

Juris AI Pocket: A Mobile- Multilingual Legal Intelligence Architecture Integrating RAG, Generative Contract Synthesis, and Judicial Informatics

*Sunitha Kanipakam¹, N. Anuradha², M. Usharani³, Dr. T.Vineela⁴, Kaveri Shukla⁵, Dr, Sarada Meduri⁶

¹Asst. Professor, Dept of Law, Sri Padmavati Mahila Visvavidyalayam, Tirupati, India.

²Asst. Professor, Dept of Education, Sri Padmavati Mahila Visvavidyalayam, Tirupati, India.

³Professor of Comp Science, Sri Padmavati Mahila Visvavidyalayam, Tirupati, India.

⁴Dept of Law, Sri Padmavati Mahila Visvavidyalayam, Tirupati, India.

⁵Research Scholar, Dept of Law, Sri Padmavati Mahila Visvavidyalayam, Tirupati, India.

⁶Asst.Professor, Anantha Law College, Tirupati, India.

PMUSHA Project-2025, Principal Investigator, Dr Sunitha Kanipakam [Corresponding Author]

Abstract: JurisAIPocket (JurisMobile) is a mobile-first multilingual legal intelligence platform designed to enhance access to legal knowledge, contract processing, and judicial informatics in emerging jurisdictions. Unlike fragmented legal technology systems, JurisAIPocket integrates Retrieval-Augmented Generation (RAG)-based contract inspection, AI-powered agreement synthesis, multilingual legal chatbots (English and Telugu), privacy redaction using NLP-based Named Entity Recognition, notarization support tools, judicial data visualization, and AI-driven legal education modules within a single lightweight ecosystem. The system leverages Meta LLaMA large language models for reasoning and generative tasks, vector-based semantic retrieval for clause-level contract analysis, and structured public judicial datasets including Courts-in-India statistics, Fast Track Special Courts (FTSC) deployment data, and district-wise Tele-Law records. The proposed architecture further introduces AI-enabled gamified legal pedagogy modules such as moot court problem generation, legal quizzes, risk clause detection, and workflow builders to support experiential learning. Designed using a React + Vite Progressive Web App (PWA) architecture with mobile-first optimization, minimal bundle size, caching mechanisms, and Cloudflare-backed delivery, JurisAIPocket demonstrates how advanced legal AI systems can operate efficiently in constrained mobile environments without subscription barriers or intrusive monetization. The work contributes to Legal Informatics by presenting a unified, multilingual, free-access, AI-driven mobile legal ecosystem supporting practitioners, students, and the general public.

Keywords: Legal Informatics; Notary Support; Contract Analysis; RAG; ML Chatbot; Document Redaction; Judicial Analytics; Tele-Law; Legal Literacy

1. INTRODUCTION

Legal technology has transformed from basic document storage solutions into advanced systems which can comprehend and produce and interpret legal documents. The legal industry continues to operate through disjointed procedures which require specialized skills for activities such as contract drafting and notarization and legal research and judicial data analysis and public legal education. The situation creates accessibility difficulties which require extended time to resolve and leads to diminished operational effectiveness. Unlike conventional desktop-centric legal AI platforms, JurisAIPocket adopts a mobile-first design philosophy, enabling lightweight deployment, rapid response



times, offline-friendly caching, and accessibility in bandwidth-constrained environments. This approach addresses the technological divide in legal service accessibility across semi-urban and rural regions.

The development of large language models (LLMs) together with Retrieval-Augmented Generation (RAG) systems allows users to perform advanced computational reasoning on legal documents while making legal information available to everyone and providing automated insights to legal professionals. The ability to speak multiple languages becomes essential for Indian jurisdictions because they still lack proper access to regional language legal materials.

The research presents JurisAIPocket (JurisMobile), as a solution to fulfill the requirements of legal professionals through its unified platform which includes multilingual AI chatbots and automated contract analysis and generation and notary workflows and judicial informatics and privacy redaction and legal literacy modules. The paper presents the system components together with the datasets and methodologies used in the evaluation process and the academic context of its contribution.

1.1 Problem Statement

Despite advancements in AI-assisted legal tools, existing systems remain fragmented, subscription-driven, and predominantly desktop-centric. Multilingual accessibility, particularly for regional languages, remains underexplored. Furthermore, few architectures integrate contract intelligence, judicial analytics, notarization assistance, privacy redaction, and legal pedagogy within a unified mobile-first ecosystem deployable in bandwidth-constrained environments. This research addresses the need for an integrated, lightweight, multilingual legal AI framework accessible to both professionals and the public.

