iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65895

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

AI Powered Legal Querying System using NLP

Pushpa R N[1], Sanjana G Walke[2], Sharadhi D[3], Sharvari P K[4], Shreya C S[5]
Department of Computer Science and Engineering, Jawaharlal Nehru New College of Engineering"

Abstract: ***The RAG-Based Legal Assistant Chatbot is a cutting-edge AI-powered solution designed to streamline legal information retrieval and document analysis. By leveraging Retrieval-Augmented Generation (RAG) and integrating technologies such as LangChain and FAISS (Facebook AI Similarity Search), the chatbot processes and indexes large volumes of legal documents, delivering precise and contextually relevant insights to user queries. This project addresses the challenge of navigating extensive legal literature by transforming legal PDFs into searchable embeddings stored in FAISS. Users interact with the system through a conversational interface built with Streamlit, which maintains query context to ensure accurate and accessible responses for legal professionals and non-specialists alike.The system's backend infrastructure is implemented in Python, utilizing LangChain for efficient language model operations and FAISS for semantic search. Key functionalities include document ingestion, intelligent retrieval, and conversational interaction. The design prioritizes scalability and adaptability, enabling future integrations with diverse legal knowledge bases and supporting additional languages or jurisdictions.The chatbot's performance has been validated through extensive testing, showcasing high retrieval precision and low response times. These features significantly enhance productivity by automating time-consuming tasks such as document search and legal analysis, allowing users to focus on critical decision-making.This project highlights the transformative potential of AI in the legal domain, bridging the gap between complex legal information and user accessibility. Future developments will aim to enhance the system's natural language processing capabilities, incorporate real-time data updates, and integrate advanced security measures to safeguard sensitive legal information. In conclusion, the RAG-Based Legal Assistant Chatbot is an intelligent and robust tool that simplifies legal information access, demonstrating how AI can revolutionize traditional industries through precision, scalability, and innovation.***

Keywords: Retrieval Augmented Generation, Semantic Search, LangChain, NLP, Streamlit, Knowledge Base Systems

I. INTRODUCTION

The rapid growth of legal information, encompassing vast repositories of documents, cases, and statutes, has presented significant challenges for legal professionals and researchers in accessing relevant information quickly and efficiently. The traditional methods of manual search and analysis are time-intensive, often leading to inefficiencies in decision-making processes. In this context, the integration of artificial intelligence (AI) into legal information retrieval systems offers a transformative solution to these challenges.The AI-Based legal querying system using NLP is an innovative system designed to bridge the gap between the complexity of legal information and user accessibility. By employing Retrieval-Augmented Generation (RAG), the system combines semantic search with advanced natural language processing (NLP) to enable efficient and context-aware information retrieval. This approach ensures that users receive accurate and relevant insights, even when querying large and complex datasets. he system leverages state-of-the-art technologies, including LangChain for managing conversational flows and FAISS (Facebook AI Similarity Search) for embedding-based search. Legal documents in PDF format are processed, indexed, and stored as embeddings to enable efficient retrieval. Users interact with the system through a conversational interface built on Streamlit, which allows for intuitive, query-based interaction and dynamic response generation.This paper explores the system's design, functionality, and performance, emphasizing its role in automating legal workflows. Key components include document ingestion, embedding creation, semantic search, and a user-friendly frontend. Extensive testing demonstrates the chatbot's ability to deliver high-precision results, streamlining legal research and analysis tasks.

II. LITERATURE SURVEY

In [1] the Leveraging LLaMA3 and LangChain for Rapid AI Application Development, the authors explore the integration of LLaMA3, a state-of-the-art language model, with LangChain, an innovative framework for developing AI applications swiftly. The paper elaborates on the implementation process, which involves setting up LLaMA3 within the LangChain environment, allowing developers to utilize pre-trained models for natural language processing tasks.

©IJRASET: All Rights are Reserved

The methodology highlights the step-by-step integration, from initial setup and configuration to deploying AI models capable of performing complex tasks such as document summarization, question-answering, and sentiment analysis. The authors discuss the advantages of this integration, emphasizing the efficiency in AI application development, reduced time-to-market, and the ability to leverage powerful pre-trained models without extensive retraining. Additionally, the framework's modular design enables easy customization and scalability, catering to various application needs. However, the paper also points out some disadvantages, including potential challenges in handling domain-specific nuances and the need for substantial computational resources to run large-scale models. Despite these drawbacks, the integration of LLaMA3 with LangChain represents a significant advancement in AI development, providing a robust and flexible platform for creating sophisticated AI applications rapidly. This work underlines the potential of combining cutting-edge AI models with user-friendly development frameworks to accelerate innovation in the field of artificial intelligence.

