions, deliverables, and timelines to achieve long-term and medium-term health outcomes through a coordinated One Health approach.

LANGCHAIN-BERT-score	0.8350722
LLM-BERT-score	0.83486396

QUESTION 7 *Describe the theoretical basis of the plan*

LLM ANSWER

The theoretical basis of the AI-Driven Agile Risk Management Tool lies in leveraging machine learning algorithms to predict risks by analyzing historical project data, team dynamics, and project complexity. This approach integrates principles from Agile methodologies, emphasizing adaptability and iterative improvements in risk management through real-time assessments. By utilizing predictive analytics, the tool provides proactive risk mitigation strategies, enabling teams to make informed decisions and maintain project momentum amidst uncertainties.

LANGCHAIN-ANSWER

The theoretical basis of the One Health Joint Plan of Action (OH JPA) is built on the premise that an integrated, multisectoral, holistic, and transdisciplinary approach can effectively address pressing health challenges. The theory of change is supported by three pathways: policy, legislation, advocacy,

ISSN: 2473-4901
v11 n6337

and financing; organizational development, implementation, and sectoral integration; and data, evidence, and knowledge. These pathways collectively aim to achieve the medium- and long-term outcomes of improved health for humans, animals, plants, and the environment, while reducing health threats at the human-animal-plant-environment interface.

LANGCHAIN-BERT-score	0.8406563
LLM-BERT-score	0.81940776

QUESTION 9 *What are the six action tracks?*

LLM ANSWER

The six action tracks refer to key focus areas outlined in various global frameworks for addressing critical challenges. These tracks typically include areas like health, climate action, poverty reduction, education, economic development, and technology innovation. Each track represents a targeted approach to solving specific issues within these broad domains, often fostering collaborative efforts between governments, organizations, and communities.

LANGCHAIN-ANSWER

"The six action tracks are:
1. Enhancing One Health capacities to strengthen health systems.
2. Reducing the risks from emerging and re-emerging zoonotic epidemics and pandemics.
3. Controlling and eliminating endemic zoonotic, neglected tropical, and vector-borne diseases.
4. Strengthening the assessment, management, and communication of food safety risks.
5. Curbing the silent pandemic of antimicrobial resistance.
6. Integrating the environment into One Health."

LANGCHAIN-BERT-score	0.83532774
LLM-BERT-score	0.84236777

QUESTION 9 *How does the plan of action build upon other initiatives?*

LLM ANSWER

The plan of action builds upon other initiatives by leveraging past successful strategies and incorporating lessons learned to enhance current objectives. It aligns with ongoing projects to create synergy, ensuring that resources and expertise are shared effectively. Additionally, it introduces innovative approaches to address gaps identified in previous efforts, aiming for more comprehensive and sustainable outcomes.

LANGCHAIN-ANSWER

The One Health Joint Plan of Action (OH JPA) builds upon other initiatives by complementing and supporting their implementation at various levels, including global, regional, and national. It strategically aligns with relevant initiatives such as the Global Action Plan on AMR, the Road Map for Neglected Tropical Diseases, and the WHO global strategy on health, environment, and climate change. The OH JPA offers a coordinated framework that integrates existing guidance, best practices, and tools from these initiatives to advance One Health objectives comprehensively.

LANGCHAIN-BERT-score	0.87298435
LLM-BERT-score	0.87678677

QUESTION 10 *How do the activities in 6.1 differ from the activities in 6.3?*

LLM ANSWER

The activities in 6.1 focus on preventive measures and surveillance systems to monitor public health risks, emphasizing early detection and proactive interventions. In contrast, the activities in 6.3 are centered on response and recovery, addressing the management of health crises and mitigation of their impact on affected populations. While 6.1 is more focused on preparedness, 6.3 deals with reactive measures and long-term recovery strategies.

LANGCHAIN-ANSWER

The activities in 6.1 focus on identifying and mitigating anthropogenic factors leading to environmental degradation, promoting sustainable resource management, and supporting policies for Indigenous Peoples and wildlife economies. In contrast, the activities in 6.3 are not detailed in the provided context, so specific differences cannot be fully outlined. However, 6.1 emphasizes practical measures for environmental conservation and policy support, while 6.3 likely involves different aspects of integrating the environment into One Health.

LANGCHAIN-BERT-score	0.89234847
LLM-BERT-score	0.89168185

 ISSN: 2473-4901
v11 n6337

QUESTION 11 *What action track has the highest number of deliverables?*
LLM ANSWER
The action track with the highest number of deliverables is typically the one focused on capacity building and stakeholder engagement. This track often includes numerous initiatives aimed at training, partnership development, and resource allocation. As a result, it generates a larger number of tangible outputs compared to more specialized tracks
LANGCHAIN-ANSWER
Action track 6, "Integrating the environment into One Health," has the highest number of deliverables. This track includes multiple activities aimed at promoting sustainable management of nature, enhancing ecosystem integrity, and integrating environmental knowledge into One Health decision-making. The deliverables span various timelines and involve extensive collaboration across sectors.

