Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

# AI Legal Assistant for Indian Laws

**A.yaswanthi sree divya[1], B. Madhuri[2], V. Manasa[3], A. Bharath Nagendra[4], S. Yaswanth Sai[5], A.V.Srinivasa rao[6]**

Department of Electronics and Communication Engineering
Sasi Institute of Technology and Engineering, Tadepalligudem,AP,India

**Abstract**

Legal information in India is difficult to access due to the complex language used, limited knowledge, and the absence of professional legal services. In this regard, this paper provides the suggestion to create an AI-Based Legal Assistant for Indian Laws that utilizes Natural Language Processing (NLP) and Large Language Models (LLMs) to provide the us- ers with understandable, accurate, and updated legal information. The developed AI Legal Assistant will allow users to formulate their requests in natural English and will be able to retrieve legal information from datasets in structured format, such as the Constitution of India and Indian Penal Code (IPC). It is worth noting that unlike existing rule- and keyword-based systems, the proposed system utilizes intelli- gent query processing and contextual understanding in order to improve response accuracy. At the same time, the implementation will take advantage of advanced AI and web development technologies in or- der to guarantee future growth and efficient operation of the system. It will substantially increase ac- cess to the law, reduce the necessity to consult with legal experts on trivial issues, and raise awareness of one's legal rights.

**Keywords:** AI,NLP,LLM,IPC,IRPC

## 1. Introduction

Indian legal system is made up of a host of laws, reg- ulations and judicial rulings. It may be difficult for common people to obtain precise legal information since they will need some assistance from legal professionals to comprehend the law. Little is known about the rights and obligations related to the laws because of the complicated legal terms.

It is very natural that individuals tend to employ law- yers to help them to resolve their legal problems, however, it is a very time-consuming and expensive process. Due to the pace of AI technology development, now more than ever before it is possible to cre- ate intelligent systems that are able to simplify complex legal information to be understood by all.This research paper outlines the development of an AI-based Legal Assistant application, which lever- ages the NLP and LLM techniques in providing in- stant legal advice. Users can ask their questions us- ing natural language, and get a clear legal answer to their question. The main contribution of this work is the creation of a smart system that will help to access the legal content.

## 2. Literature Survey

The sphere of law has also experienced a rapid evo- lution of AI, which led to the creation of numerous intelligent systems that can help users with their le- gal questions, find documents, and provide an intelligent legal answer. The application of NLP tech- niques, ma-chine learning, and deep learning methods to increase the usefulness of legal chatbots and decision support systems have been the subjects of many research works. Recent literature is primarily based on the themes of retrieved augmented generation (RAG), semantic searching using the support of the vector databases, and integration of the large language models (LLMs).

Application of AI to legal assistance has been ob- served to be effective in the re-cent past owing to the natural language processing technique applied. This is since Sharma and Iyer (2024) state that there has been the development of legal assistants whose role is to respond to legal questions posed by users based on the employed datasets prior to the development of the assistant. Although the assistant was good at providing prompt responses, it was unable to inter- pret the context that the questions were posed because of small datasets.

Research has also been carried out regarding NLP's use in retrieving legal information. Verma & Rao (2022) developed a system that was able to extract relevant laws from large chunks of data, but the system struggled to understand the complex legal ter- minology and multi-context queries. The chatbot created by Mehta et al. (2021) to provide legal ad- vice services was not intelligent in any way.

The previous methods that included the use of rule- based expert systems were able to demonstrate structural capabilities of reasoning. For example, Nair (2019) was able to automate a legal expert system by use of pre-existing rules; however, this kind of system had low adaptive capabilities. The online legal solutions offered by Reddy (2017) provided just a portion of the legal information, though artificial intelligence was not included.

The emergence of RAG models has greatly helped to improve the efficiency of legal AI. The combination of both the use of the vector retrieval and gen- erative models has been proved to be of great advantage in improving the performance of the system. However, this technology involves complicated systems and not user-friendly interfaces, rendering it unavailable to a wider audience.

The other area that can be researched on is the as- sessment of the responses that the AI produces. This has been quantified on measures like BLEU and ROUGE to measure the quality of the response and modern research has been more concerned about the legal consistency and factual accuracy of the response. Additionally, research that concentrates on AI ethics uncovers the problems related to bias, misinformation, and obscurity.

Despite remarkable gains made, there are still nu- merous drawbacks in the current systems. Current systems lack domain-specific customizations in line with the Indian law system, lack the ability to process complex queries related to the law, and lack user-friendly interfaces to users who may not be familiar with the topic. Moreover, there are many other restrictions that im-pact the performance of these systems.