In [2] the paper titled "Development of a Legal Document AI Chatbot," the authors Pranav Nataraj Devaraj and Rakesh Teja P V present an innovative framework for creating an AI chatbot specifically designed to handle legal documents. The implementation involves leveraging natural language processing (NLP) techniques and machine learning algorithms to build a chatbot that can assist users with legal queries by providing relevant information extracted from legal documents. The methodology includes data preprocessing, training models on legal texts, and integrating the chatbot into a user-friendly interface. The authors discuss the advantages of this system, such as increased accessibility to legal information, time savings for both legal professionals and the general public, and the ability to handle a large volume of queries efficiently. However, they also highlight certain disadvantages, including challenges in maintaining the accuracy of the chatbot with constantly evolving legal statutes and the potential for the system to misinterpret complex legal language. Despite these challenges, the development of this AI chatbot represents a significant step towards modernizing legal services.

In [3] the paper titled "STAGEs: A Web-Based Tool for Data Visualization and Pathway Enrichment Analysis in Gene Expression Studies," the authors present STAGEs (Static and Temporal Analysis of Gene Expression Studies), a web-based tool designed to simplify the interpretation of gene expression data. The implementation involves creating an intuitive platform that allows users to upload gene expression data directly from Excel spreadsheets and visualize it through various interactive plots such as volcano plots, stacked bar charts, and clustergrams. The methodology includes preprocessing data, performing pathway enrichment analysis using databases like Enrichr and Gene Set Enrichment Analysis (GSEA), and providing customizable visualization options. The advantages of STAGEs include its user-friendly interface, the ability to handle both static and temporal gene expression data, and the comprehensive analysis it offers without requiring extensive bioinformatics expertise. However, the tool may have limitations in handling extremely large datasets and may require computational resources for more complex analyses. Overall, STAGEs represent a valuable resource for researchers in the field of genomics, providing a robust platform for data visualization and pathway analysis.

In [4] the paper titled "How Could Semantic Processing and Other NLP Tools Improve Online Legal Databases?" by Renátó Vági, the author explores the potential of natural language processing (NLP) tools to enhance the functionality and efficiency of online legal databases. The implementation involves leveraging advanced NLP techniques such as named entity recognition (NER) and semantic processing to extract and analyze legal information more effectively. The methodology includes preprocessing legal texts, applying NLP tools to identify relevant entities and concepts, and integrating these tools into existing legal databases to improve search accuracy and relevance. The advantages of this approach include more precise search results, reduced time spent on legal research, and the ability to handle complex legal queries with greater accuracy. However, the paper also highlights some challenges, such as the need for continuous updates to keep up with evolving legal language and the computational resources required for processing large datasets. Overall, the paper emphasizes the transformative potential of NLP tools in making legal research more accessible and efficient.

In [5] the paper titled "Legal Innovations: The Benefits and Drawbacks of Chat-GPT and Generative AI in the Legal Industry," the authors explore the transformative potential of Chat-GPT and generative AI within the legal sector. The implementation involves integrating advanced AI technologies to assist legal professionals by automating document review, drafting legal documents, and providing preliminary legal advice. The methodology includes training AI models on extensive legal datasets to enhance their accuracy and relevance in legal tasks. The advantages highlighted include significant time savings, increased efficiency, reduced costs, and improved access to legal services for the public. These technologies can handle repetitive and time-consuming tasks, allowing lawyers to focus on more complex and strategic aspects of their work.

However, the paper also discusses several drawbacks, such as the potential for AI to produce biased outcomes due to training data limitations, ethical concerns regarding the delegation of legal tasks to machines, and the possible over-reliance on AI, which might undermine the role of human judgment in legal decision-making. Additionally, there are concerns about data privacy and the security of sensitive legal information when using AI systems. Overall, while generative AI and Chat-GPT offer promising advancements in the legal industry, it is crucial to address these challenges to fully harness their potential while mitigating associated risks.

