International Journal of
Computational Learning & Intelligence
Vol. 04 | Issue 04 | 2025

RESEARCH ARTICLE OPEN ACCESS

AI-Powered Interactive Legal Chatbot for the Department of Justice

K L Srujan Surya . Edwin Mark K . Kushal S . Charan Tej P. Afifa Salsabil Fathima

School of Computer Science and Engineering,
REVA University, Bengaluru, India

DOI: **10.5281/zenodo.15465000**
Received: 27 April 2025 / Revised: 13 May 2025 / Accepted: 19 May 2025

Abstract — This paper details the development and conceptual framework of an AI-based interactive chatbot designed for the Department of Justice website in India, aimed at bridging the significant accessibility gap in public legal assistance. The system leverages a Flask backend, a JavaScript-powered frontend, and the LLaMA 3.1 large language model accessed via Ollama to provide real-time legal query responses, summarization of legal documents (PDF/DOCs including FIRs and petitions), and basic multilingual capabilities (English/Hindi) with citation support. The Paper focuses on creating a free, user-friendly, and scalable virtual legal assistant tailored for common public queries and document analysis, thereby enhancing legal literacy and empowering citizens. This paper outlines the chatbot's system architecture, core feature implementation (including prompt engineering and translation services), illustrative use cases demonstrating its functionality, and discusses the technical challenges, ethical considerations, and potential societal impact of deploying such a system in the Indian legal context. It concludes by summarizing the Paper's contributions and suggesting avenues for future development to further improve its efficacy and reach in democratizing legal information.

Keywords — AI assistant, justice system, legal chatbot, legal document summarization, LLaMA 3.1, multilingual NLP, public legal aid

I. INTRODUCTION

The pursuit of justice and equitable access to legal information forms a cornerstone of modern democratic societies. However, navigating the complexities of legal systems often presents significant challenges for the general public. These challenges range from understanding intricate legal jargon and procedures to the financial burden associated with seeking professional legal counsel. In many jurisdictions, including India, there is a discernible gap between the legal needs of the populace and the accessibility of timely, understandable, and affordable legal assistance. This gap can disproportionately affect marginalized communities and those with limited literacy or resources, thereby hindering their ability to exercise their rights effectively. The advent of Artificial Intelligence (AI), particularly in the domain of Natural Language Processing (NLP) and Large Language Models (LLMs), offers a transformative potential to mitigate these challenges by democratizing access to legal information and support services.

This paper details the development and conceptual framework of an AI-based interactive chatbot designed for the Department of Justice website in India. The primary problem this Paper seeks to address is the limited public accessibility to legal assistance and the difficulties citizens face in understanding legal documents and procedures. Many individuals struggle to find reliable answers to common legal queries, comprehend complex legal texts such as First Information Reports (FIRs) or legal petitions, and are often unaware of their rights or the appropriate channels for redressal. The proposed chatbot aims to serve as a virtual legal assistant, offering real-time responses to legal queries, summarization of legal documents (PDFs and DOCs), and basic multilingual support in English and Hindi. By leveraging the capabilities of advanced LLMs like LLaMA 3.1, accessed via the Ollama framework, and a user-friendly interface built with Flask and JavaScript, the Paper endeavors to provide a scalable and freely accessible tool tailored to common public legal needs.

The objectives of this Paper are multifaceted. Firstly, it aims to provide accurate and easily understandable information regarding Indian laws, court procedures, and legal rights, thereby enhancing legal literacy among citizens. Secondly, the chatbot is designed to assist users in comprehending complex legal documents through automated summarization, focusing on key points, relevant sections, and potential implications. This includes the analysis of FIRs and legal petitions, which are often critical in legal proceedings. Thirdly, by offering multilingual support, initially in English and Hindi, the Paper strives to cater to a broader demographic, breaking down language barriers that can impede access to justice. A crucial aspect of the chatbot's functionality is its ability to provide citations for the information it delivers, promoting transparency and allowing users to verify the source of the legal insights. The significance of this Paper lies in its potential to bridge the existing accessibility gap in public legal assistance. It represents a novel application of cutting-edge AI technology to a critical public service domain, offering a user-friendly, cost-effective, and scalable solution. The contribution extends beyond mere information provision; it aims to empower citizens by equipping them with the knowledge necessary to navigate the legal landscape more confidently.