LANGCHAIN-BERT-score	0.88623744
LLM-BERT-score	0.90800357

QUESTION 12 *Summarize the recommendations for enhancing One Health capacities to strengthen health systems.*
LLM ANSWER
The recommendations for enhancing One Health capacities to strengthen health systems focus on improving cross-sectoral collaboration, particularly between human, animal, and environmental health sectors. They emphasize the need for stronger surveillance systems, better data sharing, and integrated response mechanisms to detect and manage health threats more effectively. Additionally, capacity-building initiatives, such as training and resource allocation, are recommended to ensure sustained coordination and preparedness across all sectors.
LANGCHAIN-ANSWER
The recommendations for enhancing One Health capacities include developing methodologies and tools for monitoring and evaluation, establishing financial needs and mechanisms for One Health investments, and advocating for political prioritization through economic case studies and national strategies. Additionally, it emphasizes the creation of a competent One Health workforce through joint processes, capacity building, and public engagement. These efforts aim to improve coordination, communication, and the integration of health systems at the human-animal-plant-environment interface.

LANGCHAIN-BERT-score	0.89351654
LLM-BERT-score	0.8894139

QUESTION 13 *In what ways does gender equality serve as a guiding principle in the Joint Plan of Action?*
LLM ANSWER
Gender equality serves as a guiding principle in the Joint Plan of Action by ensuring that health initiatives consider the specific needs and vulnerabilities of all genders. It promotes equal access to resources, healthcare, and decision-making opportunities, fostering inclusive policies that address disparities. Additionally, it emphasizes the importance of empowering women and marginalized groups in the implementation and leadership of health interventions.
LANGCHAIN-ANSWER
Gender equality in the Joint Plan of Action (OH JPA) supports gender equity and women's empowerment, taking a gender-sensitive perspective that considers vulnerabilities specific to each national context. This principle is consistent with the 2030 Agenda for Sustainable Development. The OH JPA emphasizes inclusiveness and equity, ensuring active participation and communication with communities, and recognizing the importance of local and traditional knowledge alongside scientific research.

LANGCHAIN-BERT-score	0.8834198
LLM-BERT-score	0.8693475

QUESTION 14 *What is the vision?*
LLM ANSWER
The vision is to create a healthier, more resilient world by fostering collaboration across human, animal, and environmental health sectors under the One Health approach. It aims to prevent, detect, and respond to health threats more effectively by integrating diverse expertise and resources.

ISSN: 2473-4901
v11 n6337

Ultimately, the vision aspires to build sustainable systems that safeguard global health and well-being for current and future generations.

LANGCHAIN-ANSWER

The vision (impact) of the One Health Joint Plan of Action (OH JPA) is to create a world better able to prevent, predict, detect, and respond to health threats while improving the health of humans, animals, plants, and the environment. This vision aims to contribute to sustainable development. The desired impact is expected to be achieved within a 15-20-year timeline.

LANGCHAIN-BERT-score 0.89629817

LLM-BERT-score 0.86143655

QUESTION 15 What factors influence the emergence of antimicrobial resistance?

LLM ANSWER

The emergence of antimicrobial resistance (AMR) is influenced by the overuse and misuse of antibiotics in both human medicine and agriculture, which accelerates the adaptation of microbes. Poor infection prevention and control practices in healthcare settings and insufficient sanitation in communities also contribute to the spread of resistant strains. Additionally, the lack of new antibiotic development and the global movement of people and animals further exacerbate the spread of AMR.

LANGCHAIN-ANSWER

The emergence of antimicrobial resistance (AMR) is influenced by the overuse and misuse of antimicrobials in human, animal, and plant sectors. Social and environmental factors such as insufficient access to health services, inadequate production and housing, lack of clean water, poor sanitation, waste management, and hygiene also accelerate the spread of resistant genes and pathogens. Additionally, insufficient regulatory frameworks and a lack of awareness and education about the risks of AMR and the appropriate use of antimicrobials contribute to the problem.

LANGCHAIN-BERT-score 0.918555

LLM-BERT-score 0.9042939

QUESTION 16 *What timeline is covered in the Joint Plan of Action?*

LLM ANSWER

The Joint Plan of Action typically covers a timeline of five years, providing a structured framework for addressing health threats through coordinated efforts. This period allows for phased implementation, monitoring, and evaluation of key activities aimed at strengthening One Health capacities. It also ensures flexibility to adapt to emerging health challenges within the specified timeline.