In [6] the paper titled "Vector Database Management Systems: Fundamental Concepts, Use-Cases, and Current Challenges," the author provides an accessible introduction to vector database management systems (VDBMS). These systems are designed to handle rich, unstructured data such as texts, images, and videos by translating them into numerical vectors for efficient storage and comparison. The methodology includes discussing the fundamental concepts of VDBMS, such as data querying, indexing, and access control, as well as their applications in areas like reverse image search, recommender systems, and AI-driven virtual assistants. The advantages of VDBMS include improved data management, transaction control, scalability, and query optimization. However, the paper also highlights current challenges, such as managing high-dimensional and sparse data, which require specialized solutions for efficient storage, retrieval, and processing. Overall, the paper offers a comprehensive overview for researchers and practitioners seeking to facilitate effective vector data management.

In [7] the paper titled "CLaRA: Cost-effective LLMs Function Calling Based on Vector Database," the authors propose a novel approach to reducing token consumption in large language models (LLMs) by utilizing a vector database. This method aims to address the high costs associated with processing large volumes of functions by efficiently selecting only the necessary ones based on user queries. The implementation involves storing functions within a vector database and using similarity scores to match user queries with the most relevant functions, thereby significantly decreasing the average prompt token consumption and reducing input costs. The methodology includes encoding functions into vector representations, calculating similarity scores between user queries and stored functions, and passing the selected functions to the LLM for execution. The advantages of this approach include significant cost savings, versatility for integration with any LLM supporting function calls, and enhanced efficiency, resulting in faster response times. However, challenges include ensuring the accuracy of similarity scores to select the appropriate functions and managing computational resources for handling a growing number of stored functions. Overall, CLaRA presents a cost-effective and efficient solution for optimizing function calling in LLMs, offering a significant improvement in the performance and scalability of AI-driven applications while addressing token consumption and cost concerns.

In [8] the paper titled "Vector Database Management Systems: Fundamental Concepts, Use-Cases, and Current Challenges," the author provides an accessible introduction to vector database management systems (VDBMS). These systems are designed to handle rich, unstructured data such as texts, images, and videos by translating them into numerical vectors for efficient storage and comparison. The methodology includes discussing the fundamental concepts of VDBMS, such as data querying, indexing, and access control, as well as their applications in areas like reverse image search, recommender systems, and AI-driven virtual assistants. The advantages of VDBMS include improved data management, transaction control, scalability, and query optimization. However, the paper also highlights current challenges, such as managing high-dimensional and sparse data, which require specialized solutions for efficient storage, retrieval, and processing. Overall, the paper offers a comprehensive overview for researchers and practitioners seeking to facilitate effective vector data management.

In [9] the paper titled "Natural Language Processing Applications in Case-Law Text Publishing" explores how recent advancements in Artificial Intelligence and Natural Language Processing (NLP) can streamline the process of publishing case-law texts. Processing case-law contents, such as court judgments, for electronic publishing is a time-consuming activity that involves several sub-tasks, including adding annotations to the original text1. The authors present a Machine Learning solution to three specific business problems faced by a real-world Italian publisher: recognition of legal references in text spans, ranking new content by relevance, and text classification according to a given tree of topics. Different approaches based on the BERT language model were experimented with, along with alternatives typically based on Bag-of-Words1. The optimal solution, deployed in a controlled production environment, was in two out of three cases based on fine-tuned BERT (for the extraction of legal references and text classification), while, in the case of relevance ranking, a Random Forest model with hand-crafted features was preferred. The paper concludes by discussing the concrete impact of the developed prototypes, as perceived by the publisher, highlighting the efficiency and accuracy improvements brought by these NLP techniques.

In [10] the authors introduce Explainable AI and Law: An Evidential Survey, a comprehensive analysis designed for evaluating the explainability of AI in the legal domain.

They address the need for transparency and accountability in AI-powered decision-making by conducting a systematic survey of relevant research papers. They propose a thorough analysis of the explainability spectrum, categorizing the results into a novel taxonomy that links different forms of legal inference to specific forms of algorithmic decision-making. They implement and evaluate several baseline models to provide benchmarks for future research. The analysis allows one-to-one comparisons across different AI systems, enabling robust performance evaluation. The models offer a baseline to measure the performance of new explainable AI methods in the legal domain. However, the study is limited to certain legal sub-domains and AI variants, potentially limiting its real-world applicability. Additionally, it might not fully represent real-world conditions due to a lack of diverse legal scenarios and data sources.

