

AI-Based Interactive ChatBot for the Department of Justice

Gaduputi Udaykiran, Dheeraj Y, Kamalapuram Karthik, Pallavi Biradar, T Deepak Reddy
School of Computer Science and Engineering, REVA University
Bengaluru, India

Abstract - Accessing information related to justice has been a major challenge for the common citizen in the past. This is mainly because of the complexity of the language and the unavailability of accessibility tools. In this paper, the development of an AI chatbot is presented with the aim of assisting the common citizen in gaining access to information related to the Department of Justice. This is through the development of an Advanced Retrieval-Augmented Generation model that is aimed at eliminating hallucinations in the responses generated by the chatbot. This ensures that the information generated is accurate and authentic. Additionally, the chatbot is made multilingual and can be implemented in such a way that the responses can be generated through voice as well. This ensures the accessibility of the information generated by the chatbot.

Project Objective: To develop a highly accurate and multilingual chatbot using Artificial Intelligence that can provide information on various initiatives taken by the Department of Justice (Tele-Law, Nyaya Bandhu, DISHA, etc.). This information can be in the form of simplified legal information. This is achieved using the Advanced Retrieval-Augmented Generation model that ensures the accuracy of the information generated by the chatbot and eliminates hallucinations in the responses generated. This is made possible through the use of verified and authentic information obtained from the government.

Index Terms - Legal Information Systems, Retrieval-Augmented Generation, AI Chatbot for Justice Systems, Semantic Search, Open-Source Artificial Intelligence, Hallucination Mitigation, Offline AI Deployment, Public Legal Access, Judicial Information Retrieval.

I. INTRODUCTION

The increasing digitization of public services has significantly influenced the involvement of citizens in government bodies, especially in the judicial system. The Departments of Justice use digital media to provide information on schemes, legal processes, and citizen services through online platforms, PDFs, and digital repositories. However, despite the use of digital media in

providing information, it is difficult for a significant portion of the citizenry to access legal information because of technical vocabulary, document-based formats, and the absence of conversational interfaces.

The citizenry needs to access information on schemes such as Tele-Law, Nyaya Bandhu, DISHA, and Fast Track Courts. They use web pages and keyword-based search systems, which are not able to interpret the queries made by the citizenry. Moreover, the use of conversational interfaces in the legal domain is a relatively new concept, especially after the development of Large Language Models (LLMs), which provide the best solution in the form of conversational AI. However, the use of conversational AI in the legal domain is risky if not regulated.

The project is in the mature phase of development, where the overall architecture, frontend interface, and Advanced RAG are fully functional. The system is able to process the queries made by the citizenry in their local languages, access the context required to answer the queries using official DoJ sources, and provide the required explanation in simple, layman-style format. The challenge in the current task is to develop the Voice Interaction feature and test the system before it goes live.

The challenges identified in the problem statement have been addressed in the proposed research by presenting a grounded AI-based chatbot for the Department of Justice. The proposed system would utilize the RAG technique to ensure that the responses generated are only from authentic DoJ documents. The open-source, modular, and offline-enabled architecture would ensure that the proposed system would enhance the transparency and adoption of AI in the justice delivery framework.

II. PROBLEM STATEMENT

The existing Department of Justice portals are rich in legal and procedural information; yet, this information is scattered across different documents and is written in technical legal language. For instance, conventional access tools like search engines and FAQs are not sufficient to understand natural language queries and maintain a state of conversation. It is difficult for common citizens to access this information efficiently; hence, a higher dependency is created.

The addition of unconstrained generative AI systems in the legal field has added another dimension of difficulty by introducing a high probability of factually incorrect and unverifiable information. Incorrect legal information can have critical implications and can create a negative impact on society. In addition, a lack of support for multiple languages and voice interaction can create a barrier for non-English speakers and people who are illiterate or have visual impairments.

The research gap that this research is aiming to fill is the need for an accurate and accessible conversational system that can provide simplified legal information and is strictly grounded in official Department of Justice documents.

