IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, November 2026
Impact Factor: 8.2

Law Assist AI: Smart AI Chatbot for Legal Guidance

Trushal Hiremath[1], K. V. Anish[2], Chandan M. A.[3], Prof. Lakshmi Sharma K. M.[4]
[1],[2],[3] Students, Department of Computer Science – Artificial Intelligence and Machine Learning
[4]Project Guide, Department of Computer Science – Artificial Intelligence and Machine Learning
Proudhadevaraya Institute of Technology, Hosapete, Karnataka, India

Abstract : *Looking up Indian statutes or court steps is difficult without a lawyer. Most people do not know which section applies to a dispute, and government portals only host raw PDFs. Here we lay out the Phase 1 design of Law Assist AI, a chatbot built around retrieval-augmented generation (RAG) and natural language processing. Instead of letting a language model recite laws from memory, every reply pulls context from an indexed store of Indian acts—including IPC, BNS, CrPC, BNSS, and court rulings. Users type or speak in English or regional languages and toggle between formal statutory text and an "Explain Like I'm 10" summary. The system also suggests First Information Report (FIR) sections, displays risk and confidence scores, checks uploaded notices, and drafts basic templates. Because this manuscript covers Phase 1 design, it defines the software stack, routing logic, and test plan without reporting unmeasured accuracy numbers.*

Keywords: legal chatbot, retrieval-augmented generation, Indian law, multilingual voice UI, explainable AI

I. INTRODUCTION

Most citizens only read Indian law when a dispute, consumer complaint, or police matter lands on their doorstep. Bare acts are dense and hard to parse without legal training. Paying an advocate just to understand basic rights or filing steps costs time and money many households cannot spare.

Public websites do publish statutes and forms, yet a search fails unless the user already knows the exact act name. Generic AI chatbots can hold a conversation, though they carry a serious flaw: they invent section numbers when forced to answer from parametric memory. Citing a repealed clause in a legal dispute can push someone toward the wrong action.

Law Assist AI separates language fluency from legal facts. Our Phase 1 architecture routes typed or spoken queries—including regional Indian languages—into a vector index of Indian statutes and judgments. Replies appear in standard statutory form alongside an "Explain Like I'm 10" view, backed by FIR hints, risk badges, and practical checklists. Live user trials are planned for Phase 2.

II. RELATED WORK

Earlier studies on legal chatbots point to two core needs: regional language support and factual grounding. Rajeshwar et al. built a multilingual bot for Indian law and found that non-English support widens reach outside metro cities [1]. To stop models from inventing rules, Mehta et al. tied language models to external legal stores using RAG [2].

Gupta et al. later used LangChain to organize statutory document retrieval across multi-step queries [3]. Benchmarks by Iyer et al. showed that RAG beats standalone models on citation precision, though response latency still needs tuning [4]. On the ethics side, Fernandez et al. recommended Explainable AI (XAI) indicators like confidence scores so users know how much to trust an automated answer [5]. Sharma et al. documented how messy, multilingual Indian legal corpora must be cleaned before indexing [6]. Table I lists how these six papers guide our architecture.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, November 2026
Impact Factor: 8.2

TABLE I: SUMMARY OF RELATED WORK

Ref.	Methodology	Primary Focus	Contribution to Current Design
[1]	Multilingual Indian law bot	Non-English user access	Regional text & voice support
[2]	RAG over legal databases	Cutting invented citations	Statute-grounded answers
[3]	LangChain query pipelines	Structuring legal documents	Modular retrieval workflow
[4]	Standalone vs. RAG tests	Measuring answer precision	Mandatory context check
[5]	Explainable AI (XAI) cues	Bias & trust transparency	Confidence & risk badges
[6]	Indian legal NLP survey	Cleaning raw statutory text	Corpus indexing strategy

III. RESEARCH GAP AND PROBLEM STATEMENT

Current legal tools rarely solve lookup, plain-language explanation, and procedural guidance together. Static portals list acts without answering follow-up questions. Many prototype bots also miss regional Indian languages, voice input, and the recent shift from IPC/CrPC to BNS/BNSS. Rural and non-English speakers still lack a free tool that explains what a law means and what step comes next.