## 3. Methodology

### A. Data Collection and Preprocessing

The efficacy of the AI-powered legal assistant de- pends heavily on the quality and structure of the legal database. In our study, legal data will be sourced from reliable and genuine sources of Indian leg-al law, which include but are not limited to the Consti- tution of India, Indian Penal Code (IPC), Criminal Procedure Code (CrPC), Civil Procedure Code (CPC), and other legally valid information available publicly through government websites and legal documents.

Preprocessing is carried out on the gathered data to make it more conducive for intelligent queries processing. First of all, text normalizing processes like tokenizing, stop-word elimination, and lemmatization are used to normalize the legal documents with- out losing relevant terminologies within them. Furthermore, entities extraction is also done for extract- ing legal aspects such as acts, sections, etc.

In order to optimize system efficiency, unnecessary and redundant information is filtered out. Once the cleaning process is completed, the information will be systematically structured so as to facilitate easy retrieval of the required legal information. In con- trast to the existing systems which solely depend on keyword searching techniques, the suggested ap- proach takes a different dimension of understanding users queries.

### B. Model Design and Query Processing

The proposed model uses an artificial intelligence (AI) based method for processing queries submitted by the users and providing relevant legal answers. Once the query is received from the user, NLP is performed on the query to comprehend its content.

The query is then sent to a Large Language Model (LLM) to be interpreted. The model breaks down the query context and generates an output, which is hu- man understandable. As opposed to rule-based systems, the LLM does not need pre-determined rules but responds to the query and existing legal data through dynamic generation of response.

This enables it to be more adaptable than traditional rule-based systems and enables the assistant to address legal issues that are both relatively easy and fairly difficult.

### C. System Architecture and Framework

The whole system is built on a modular architecture comprising frontend, backend, and database modules with AI capabilities.

The front-end is developed with the help of HTML, CSS, and JavaScript, therefore, offering an interactive system where the questions can be posed about the law. On the other hand, the back-end has been made with the help of Node.js and Express API.

Implemented on MongoDB, the database contains legal data, user queries, and system re-responses. This will provide effective data management and en- able the system to record query logs to enhance in the future.

The system works in the following way:

1. The interface gets a query of law entered by the user. The backend processes the query using NLP techniques.
2. The AI model analyzes the query and re- trieves relevant legal information.
3. A response is generated and displayed to the user in simplified language.

The system is deployed using platforms such as GitHub and cloud services to ensure scalability and accessibility.

### D. Response Generation and Accuracy Enhance- ment

In order to provide reliable and relevant responses, the technology is enhanced by sophisticated prompt engineering and dynamic response parsing capabili- ties. Rather than generating responses in an open-ended manner, the artificial intelligence model is re- stricted within a tightly structured five-part Markdown format through the use of system prompts: **Case Summary, Explanation, Risky Clauses, Important Terms, and Safe Clauses**.

Furthermore, accuracy and relevance are improved by the following means:

- **Adaptive Persona:** The system permits us- ers to choose particular interaction styles, such as a simple explanation, legal expert, or action-oriented mode. The back end in- tegrates this setting into the AI's context window, making the language more or less sophisticated based on the user's understanding.
- **Risk Scoring Automation:** To convert le- gal analysis into immediate action, the frontend contains a custom parser engine (LegalResponseRenderer) that parses the Markdown response from the AI. The Risk Dashboard is automated scoring based on the quantity of bullet points that are present in the Risky Clauses and Safe Clauses sec- tions, which gives a Safety Score percent- age.
- **Domain Restriction:** To avoid hallucina- tions and remain focused, the AI is directly told through system prompts that it should categorize the request. In case the user en- ters an illegal request, the system activates a pre-set refusal procedure, making sure that the assistant only works in the Indian legal domain.

## 4. Implementation and System Architecture

The AI Legal Assistant suggested above utilizes a scalable client-server architecture that employs a MERN stack (MongoDB, Express.js, React.js, Node.js) and incorporates sophisticated multimodal language models for processing legal document and natural language queries.

1. **AI Processing at Core Level and Multi- modal Integrations:**

The orchestration of models takes place at the core level using Ollama. For maximum efficiency and effectiveness, the backend dynamically directs queries to the appropri- ate large language models depending on the type of inputs. For standard text and docu- ment queries, the high-compute text model (gpt-oss:120b-cloud) is used. Queries that involve image uploads are seamlessly di- rected to the vision-language model (qwen3-vl:235b-cloud).

2. **Documentation Ingestion Pipeline:** While most of the text-only chatbots rely on an inefficient approach, our AI has its own multi-modal documentation ingestion pipeline.

- Files are uploaded through multer.
- Legal documentation in image format is first optimized and compressed with the help of the sharp library before being ana- lyzed by AI.
- Txt files are processed by specific libraries, pdf-parse that can extract text from PDF files (automatically set limits to avoid to- ken overflows), and mammoth that can ex- tract text from .docx files.