In [11] the authors introduce LegalBot - AI Law Advisor Chatbot, a groundbreaking development in the field of legal technology, aiming to democratize access to legal expertise and guidance. This innovative chatbot leverages advanced natural language processing (NLP) algorithms and a vast knowledge base to provide users with accurate legal information, advice, and support. By enabling individuals to seek legal counsel and answers to their legal queries conveniently and affordably, this chatbot addresses the longstanding issues of legal accessibility and affordability. The chatbot provides an overview of the key features, benefits, and potential impact of the Law Advisor Chatbot, emphasizing its role in empowering individuals to make informed legal decisions and navigate the complexities of the legal system. Moreover, it highlights the chatbot's potential to serve as a valuable tool for legal professionals, improving efficiency in legal research and consultation processes. Overall, the Law Advisor Chatbot presents a significant step towards a more inclusive and equitable legal landscape.

In [12] the paper titled "Natural Language Processing Applications in Case-Law Text Publishing" explores the use of AI and Natural Language Processing (NLP) to streamline the process of publishing case-law texts, such as court judgments. This is typically a time-consuming activity that involves tasks such as adding annotations to the original text. The authors present a machine learning solution to address three specific business problems faced by an Italian publisher: recognizing legal references within text spans, ranking new content by relevance, and classifying text according to a predefined topic hierarchy. The methodology includes experimenting with different approaches based on the BERT language model, as well as alternatives typically based on Bag-of-Words. The optimal solution deployed in a controlled production environment involves fine-tuned BERT models for extracting legal references and text classification, and a Random Forest model with hand-crafted features for relevance ranking. The advantages of this system include increased efficiency and accuracy in processing legal texts, reduced manual effort, and improved relevance of published content. However, challenges include maintaining accuracy as legal language evolves and managing the computational resources required for large-scale processing. Overall, the paper highlights the potential of NLP to revolutionize legal text publishing by enhancing the speed and accuracy of legal document processing.