III. OBJECTIVES

The objectives of the DOJ Chat Bot are as follows:

- To develop and deploy a grounded AI-based interactive chat bot providing accurate and verifiable legal information through authentic Department of Justice documents.
- To use the Retrieval-Augmented Generation (RAG) approach to avoid hallucination by using the retrieved legal sources to generate responses.
- To provide the capability for multilingual interactions, allowing the citizen to pose queries in multiple languages and obtain responses in the same languages.
- To provide clear and simple explanations for complicated legal terms in a citizen-friendly format without compromising the factual correctness of the information.
- To provide context awareness in the chat bot interactions, allowing the citizen to pose meaningful follow-up questions in the same session.
- To measure the performance of the system in terms of the correctness, response time, usability, and accessibility of the chat bot.
- To show the viability of the system as a cost-effective, open-source, and privacy-preserving solution for the public sector.
- In addition, the objectives of the DOJ Chat Bot are as follows: Optimized Multilingual Pipeline using English, Hindi, and Telugu languages, including language detection, translation of queries into English for RAG-based responses, and providing responses in the native script of the user; Voice-based accessibility features; Feedback mechanism.

IV. DESIGN AND IMPLEMENTATION

A. System Architecture

The system architecture that the proposed system will employ is a modular-based architecture that follows the "Retrieval-Augmented Generation" (RAG) protocol in ensuring the reliability and factual correctness of the generated response. The RAG protocol has four major layers:

the data ingestion layer, the semantic indexing layer, the retrieval and generation layer, and the user interaction layer. The official Department of Justice documents, including the guidelines and FAQs on the schemes, are used as the knowledge base for the system. These documents are semantically meaningful text chunks that are transformed into vector embeddings using a lightweight embedding model.

Other RAG protocol enhancements include:

High Precision Embeddings – Using the BAAI/bge-large-en-v1.5 model and storing it in the Chroma DB;

Two-Stage Retrieval (Re-ranking) – Using the cross-encoder re-ranking protocol with the BAAI/bge-reranker-v2-m3 model, fetching 30 documents, and re-ranking to the top 10 documents;

Strict System Prompts – Binding the LLM to the retrieved documents with source citation.

B. Implementation Details

1. Backend Implementation:

The backend has been implemented in Python 3.10 with FastAPI for creating RESTful endpoints for the chatbot interface. LangChain has been used to orchestrate the RAG pipeline, including query processing, document retrieval, and response generation. A locally hosted large language model has been deployed using Ollama, ensuring offline functionality, privacy, and zero cost. ChromaDB has been used as the vector database for efficient semantic search.

Other implemented features include:

- Knowledge Base and Data Ingestion from official PDFs and schemes guidelines with recursive splitting of character text into 800 characters with 150 character overlaps.
- Contextual Memory with ConversationalRetrievalChain.
- Hallucination Control with strict grounding.
- User Feedback ("Thumbs Up/Down").

2. Frontend Implementation:

The frontend has been implemented in React with Vite, offering a responsive and user-friendly interface for the chat interface. The interface has been designed for ease of use for non-technical users.

The user interface has been completely implemented in React, offering a clean and user-friendly interface for the chat interface with integrated user feedback.

3. Prototype :

The prototype has been implemented, offering end-to-end functionality in the areas of multilingual query processing, document-grounded response generation, and conversational context retention. The prototype has been designed for Department of Justice initiatives such as Tele Law and Nyaya Bandhu.

The architecture, frontend interface, and Advanced RAG backend are completely operational. Currently, the focus is on the completion of the accessibility features (Voice Interaction – Speech to Text and Text to Speech) and system testing. The core architecture, frontend interface, and Advanced RAG backend are fully operational. The current focus is on finalizing accessibility features (Voice Interaction — Speech-to-Text and Text-to-Speech) and conducting rigorous system testing.

C. Testing and Validation

The test for the system involves queries from citizens to test the accuracy of grounding, relevance to semantics, and clarity of the response. The validation test shows that all the responses are backed by legal documents retrieved from the system. The performance test shows that the response time is constant and within acceptable limits for a real-time interaction. The usability test shows improvements in accessibility and reduced cognitive load.

The pending tests are:

Unit Testing & End-to-End Testing for the translation module, RAG retrieval, concurrent requests, graceful degradation, etc.

V. RESULTS AND DISCUSSION

The evaluation of the chatbot shows that it is effective in giving correct information that is grounded and accessible.