Problem Statement: Citizens need an affordable, voice- and text-capable legal assistant that handles regional Indian languages, fetches current statutory rules accurately, and pairs plain next steps with clear risk warnings.

IV. OBJECTIVES

Five engineering targets guide this Phase 1 work:

1) Build an NLP and RAG pipeline that links everyday citizen questions to matching Indian statutes.
2) Provide two output modes per reply: formal legal text and a plain “Explain Like I’m 10” summary.
3) Suggest relevant IPC/BNS sections and FIR filing steps based on the user’s situation.
4) Add regional Indian translation and voice input/output for non-English speakers.
5) Include risk badges, confidence scores, basic document checks, and template drafts, backed by a Phase 2 test plan.

V. PROPOSED SYSTEM ARCHITECTURE

Fig. 1 splits the system into two cooperating paths—a RAG and Knowledge Loop for statutory lookup, and an Interaction and Guidance Loop for language, simplification, and next steps—run by backend APIs.

TABLE I : ARCHITECTURAL OVERVIEW OF THE LAW ASSIST AI SYSTEM

RAG & Knowledge Loop	Interaction & Guidance Loop
Legal Corpus (IPC 1860, BNS 2023, CrPC, BNSS)	Speech-to-Text / IndicTrans2 Translation
Sentence Transformers Embeddings	Standard / “Explain Like I’m 10” Mode
Vector Database (FAISS / Pinecone)	FIR Suggestions & Risk / Confidence Scoring
Retrieval-Augmented LLM (LangChain)	Actionable Guidance & Legal Templates

Fig. 1 Architectural overview of the Law Assist AI System (coordinated by FastAPI, PostgreSQL, and Redis).

A. Legal Knowledge Base and RAG Retrieval

The corpus stores IPC 1860, BNS 2023, CrPC, BNSS 2023, CPC, Limitation Act, Transfer of Property Act, Constitution of India, Consumer Protection Act 2019, IT Act 2000, DPDP Act 2023, Hindu Marriage Act, and selected court judgments. Files are split into chunks, encoded with Sentence Transformers, and indexed in FAISS or Pinecone via

B. LangChain so every reply starts from retrieved text.

Multilingual Query and Voice Input Processing

Users send questions by typing or speaking on web or mobile devices. Speech APIs convert audio into text, and IndicTrans2 handles regional Indian languages such as Hindi, Tamil, and Bengali.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 7, November 2026 Impact Factor: 8.2

C. AI Legal Response and Simplified Explanation Engine

Matched clauses and the user question go to a retrieval-augmented LLM. Prompt templates return two views: a formal summary with section numbers, and an "Explain Like I'm 10" note written in simple words.

D. Document Analysis, Risk Scoring, and Actionable Guidance

For uploaded notices, the backend checks basic clauses and fills draft templates. Each reply carries a confidence score, a risk-level badge, FIR hints where criminal rules apply, and a short checklist of where to file and what records to bring.

VI SYSTEM WORKFLOW AND DECISION LOGIC

Each turn runs through a deterministic state function based on detected language, reading mode, and retrieval confidence. Let R_t be the output state at step t:

$$R_t = g(q_lang, m_mode, c_conf) \quad (1)$$

Here q_lang holds the normalized query and modality, m_mode picks standard prose or "Explain Like I'm 10" mode, and c_conf tracks vector similarity and risk tier. Safety overrides fluency: if c_conf < c_threshold, the bot holds back firm statutory claims, flags low confidence, and tells the user to consult a qualified advocate.

VIII. IMPLEMENTATION REQUIREMENTS

Table II lists the hardware, backend tools, and NLP libraries chosen for prototype assembly.