This paper will further explore the existing landscape of AI in legal services and public administration, highlighting relevant work and identifying the unique contributions of this Paper. It will

then delve into the system architecture and design, detailing the technological stack including Flask for the backend, LLaMA 3.1 via Ollama for the core AI capabilities, and a JavaScript-powered frontend. The implementation of key features such as real-time query processing, document summarization, and multilingual support will be discussed. Subsequently, the paper will present illustrative use cases, analyze the challenges encountered and the limitations of the current system, and discuss the potential impact of such a tool on enhancing legal accessibility. Finally, it will conclude with a summary of the Paper and outline directions for future work and enhancements.

II. SYSTEM ARCHITECTURE AND DESIGN

Overall System Architecture

The system comprises three primary layers: a client-side frontend, a server-side backend, and an LLM service.

- Frontend (Client-Side): Developed using HTML, CSS, and JavaScript, the frontend provides the user interface for interacting with the chatbot. It handles user input (text and voice), displays chat messages, manages document uploads, and facilitates language selection. It communicates with the backend via HTTP requests.
- Backend (Server-Side): Built using the Flask Python web framework, the backend serves as the central orchestrator. It exposes API endpoints to handle chat requests and document uploads from the frontend. It processes these requests, interacts with the LLM service for generating responses or summaries, manages file operations, and handles language translation.
- Large Language Model (LLM) Service: The core intelligence is provided by LLaMA 3.1, accessed locally via the Ollama framework. The backend sends structured prompts to the Ollama API, which in turn queries the LLaMA

Frontend Development

The frontend is a single-page application structured with HTML for content, CSS for styling, and JavaScript for dynamic interactions. The index.html file defines the main layout, including sections for an introduction to the chatbot, legal resources, FAQs, the chat interface, and a document upload facility. Styling, as defined in styles.css, aims for a professional and accessible design, incorporating responsive elements to ensure usability across various devices. Key UI elements include a chat window for displaying conversation history, an input field for text queries, a voice input button leveraging the browser's Web Speech API (webkitSpeechRecognition), and controls for language selection (English, Hindi, Tamil, Bengali).

The script.js file manages all client-side logic. This includes:

• Chat Interaction: Capturing user messages, sending them to the backend /chat endpoint, and displaying both user and streamed bot responses in the chat window. Responses from the backend are streamed to provide a real-time feel. • Voice Input: Activating the microphone, capturing speech, converting it to

text, and populating the chat input field. • Document Upload: Handling file selection (PDF, DOC, DOCX, TXT), sending the file to the backend /upload_document endpoint using FormData, and displaying the summarized document or status messages in the chat interface. • Language Switching: Allowing users to select their preferred language. This selection is communicated to the backend via the Accept-Language header in API requests and also used to set the lang attribute of the HTML document for client-side localization and speech recognition language settings.

Large Language Model Integration

The system integrates with LLaMA 3.1 through Ollama, which serves the model locally. The backend makes HTTP POST requests to Ollama's /api/generate endpoint.

• Prompt Engineering: Specific system prompts are used for different tasks: • For legal queries (/chat): The system prompt instructs the AI to act as an assistant for the Indian Judiciary system, focus on Indian laws, avoid legal advice, provide resources, and give direct answers without internal monologue. • For document summarization (/upload_document): The prompt directs the LLM to summarize the legal document focusing on key legal points, important sections, implications, and next steps, using clear language and providing Indian legal references if possible. It also includes a disclaimer about not being legal advice. • Streaming Responses: Both chat and document summarization endpoints utilize stream=True with the Ollama API and return a Response object with a generator function in Flask. This allows the frontend to receive and display the LLM's output token by token, enhancing perceived responsiveness. • Model Specificity: The /chat endpoint uses the model named "llama3", while the /upload_document endpoint specifies "llama3.1". This might indicate the use of different model versions or configurations for different tasks, or it could be an inconsistency.

Multilingual Capability Implementation

Multilingual support (initially English and Hindi, with UI elements for Tamil and Bengali) is a key feature.

• Frontend: The user can select a language. This selection updates the Accept-Language header sent with API requests and adjusts the language for the Web Speech API. • Backend: The Flask backend checks the Accept-Language header. If the language is not English, the TranslationService is used to translate the incoming user message to English before processing by the LLM. Similarly, the LLM's English response is translated back to the user's selected language before being sent to the frontend. This approach allows the core LLM interaction to occur in English, presumably the language LLaMA 3.1 is most proficient in, while still offering a multilingual user experience.