3. **Backend Architecture & Security**

The backend is based on Node.js and the Ex- press.js library and works as an intermediary component between the client, the database, and the APIs of the large language models (LLM). Authentication and security are handled via JSON Web Tokens and bcryptjs hashing. The system has a Role Based Access Control mech- anism, where regular users differ from adminis- trators who can access a special dashboard page called AdminDashboard.jsx.

4. **Frontend Interface and Database**

The front-end interface is constructed as a Single Page Application (SPA). It uses a modern and re- sponsive user interface that has a design that sup- ports CSS glassmorphism and dual themes (light/dark). The chat application makes use of re- act-markdown to display the structured and com- plex outputs from the AI.

MongoDB is used as the persistent storage system. The schema has been created to facilitate management of user profiles, create unique session ID’s for the legal consultation, and manage chat history.

## 5. Results and Evaluation

The AI Legal Assistant was assessed by evaluating its ability to handle multimodal inputs, maintain structure, and quantify legal risks through several scenarios of real-life legal cases.

### E. Multimodal Inputs Handling and Dynamic Route Selection

The system proved robust when handling unstruc- tured inputs from legal processes. The user uploaded two types of inputs - image files and TXT files. When the backend detected that the input contained an image, it routed the request to the Vision Lan- guage model(qwen3-vl:235b-cloud) rather than the normal text model(gpt-oss:120b-cloud). Prior to be- ing routed to the AI, the image was resized and com- pressed using sharp library to boost performance. On the other hand, the plain text input was extracted as is from the TXT files using file system calls. This means the user could upload a complaint or FIR as either an image file or a TXT file without breaking the application.

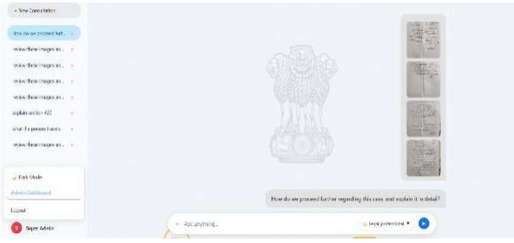

Fig. 1: User Input Interface with Uploaded Le- gal Documents

### F. Structured Legal Analysis Generation

In contrast to conventional conversational models that generate random responses, the model followed the system prompt guidelines flawlessly. Upon be- ing shown legal texts, such as complaints according to the Bharatiya Nyaya Sanhita or civil agreements, the AI generated the response strictly within the re- quired five-part structure:

**Case Summary:** Giving a quick overview of the strength of the case and the main le- gal issue involved.

- **Explanation:** Converting complex legal terminology and procedures (for example, registration of FIR or preparation of charge sheet) to the user's selected degree of un- derstandability, on a spectrum between a plain explanation to a legal-professional dissection.

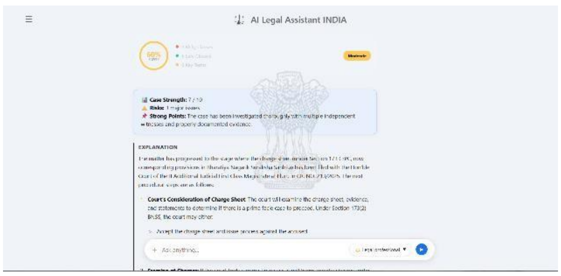


Fig 2: Clear explanation and analysis of Case

- **Risky Clauses:** Spotting issues related to procedural delays, absence of evidence, and risky contract provisions.


Fig 3: Identification of Risky and Important Terms regarding the case

- **Important Terms:** Outlining important re- sponsibilities, deadlines, and applicable laws and statutes.
- **Safe Clauses:** Noticing relevant legal pro- tection provisions and compliance require- ments.


Fig 4: Identification of safe clauses regard- ing the case

### G. Quantitative Risk Assessment (The Risk Dash- board)

The most prominent feature revealed by the evalua- tion of the system was its ability to convert qualita- tive AI content into quantitative visuals. In particu- lar, the LegalResponsibleRenderer was able to read the Markdown content generated by the AI in real- time. Using the ratio between bullet points under Safe Clauses and Risky Clauses sections, it could render an interactive Risk Dashboard. The dash- board would then automatically present the safety score percentage having a suitable color like red (high risk) and green (safe).

### H. System Performance and Latency

The experimentally obtained results confirm the ef- ficiency of using MERN technology for implement- ing systems with high computational complexity of LLMs. Specifically, for the frontend, a function for displaying a dynamic typing effect was imple- mented, which provided visual confirmation to the user that the AI analysis was in progress. More-over, JWT-based authentication and RBAC mechanisms used in the system effectively separated the user ses- sions, ensuring that personal data during consulta- tion remained confidential in the MongoDB data- base.