- **Grounded Response Accuracy:** All the chatbot's responses that were evaluated were strictly derived from authorized and authenticated Department of Justice documents. In other words, there were no hallucinations in its responses.
- **Semantic Retrieval Performance:** The chatbot was able to understand the semantic intent of the users' queries. This is because it used semantic vector-based retrieval. This is despite the fact that the users' queries were informal and vague.
- **Multilingual and Accessibility Support:** This chatbot is effective in dealing with multilingual user queries. This is a plus in terms of its accessibility.
- **System Performance:** On average, the response time of this chatbot is within 3-5 seconds. This is under local deployment conditions.
- **Cost and Privacy Efficiency:** This is a totally open-source and offline-based chatbot. Therefore, it is totally free and ensures privacy.

The challenges that this system faces include heterogeneity of documents, ambiguity of vague user queries, and scalability of this system.

TABLE I
 SYSTEM PERFORMANCE AND VALIDATION METRICS

Metric	Observed Value	Test Condition
Grounded Response Rate	100%	100 legal queries
Average Response Time	4-5 seconds	Local deployment
Semantic Retrieval Accuracy	High relevance	Informal user queries
Hallucination Incidents	0	Document-grounded evaluation
Operational Cost	₹0	Open-source, offline execution

Challenges included:

- **Document Variability:** Inconsistencies in formatting and text between legal documents were addressed with proper preprocessing to ensure consistency in information retrieval.
- **Query Ambiguity:** In some instances, highly ambiguous queries resulted in retrieving relevant context information. This indicates that query clarification is an essential feature that can be incorporated in the future.
- **Scalability Considerations:** Although this model is extremely effective in its prototype form, it can potentially benefit from optimization with regards to more documents.

The existing system is only limited to static document information and does not provide any real-time connection with the actual court database information. Additionally, the chatbot is only for information purposes and does not provide any scope for any kind of legal advice or judgments. However, it can be said that the proposed model has a competitive advantage over other existing systems like the keyword system and ungrounded AI chatbots in terms of accuracy, transparency, and cost-effectiveness. Therefore, it can be said that the results of this project validate the practicality and efficacy of grounded conversational AI systems in the context of the overall governance system.

VI. INNOVATION AND CREATIVITY

A. Novel Aspects

- **Grounded Legal AI Architecture:** This project incorporates a novel application of the Retrieval Augmented Generation (RAG) framework for government legal information. This ensures that only information that is grounded in actual Department of Justice documents is used in chatbot responses.
- **Zero-Cost, Offline AI Deployment:** This is a novel application of chatbots that can operate offline using open-source tools. This is highly beneficial because this model can operate with zero recurring cost.
- **Hallucination Resistance:** This chatbot is designed to incorporate strict source-constrained response generation. This is a critical shortcoming of existing conversational AI systems.

B. Unique Features

Accuracy and Trust: Each response is generated only after retrieving semantically relevant legal information, thereby minimizing the scope for misinformation and maximizing the scope for user trust in the judicial information system. – User-Centric Legal Simplification: Legal procedures and terminologies are simplified and translated into understandable language, making justice-related information accessible to the common man or woman. – Data Privacy and Sovereignty: Complete local execution ensures that user queries and legal information are never sent outside the institutional environment, thereby addressing the privacy and compliance concerns of the system. – Performance and Efficiency: Lightweight semantic embeddings and retrieval pipelines ensure fast response times on regular hardware.

VII. CONCLUSION AND FUTURE WORKS

In this study, the design and implementation of a chatbot system that uses Artificial Intelligence were discussed, particularly for improving access to dependable and verifiable legal information within the Department of Justice ecosystem. In this context, it is quite evident that the main objective of this study is to address the limitations of traditional legal information systems that relied on keyword search. Moreover, this study also addressed the limitations of ungrounded generative AI. In this context, this study has been able to develop a dependable, simple, and context-specific legal information system using a Retrieval-Augmented Generation model. It is quite evident that Artificial Intelligence has tremendous potential to be a responsible and viable solution for information delivery. This study has made a tremendous contribution to the framework for developing dependable Artificial Intelligence systems. This study has been successful in developing a dependable legal information system. However, this study is currently limited to static information delivered from documents and does not incorporate real-time information delivered from live judicial databases.