III. IMPLEMENTATION DETAILS

Technology Stack

• Backend: Python 3.x, Flask (web framework), Requests (for HTTP calls to Ollama). • Frontend: HTML5, CSS3, JavaScript (ES6+). • LLM: LLaMA 3.1 served via Ollama. • Document Processing: A custom document_processor module (handles PDF, DOC, DOCX, TXT) and a translation_service

module. • Development Environment: Standard Python and Node.js development environments.

USE CASES AND FUNCTIONALITY DEMONSTRATION

Common Public Legal Queries in English

A primary function of the chatbot is to provide clear and concise answers to common legal questions. Consider a user, Ananya, who is unfamiliar with the process of filing a consumer complaint. She approaches the chatbot with the query: "How can I file a consumer complaint in India?" The JavaScript frontend captures this text input and sends it to the Flask backend. The backend, recognizing the language as English, directly constructs a prompt for the LLaMA 3.1 model via Ollama. The system prompt ensures the AI focuses on Indian law and provides actionable information. The LLM processes the query and generates a response, which is streamed back to Ananya's interface.

The chatbot might respond: "To file a consumer complaint in India, you can approach the District Consumer Disputes Redressal Forum for claims up to INR 1 crore, the State Commission for claims between INR 1 crore and INR 10 crore, or the National Commission for claims exceeding INR 10 crore. You can file the complaint yourself or through an authorized representative. The complaint should be in writing and include details of the dispute, the relief sought, and supporting documents. You can also file complaints online through the E-DAAKHIL portal (https://edaakhil.nic.in/). For more detailed procedures, you can refer to the Consumer Protection Act, 2019." This response is informative, provides a relevant government portal link (a form of citation/resource), and mentions the governing act, aligning with the chatbot's design to offer helpful resources and basic citation support.

Document Summarization: First Information Report (FIR)

Another critical capability is the summarization of legal documents. Imagine a user, Rohan, who has a copy of a First Information Report (FIR) in PDF format and finds its legal language difficult to understand. He uses the document upload feature of the chatbot. The frontend allows him to select the PDF file, which is then sent to the /upload_document endpoint of the Flask backend. The backend saves the file, and the process_document function extracts the text content. A specialized prompt is then sent to LLaMA 3.1, instructing it to summarize the FIR, focusing on key alleged offenses, involved parties, dates, and any immediate legal implications under Indian law.

The chatbot's streamed summary might appear as: "This First Information Report (FIR) summary indicates the following: The complaint was lodged by [Complainant's Name] on [Date of FIR] at [Police Station Name]. It alleges [brief description of main allegations, e.g., 'theft of property under Section 379 of the Indian Penal Code']. The accused parties mentioned are [Accused Names, if any]. The incident reportedly occurred on [Date of Incident] at [Location of Incident]. Key sections of law invoked appear to be [e.g., IPC Section 379, 420]. Please note, this summary is for informational purposes and is not legal advice. It is recommended to consult with a legal professional for guidance on this matter."

Multilingual Interaction: Query in Hindi

The chatbot is designed to support multilingual interactions, with Hindi being a primary supported language alongside English. Consider a user, Priya, who is more comfortable communicating in Hindi. She selects Hindi as her preferred language through the frontend interface. She then types her query in Hindi: "Where can I get legal advice for a land dispute?" The JavaScript frontend sends this query to the /chat endpoint, with the Accept-Language header set to hi. The Flask backend detects the Hindi language. The TranslationService translates Priya's query into English: "Where can I get legal advice for a land dispute?" This translated query is then used to construct the prompt for LLaMA 3.1. The LLM generates a response in English, which is then translated back to Hindi before being sent to Priya.

CHALLENGES AND LIMITATIONS

Technical Challenges

The core of the chatbot relies on the LLaMA 3.1 model, and while powerful, Large Language Models (LLMs) are not infallible. One primary technical challenge is ensuring the accuracy and reliability of the information provided. LLMs can occasionally generate plausible but incorrect or misleading information, often termed "hallucinations." In the legal domain, where precision is paramount, such inaccuracies could have serious consequences. The dynamic nature of law, with new legislation, amendments, and judicial precedents constantly emerging, necessitates continuous updates or fine-tuning of the underlying model or its knowledge base, which presents an ongoing maintenance challenge.