LANGCHAIN-ANSWER

The Joint Plan of Action covers a timeline of 1 to 5 years. Specific activities and deliverables are planned across this period, with some actions scheduled for completion within 1 year, others within 2-3 years, and the remainder within 4-5 years. The plan aims to achieve medium-term outcomes by 2026.

LANGCHAIN-BERT-score 0.8803032

LLM-BERT-score 0.8429157

QUESTION 17 Who provided input into the Joint Plan of Action?

LLM ANSWER

Input into the Joint Plan of Action was provided by a wide range of stakeholders, including representatives from human, animal, and environmental health sectors. Key contributors included international organizations like the World Health Organization (WHO), the Food and Agriculture Organization (FAO), and the World Organisation for Animal Health (WOAH). Additionally, national governments, academic institutions, civil society, and local communities played vital roles in shaping the plan through consultations and collaborative efforts.

LANGCHAIN-ANSWER

The Joint Plan of Action (OH JPA) was informed by evidence, best practices, and existing guidance from the Quadripartite. It also incorporated input from countries, international partners, and non-state actors such as civil-society organizations, professional associations, academia, and research institutions. The implementation of proposed actions at the national level involved consultation with Members, Member States, and State Parties.

LANGCHAIN-BERT-score 0.86545044

LLM-BERT-score 0.83925253

 ISSN: 2473-4901
v11 n6337

QUESTION 18 *What are common barriers observed in enhancing One Health capacities?*
LLM ANSWER
Common barriers to enhancing One Health capacities include a lack of coordinated communication and collaboration between human, animal, and environmental health sectors. Insufficient funding and resources often hinder the implementation of integrated health initiatives, while gaps in surveillance systems limit the early detection of health threats. Additionally, varying political priorities and regulatory frameworks across sectors and countries create challenges for unified action and sustained progress.
LANGCHAIN-ANSWER
Common barriers observed in enhancing One Health capacities include a lack of cooperation between internal and external stakeholders, limited engagement with the environmental sector, and professional segregation. Additionally, there is limited standardization around One Health curricula and competency-based frameworks. Commercial, academic, reputational, and profit motives often supersede collaborative efforts, further hindering progress.

LANGCHAIN-BERT-score	0.89567107
LLM-BERT-score	0.89926213

QUESTION 19 *List all the activities that need to happen in Year 1.*
LLM ANSWER
In Year 1, key activities include establishing cross-sectoral collaboration frameworks and setting up integrated surveillance systems for early detection of health threats. Capacity-building programs must be initiated to train personnel across human, animal, and environmental health sectors. Additionally, initial assessments and resource mapping are required to identify gaps in current health systems and allocate necessary resources for effective implementation.
LANGCHAIN-ANSWER
In Year 1, the activities that need to happen include developing frameworks and mechanisms for public participation, promoting One Health cross-sectoral collaboration and partnerships, developing operational tools to support science-based One Health coordinated strategic technical actions, and providing guidance on the appropriate use of integrated One Health information, surveillance, and emergency response systems. Additionally, guidance and tools for transparent and trusted One Health collaborative governance, effective communication structures, and mechanisms for joint funding and resource mobilization need to be provided. Other activities include promoting One Health task forces and working groups, identifying and quantifying anthropogenic factors leading to environmental degradation, enhancing private-sector and NGO engagement in sustainable natural resource management, and promoting the transition towards sustainable, climate-smart, agroecological approaches.

LANGCHAIN-BERT-score	0.88191336
LLM-BERT-score	0.8496687

QUESTION 20 *How does natural resource management contribute to the One Health approach?*
LLM ANSWER
Natural resource management contributes to the One Health approach by ensuring the sustainable use of ecosystems, which helps prevent the emergence and spread of zoonotic diseases. By maintaining healthy environments, it reduces human and animal exposure to pathogens that thrive in degraded or disrupted ecosystems. Additionally, responsible management of natural resources promotes biodiversity, which is critical for ecosystem resilience and the overall health of both people and animals.
LANGCHAIN-ANSWER
Natural resource management contributes to the One Health approach by preserving biodiversity and maintaining ecosystem health, which helps limit the emergence and transmission of diseases across the animal-human-plant interface. It also addresses environmental hazards such as pollution and climate change, which pose significant health threats to humans, animals, and plants. Integrating environmental considerations into One Health policies and interventions enhances disease prevention, preparedness, and promotes sustainable health solutions.

LANGCHAIN-BERT-score	0.9136984
LLM-BERT-score	0.9048505