2. LITERATURE REVIEW

Legal Informatics has evolved from rule-based expert systems toward data-driven and large language model (LLM)-powered architectures. Early systems focused on structured rule encoding and deterministic reasoning, which limited adaptability to heterogeneous legal language. Modern LLM-based systems such as GPT-series models, LLaMA architectures, and PaLM have demonstrated improved capabilities in legal summarization, clause extraction, and contextual reasoning.

Recent advancements in Retrieval-Augmented Generation (RAG) have significantly improved factual grounding in generative outputs by combining vector-based semantic search with large language models. In legal domains, RAG architectures have been applied to contract review, statute interpretation, and case-law referencing to reduce hallucination risks.

Benchmarks such as LegalBench and ContractNLI have further highlighted the challenges of domain-specific reasoning in legal language models, particularly in long-document comprehension and logical consistency.

Within the Indian context, research has largely focused on judiciary digitalization initiatives including e-Courts, National Judicial Data Grid (NJDG), and Tele-Law programs. However, existing research primarily addresses data transparency and administrative digitization rather than integrated multilingual AI-driven legal ecosystems. Enterprise platforms such as Harvey AI and Lexis+ AI demonstrate the growing adoption of AI in legal drafting and research workflows. However, these systems are typically subscription-based and enterprise-oriented, with limited focus on multilingual regional accessibility or mobile-first deployment strategies.

There remains limited academic exploration of unified, multilingual, mobile-first legal AI architectures integrating contract inspection, judicial analytics, notarization support, privacy redaction, and legal pedagogy within a single deployable ecosystem. JurisAIPocket addresses this gap.

3. CONTRIBUTION

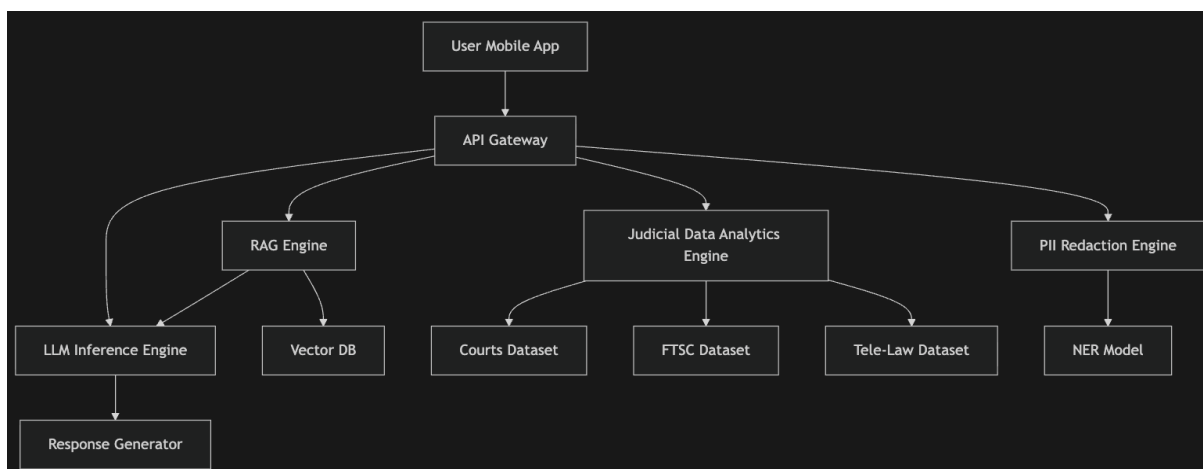
This work makes the following contributions to the domain of Legal Informatics:

1. Proposes a unified mobile-first legal AI architecture integrating RAG-based contract analysis, generative agreement synthesis, multilingual legal chatbots, and judicial data analytics.
2. Introduces AI-driven notarization assistance including a structured Andhra Pradesh Notary Officer Directory and automated notary document generation.
3. Develops a multilingual (English–Telugu) legal support chatbot leveraging Meta LLaMA models to improve accessibility in regional jurisdictions.
4. Integrates NLP-based privacy redaction mechanisms for automated masking of personally identifiable information (PII) in legal documents.
5. Presents a judicial informatics visualization module incorporating Courts-in-India, FTSC, and Tele-Law datasets for demographic-aware analysis.
6. Proposes an AI-powered legal gamification framework supporting experiential learning through moot court generators, quizzes, clause-risk detection, workflow builders, and interactive term-matching modules.
7. Demonstrates the feasibility of deploying large language model–based legal intelligence systems within lightweight Progressive Web App environments optimized for constrained mobile ecosystems.