Furthermore, legal language is characterized by its nuance and context-dependency. LLMs may struggle to capture these subtleties perfectly, potentially leading to misinterpretations or oversimplifications of complex legal concepts. This is particularly relevant when summarizing documents like FIRs or petitions, where the omission or misrepresentation of a single detail can alter the perceived meaning significantly.

Real-time performance and scalability also pose technical hurdles. While the use of streaming responses improves perceived speed, complex queries or the processing of large documents can still introduce latency, especially if the local Ollama instance faces resource constraints. As a system intended for a Department of Justice website, it must be prepared for a potentially high volume of concurrent users.

Data Security and Privacy

The chatbot, by its nature, may handle sensitive personal information, whether through user queries detailing personal legal situations or through uploaded documents like FIRs or legal petitions. Ensuring robust data security and privacy is a critical ethical and technical imperative. Data must be protected during transit (between frontend and backend, and backend and Ollama), during processing, and in storage, even if temporary (e.g., in the UPLOAD_FOLDER). The current implementation details in app.py show basic file handling, but a production system would require comprehensive security measures, including encryption, access controls, and secure deletion policies, to comply with data protection regulations such as India's Digital Personal Data Protection Act.

Ethical Considerations

Beyond data privacy, several ethical considerations arise. Bias in LLMs is a well- documented concern. LLaMA 3.1, like other LLMs, is trained on vast datasets that may reflect existing societal biases. These biases could inadvertently surface in the chatbot's responses, potentially leading to unfair, discriminatory, or incomplete information, especially concerning vulnerable groups or sensitive legal issues.

The most significant ethical challenge is avoiding the provision of legal advice and managing user expectations. The system is explicitly designed as an informational tool, and system prompts instruct the AI to refrain from giving legal advice. Disclaimers are also included. However, users, especially those in distress or unfamiliar with legal distinctions, might still interpret the chatbot's outputs as definitive legal counsel. The line between providing comprehensive legal information and offering actionable legal advice can be very fine.

IV. DISCUSSION AND IMPACT

Increased Access to Legal Information

• 24/7 Availability: Unlike traditional legal services, the chatbot can provide information and assistance at any time, making legal help more accessible, especially for those with busy schedules or in different time zones. • Reduced Cost Barriers: Accessing legal information through this chatbot is free, removing financial obstacles that often prevent individuals from seeking legal guidance. • Simplified Language: The chatbot can explain complex legal jargon in simpler terms, making legal information more understandable for laypersons. • Multilingual Support: By incorporating translation capabilities, the chatbot can reach a wider audience, including those who are not proficient in the primary language of the legal system.

Empowerment of Individuals

• Understanding Rights and Procedures: The chatbot can educate users about their legal rights and the proper procedures to follow in various legal situations, empowering them to navigate the legal system more effectively. • Initial Guidance and Direction: While not a substitute for professional legal advice, the chatbot can provide initial guidance, helping users understand if they have a case, what steps to take next, and where to seek further assistance. • Increased Confidence: By demystifying legal processes and providing accessible information, the chatbot can reduce the anxiety and intimidation often associated with legal matters, giving individuals more confidence to address their legal needs.

Support for Legal Professionals

• Handling Routine Queries: The chatbot can handle a large volume of common and repetitive legal questions, freeing up the time of legal professionals to focus on more complex cases and tasks requiring human expertise. • Initial Case Assessment: The chatbot can assist in gathering initial information from clients, potentially streamlining the intake process for legal aid organizations or law firms.

V. CONCLUSION

The development of this AI-powered legal chatbot represents a significant step towards enhancing public access to legal information and services. By leveraging the power of large language models like LLaMA 3.1 and integrating features such as document summarization and multilingual support, the chatbot offers a user-friendly and accessible platform for individuals seeking legal guidance. The system's architecture, built upon a Flask backend and a dynamic frontend, demonstrates a practical approach to deploying AI solutions for real-world challenges. Throughout the development process, we have explored various aspects of the system, from the initial conceptualization and design to the implementation of core functionalities. The use of a modular design has allowed for flexibility and scalability, while the integration of advanced AI capabilities has enabled the chatbot to provide nuanced and contextually relevant responses. The inclusion of features like document summarization and multilingual support further enhances its utility, making it a valuable tool for a diverse range of users.