TABLE II: PLANNED HARDWARE AND SOFTWARE STACK

Component	Specification	Role in System
Dev. & Test Hardware	Computer/Laptop, Mobile Device	Coding, local tests, and mobile checks
Audio I/O Hardware	Microphone, Speakers/Headphones	Spoken query capture and audio output
Backend & API	Python, FastAPI, Flask/Django, JWT, Redis	API routing, login tokens, and caching
AI & RAG Stack	LangChain, Sentence Transformers, FAISS/Pinecone	Chunk embeddings, vector lookup, and LLM calls
NLP & Speech	NLTK, SpaCy, Speech Recognition, IndicTrans2	Text cleaning, STT, and Indian-language translation
Database	PostgreSQL / MongoDB / MySQL	Query history and user metadata
Frontend UI	React.js / Next.js (HTML, CSS, JS)	Web chat and document upload UI

VIII. EVALUATION PLAN

No live user benchmarks are reported in this Phase 1 blueprint. During Phase 2, the working prototype will be tested against the five metrics in Table III.

TABLE III: PHASE 2 VERIFICATION METRICS

Criterion	Measurement Goal	Evaluation Focus
Retrieval Relevance	Precision of fetched IPC/BNS & judgment chunks	Grounding accuracy
Query Understanding	Intent & entity match on citizen prompts	NLP reliability
Response Clarity	Readability in standard vs. "ELI10" views	Layperson comprehension
Multilingual & Voice Accuracy	Translation fidelity & STT error rate	Regional access quality
Risk & Document Output	Validity of risk flags, FIR hints & templates	Guidance usefulness

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, November 2026
Impact Factor: 8.2

IX. EXPECTED OUTCOMES AND DISCUSSION

Once built, Law Assist AI should give citizens a fast, low-cost way to understand legal rules and filing steps before visiting a court or police station. Keeping retrieval (FAISS/Pinecone with LangChain) separate from text generation cuts down invented citations while still allowing plain "Explain Like I'm 10" summaries.

Phase 2 trials will also check how confidence scores and risk badges shape user choices. Clear warnings help people tell routine informational questions apart from serious disputes that genuinely need a practicing lawyer.

X. LIMITATIONS AND FUTURE WORK

Because this paper stops at Phase 1 design, live latency, retrieval precision, and user comprehension remain unmeasured. Calling cloud speech, translation, and LLM endpoints assumes steady internet access, which is uneven in rural areas. Ongoing court rulings around the IPC-to-BNS transition also mean the vector index must be updated regularly.

Phase 2 work will assemble the FastAPI and React app, measure retrieval precision across Indian acts, widen regional voice support, and ask legal practitioners to review FIR hints and template outputs.

XI. CONCLUSION

We presented the Phase 1 design of Law Assist AI, a legal guidance chatbot combining Indian statutory RAG search with multilingual text and voice access. Through dual standard and "Explain Like I'm 10" views, confidence gating, risk badges, and basic document checks, the design aims to make everyday legal information approachable without giving up statutory accuracy.

ACKNOWLEDGMENT

The authors thank the project guide, Prof. Lakshmi Sharma K M, the Head of the Department, the Project Coordinator, and the Department of CS-AI&ML, Proudhadevaraya Institute of Technology, Hosapete, for their guidance and support.

REFERENCES

[1] S. Rajeshwar et al., "Developing an AI-Powered Legal Chatbot for Multilingual Indian Law Queries," Int. Conf. Artificial Intelligence and Law (ICAIL), pp. 180–185, 2024.

[2] A. Mehta et al., "Enhancing Legal AI Chatbots with Retrieval-Augmented Generation: Techniques and Applications," J. Legal Technology and Practice (JLTP), vol. 15, no. 2, pp. 102–110, 2024.

[3] R. Gupta et al., "LangChain and Large Language Models for Legal Knowledge Retrieval: A Case Study," AI Ethics and Law Review (AIELR), vol. 9, no. 1, pp. 45–55, 2025.

[4] M. Iyer et al., "Evaluating RAG-Based Legal AI Assistants: A Comparative Study," Int. J. Artificial Intelligence & Law (IJAIL), vol. 10, no. 3, pp. 75–85, 2025.

[5] T. Fernandez et al., "Towards Fair and Explainable AI in Legal Chatbots: Addressing Bias and Hallucinations," J. Law and Artificial Intelligence (JLAI), vol. 16, no. 2, pp. 125–135, 2026.

[6] N. Sharma et al., "A Survey of Indian Legal Corpus and NLP Techniques for Developing AI Chatbots," Indian J. Legal Informatics (IJLI), vol. 8, no. 4, pp. 205–214, 2026.