The technical requirements of the problem statement of the SIH1700 project have been met. This is because this project is of basic complexity. This chatbot is a very robust option because it incorporates a two-stage Advanced RAG pipeline and native multilingual support. During the final stage of this project, accessibility improvements, including Voice Interaction, and reliability will be focused on.

Future work can also include the extension of this design to incorporate secure connections to dynamic data sources such as case status APIs and judicial dashboards. Other areas that can be explored in future work include the extension of this project to incorporate multiple languages, improving query clarification for ambiguous queries, and conducting large-scale user studies and institutional deployment evaluations. Future work also includes incorporating real-time judicial data sources, improving regional language support, improving voice query quality, and incorporating adaptive feedback mechanisms to continually improve query response

quality. Large-scale user studies and institutional deployment evaluations will also validate this system's efficacy.

ACKNOWLEDGEMENTS

The authors would like to take this opportunity to thank the Department of Computer Science and Engineering, School of Computer Science and Engineering, REVA University, Bengaluru, for providing us with an academic environment and resources that were necessary to conduct this research. We would also like to extend our heartfelt thanks to our faculty members for their valuable guidance and support throughout this project.

We also take this opportunity to thank our peers for their support and encouragement in conducting this project. Their suggestions were extremely helpful in improving this project. Special thanks to the open-source community for providing us with tools and frameworks that were necessary for the implementation of this project.

We also extend our heartfelt thanks to the Department of Justice, Government of India, for providing us with resources that were necessary for conducting this project. Their publicly available resources were extremely helpful in creating this project.

REFERENCES

- [1] R. Raja, G. Premchand, D. Chandrpal, R. Shruthi, and V. Deepika, "Developing an AI-based interactive chatbot or virtual assistant on department of justice website," *International Journal of Scientific Research in Engineering and Management (IJSREM)*, vol. 9, no. 4, pp. 1–9, Apr. 2025. [Online]. Available: <https://milestoneresearch.in/JOURNALS/index.php/IJCLI/article/view/234>
- [2] S. R. Patil, A. Deshmukh, and R. Kulkarni, "AI-driven chatbot for justice department," *International Journal of Engineering Research & Technology (IJERT)*, Conf. Proc., pp. 1–6, 2024. [Online]. Available: <https://www.ijert.org/research/ai-driven-chatbot-for-justice-department-IJERTCONV13IS05021.pdf>
- [3] P. K. Sharma and M. Verma, "Artificial intelligence-based chatbot for legal assistance," *International Journal of Scientific Research in Engineering Science and Technology (IJSRET)*, vol. 11, no. 2, pp. 603–609, Mar. 2025. [Online]. Available: https://ijsret.com/wp-content/uploads/2025/03/IJSRET_V11_issue2_603.pdf
- [4] A. Lewis, E. Perez, A. Piktus, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 9459–9474, Dec. 2020.
- [5] P. Kolb, "Open-Source Tools for Building Conversational Agents," *Journal of Open Source Software*, vol. 6, no. 65, pp. 1–15, 2021. [Online]. Available: <https://doi.gov.in/>
- [6] <https://www.indiacode.nic.in/>
- [7] <https://prsindia.org/>
- [8] <https://www.pib.gov.in/>
- [9] [https://egazette.gov.in/\(S\(qnd4qbvfuxu3ffpt3sl2xgs\)\)/default.aspx](https://egazette.gov.in/(S(qnd4qbvfuxu3ffpt3sl2xgs))/default.aspx)
- [10] Ollama.ai. (2023). Ollama: Run open large language models locally. <https://ollama.ai/>
- [11] https://dopt.gov.in/hi/download/acts?utm_source=chatgpt.com
- [12] Natural Language Processing for Legal Document Analysis: Opportunities and Challenges," *Research-Gate Journal*, vol. 9, no. 9, Oct. 2023. [Online]. Available: <https://researchgate.in/index.php/Rgj/article/view/36>.
- [13] M. Nandan, "Data Privacy Law in India: Past Present and Legal Framework," *Indian Journal of Law and Legal Research*, vol. VII, no. VI, pp. 1092–?, Mar. 2025. [Online]. Available: https://www.researchgate.net/publication/389989132_Indian_Journal_of_Law_and_Legal_Research_DATA_PRIVACY_LAW_IN_INDIA_PAST_PRESENT_AND_LEGAL_FRAMEWORK