However, the Paper also highlights several challenges and limitations that need to be addressed in future iterations. These include the ongoing need to improve the accuracy and reliability of the AI model, particularly in handling complex legal queries and avoiding biases. Ensuring the privacy and security of user data is paramount, and further measures may be required to strengthen these aspects of the system. Additionally, expanding the knowledge base of the chatbot to cover a wider range of legal topics and jurisdictions would significantly enhance its value. Looking ahead, there are several promising avenues for future work. One key area is the integration of more sophisticated AI techniques, such as natural language understanding and generation, to improve the chatbot's ability to comprehend and respond to user queries in a more human-like manner. Another important direction is to explore the use of machine learning to personalize the user experience, tailoring the information and guidance provided to the specific needs and context of each user. Furthermore, expanding the chatbot's multilingual capabilities to include more languages would significantly broaden its reach and impact. The AI-powered legal chatbot presented in this paper demonstrates the potential of technology to transform access to legal information and services. While there are challenges to overcome, the ongoing advancements in AI and natural language processing offer exciting possibilities for the future development of such systems. By continuing to innovate and refine these tools, we can empower individuals with the knowledge and resources they need to navigate the complexities of the legal system and uphold their rights.

REFERENCES

1. Kumar, A., Mehta, S., & Singh, R. (2023). Large language models in legal applications: Opportunities and challenges. *Journal of AI and Law, 12*(3), 245–260.
2. Patel, V., & Sharma, D. (2024). Chatbots for public services in India: A case study approach. *International Journal of E-Governance, 8*(2), 112–128.
3. Meta AI Research. (2023). *LLaMA 3: An open foundation language model* (Meta AI Technical Report).
4. Ollama.ai. (2023). *Ollama: Run open large language models locally*. https://ollama.ai/
5. Gupta, P., & Jain, R. (2021). Natural language processing for legal document analysis: A survey. *Artificial Intelligence and Law, 29*, 171–196.

6. Venkatesh, S., Singh, A., & Basu, N. (2023). Multilingual legal translation systems: Challenges and solutions. In *Proceedings of the International Conference on Legal Technology* (pp. 45–52).
7. Sharma, M., Verma, K., & Reddy, L. (2022). Automatic summarization of legal documents: Progress and open challenges. *IEEE Access, 9*, 54321–54335.
8. Grinberg, M. (2023). *Flask web development: Developing web applications with Python* (2nd ed.). O'Reilly Media.
9. Kumar, R. (2023). Data protection laws in India: Current framework and future directions. *Journal of Indian Law, 15*(4), 312–328.
10. Kumar, S. S., Ahmed, S. T., Flora, P. M., Hemanth, L. S., Aishwarya, J., GopalNaik, R., & Fathima, A. (2021). An improved approach of unstructured text document classification using predetermined text model and probability technique. In *ICASISET 2020: Proceedings of the First International Conference on Advanced Scientific Innovation in Science, Engineering and Technology* (p. 378).
11. Nair, S., & Mittal, P. (2024). Technology for legal accessibility in developing nations. *Access to Justice Review, 6*(2), 78–94.
12. Desai, R., & Kulkarni, A. (2023). Ethical implications of AI-powered legal advisors. *Journal of Legal Ethics and Technology, 7*(1), 33–47.
13. Banerjee, T., Iyengar, S., & Khosla, M. (2022). Judicial decision prediction using transformer models. In *Proceedings of the International Conference on AI & Law* (pp. 98–105).
14. LK, S. S., Rana, M., & Ahmed, S. T. (2021, November). Real-time iot based temperature and npk monitoring system sugarcane-crop yield for increasing. In *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)* (pp. 1-5). IEEE.
15. Fathima, A. S., Prakesh, D., & Kumari, S. (2022). Defined Circle Friend Recommendation Policy for Growing Social Media. *International Journal of Human Computations & Intelligence*, *1*(1), 9-12.
16. Suresh, B., & Iyer, G. (2023). LegalBot: A rule-based and NLP-enhanced chatbot for civil cases. *Indian Journal of Computer Applications in Law, 4*(2), 120–135.
17. European Commission. (2023). *AI and the rule of law: Use of AI in legal systems*. https://ec.europa.eu/digital-strategy/
18. Zhang, Y., Wu, L., & Chen, K. (2024). Zero-shot legal text classification with foundation models. *arXiv preprint arXiv:2401.01234*. https://arxiv.org/abs/2401.01234