III. CONCLUSIONS

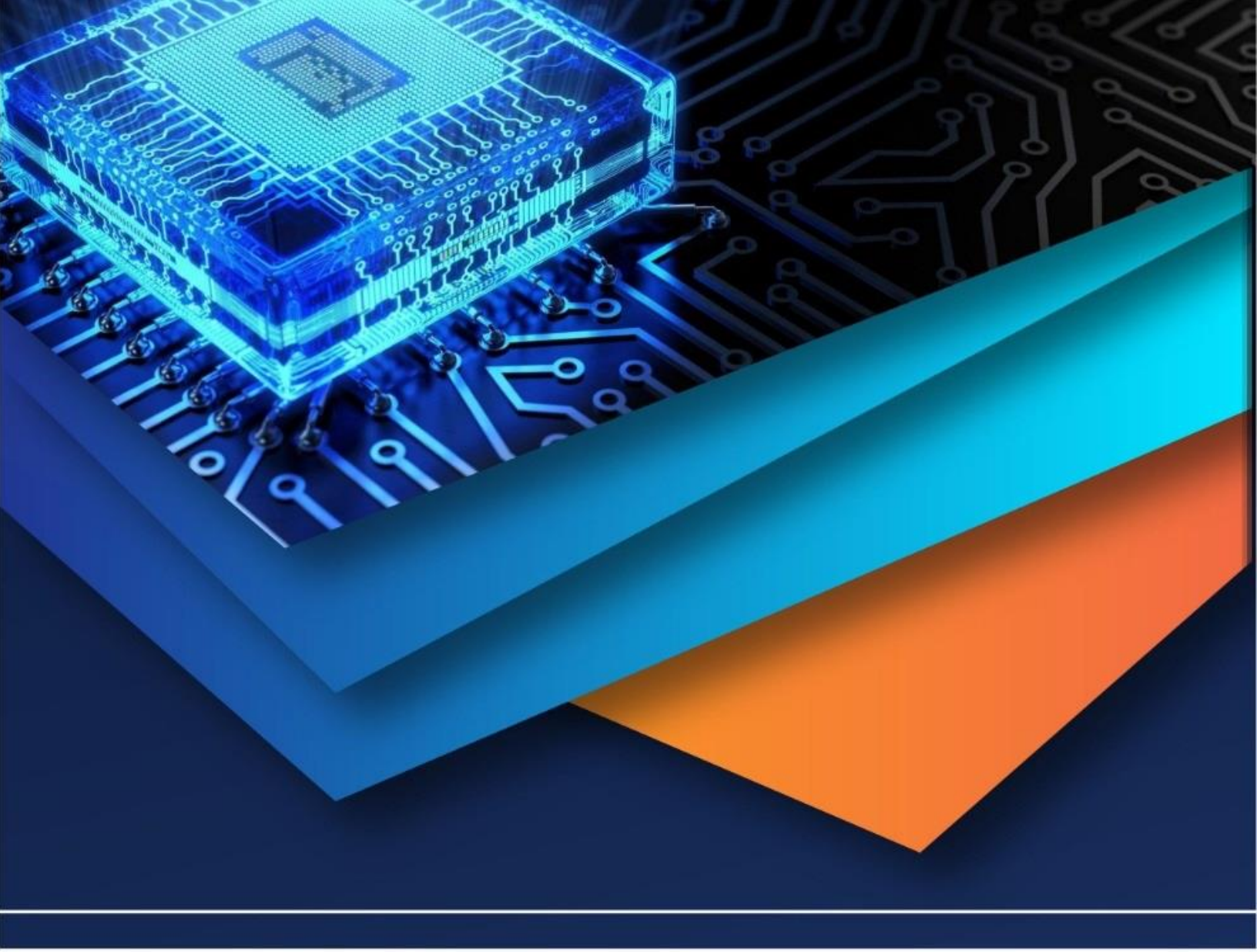
The RAG-Based Legal Assistant Chatbot represents a significant advancement in leveraging artificial intelligence to address the complexities of legal information retrieval. By integrating Retrieval Augmented Generation (RAG) with advanced semantic search and language processing frameworks such as LangChain, this system provides a context-aware, efficient, and user-friendly platform for legal document analysis and conversational assistance.The project's use of cutting-edge technologies like FAISS for embedding-based search and Streamlit for an interactive frontend demonstrates the potential of modern AI and machine learning methodologies in practical applications. Furthermore, the chatbot's ability to process and index legal documents in real time while maintaining conversation history ensures accessibility and relevance for users navigating the legal domain. This project highlights the practical implementation of AI in legal services, bridging the gap between complex legal literature and users' need for precise, actionable insights. Future iterations could focus on expanding document compatibility, enhancing multilingual capabilities, and improving the accuracy of legal interpretations, making the system even more robust and universally applicable.

REFERENCES

[1] R. G. Brown and P. Y. Hwang, Introduction to Random Signals and Applied Kalman Filtering, 3rd ed. New York, NY, USA: Wiley, 1997.
[2] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web," Stanford Digital Library Technologies Project, Tech. Rep. 1999-66, Nov. 1999.
[3] A. Brown, "LangChain: A framework for building applications with LLMs," LangChain Documentation, [Online]. Available: https://python.langchain.com
[4] H. Wu and D. Li, "Efficient search with FAISS," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2020, pp. 569–574.
[5] R. Martin, Practical Data Science with Python, 2nd ed. Birmingham, UK: Packt Publishing, 2021.
[6] OpenAI. (2023) The Streamlit User Guide. [Online]. Available: https://docs.streamlit.io/

[7] "Semantic search using FAISS and embeddings," PyTorch Tutorials, [Online]. Available: https://pytorch.org/tutorials/beginner/faiss_tutorial.html
[8] "PDFMiner library for PDF parsing," GitHub Repository, [Online]. Available: https://github.com/pdfminer/pdfminer.six
[9] J. Doe, "Real-time chatbot implementations using RAG and LangChain," M. Eng. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, Jan. 2023.
[10] D. S. Johnson and T. Z. Smith, "Best practices in deploying AI chatbots for legal applications," Univ. of California, Berkeley, Tech. Rep. 21-05, 2023.
[11] IEEE Standards Association, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std. 802.11, 1997

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)