## 6. Conclusion

The process of designing and analyzing the AI Legal Assistant for Indian laws is an excellent example of a perfect blend of cutting-edge web technology with LLMs. Using the MERN stack architecture and mul- timodal AI routing (Ollama, GPT-OSS, Qwen3- VL), the tool has successfully managed to bridge the gap between the intricacies of the legal structure and the common citizen.

The use of prompt engineering makes sure that there is a reduction of hallucinations by AI and consistent structuring of response messages into actionable cat- egories. Also, the introduction of a quantitative tool called Risk Dashboard helps make complex legal analysis more tangible and visual. Overall, the sys- tem shows how effective it is to use AI in processing images and texts in law through legally structured assistance provided to Indians without any knowledge of the law.

**References**

1. Liu, B. Feng, B. Xue, B. Wang, B. Wu, C.Lu, C. Zhao, C. Deng, C. Zhang, and C. Ruan, DeepSeek-V3 Technical Report, arXiv preprint arXiv:2412.19437, 2024.
2. M. Douze, A. Guzhva, C. Deng, J. Johnson, G. Szilvasy, P.-E. Mazaré, M. Lomeli, L. Hosseini, and A.Jégou, The FAISS Library, arXiv preprint arXiv:2401.08281, 2024.
3. D. Khurana, A. Koli, K. Khatter, and S. Singh, Natural Language Processing: State of the Art, Cur- rent Trends and Challenges, Multimedia Tools and Applications, vol. 82, no.3, pp. 3713-3744, 2023.
4. Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi,Y. Dai, J. Sun, H. Wang, and H. Wang, Retrieval- Augmented Generation for Large Language Models: A Survey, arXiv preprint arXiv:2312.10997, 2023.
5. S. S. Manathunga and Y. Illangasekara, Re- trieval-Augmented Generation and Representative Vector Summarization for Large Unstructured Tex- tualData, arXiv preprint arXiv:2308.00479, 2023.
6. Ryu et al.,Retrieval-Based Evaluation for LLMs: A Case Study in Korean Legal QA, in Proc. Natural Legal Language Processing Workshop, Sin- gapore, 2023, pp. 132-137.
7. A.Devlin, M.-W. Chang, K. Lee, and K. Toutanova, BERT: Pre-training of Deep Bidirec- tional Transformers for Language Understanding, in Proc. NAACL-HLT, Minneapolis, USA, 2019, pp. 4171-4186.
8. Danopoulos, C. Kachris, and D. Soudris, Ap- proximate similarity search with FAISS framework using FPGAs on the cloud, 2019, pp. 373–386.
9. Y. Liu et al., RoBERTa: A robustly optimized BERT pre-training approach, arXiv preprint arXiv:1907.11692, 2019.
10. H. Arabi and M. Akbari, Improving plagiarism detection using hybrid weighted similarity, Expert Systems with Applications, vol. 207, 2022.
11. A.Papineni, S. Roukos, T. Ward, and W.-J. Zhu, BLEU: A Method for Automatic Evaluation of Machine Translation, in Proc. ACL, 2002, pp. 311– 318.
12. C.-Y. Lin, ROUGE: A Package for Automatic Evaluation of Summaries, in Proc. ACL Workshop, 2004, pp. 74–81.
13. J. Yuan et al., RAG-Driver: Generalisable Driv- ing Explanations with Retrieval-Augmented Learn- ing, arXiv preprint arXiv:2402.10828, 2024.
14. J. Xue et al., Bias and Fairness in Chatbots: An Overview, APSIPA Transactions on Signal and In- formation Processing, vol. 13.
15. A.Min, Artificial Intelligence and Bias: Chal- lenges, Implications, and Remedies,2023.
16. H. Zeng et al., Scalable and Effective Genera- tive Information Retrieval, in Proc. ACM Web Con- ference, 2024, pp. 1441–1452.
17. Singh, Legal Documents Text Analysis Us- ing NLP, in Proc. ICSSAS, 2024, pp. 1302–1307.
18. P. Weichbroth, AI and the Law: Evaluating ChatGPT's Performance in Legal Classification,

2025.
19. S. Yu et al., Disparity-Preserved Deep Cross- Platform Association, arXiv preprint arXiv:1901.00171, 2019
20. Schwartz, Ensemble of MRR and NDCG Models for Visual Dialog, arXiv preprint arXiv:2104.07511, 2021.
21. Y. Wang et al., Equality Before the Law: Legal Judgment Consistency Analysis, arXiv preprint arXiv:2103.13868, 2021.
22. S. Es et al., RAGAs: Automated Evaluation of Retrieval-Augmented Generation, in Proc. EACL, 2024, pp. 150-158.