4. SYSTEM OVERVIEW

JurisAIPocket is an integrated legal informatics system with the following core objectives:

1. Automate contract-centric tasks (analysis, generation, clause extraction).
2. Support multilingual legal interaction, especially in English and Telugu.
3. Provide notarization infrastructure, including officer databases and affidavit generators.
4. Enable document privacy and security via PII redaction.
5. Enhance legal literacy through structured educational modules.
6. Visualize judiciary-scale datasets for research and analysis.
7. Deliver AI-powered legal services through a lightweight mobile-first architecture optimized for constrained device environments.



The platform supports both legal professionals and non-expert end-users across research, education, and field practice domains.

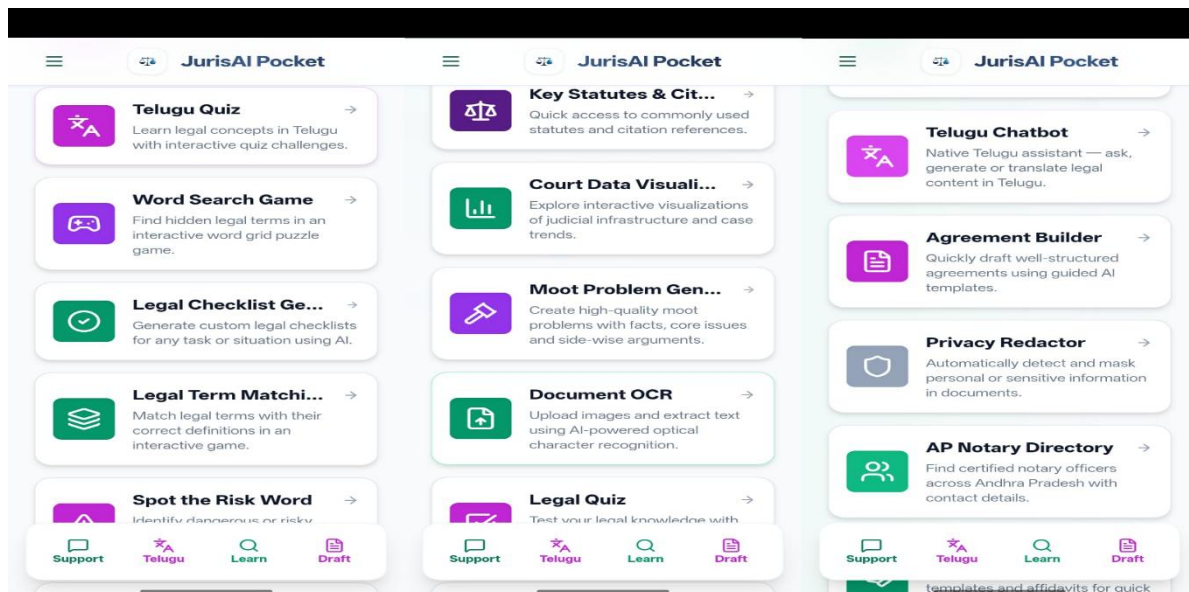


Figure. JurisAIPocket Mobile Application Screenshots

Module Name	Core Technology	Purpose	Output Type
Contract Inspector	RAG + Vector Embeddings + LLaMA	Clause-level contract analysis & Q&A	Structured Insights
Agreement Builder	Prompt Engineering + LLaMA	NDA, Lease, Employment, Custom Agreements	DOCX / Text
Legal Support Chatbot (English)	LLaMA Instruct	Legal Q&A and Explanation	Conversational Response
Telugu Legal Chatbot	LLaMA Multilingual	Regional Legal Assistance	Telugu Text Output
Privacy Redactor	NLP + NER Models	PII Detection & Masking	Redacted Document
AP Notary Directory	Structured Dataset	Notary Officer Lookup	Directory Listing
Notarization Template Generator	LLaMA Generative Model	Affidavits & Notary Documents	Printable Template
Key Statutes & Citations Repository	Structured Legal Database	Top 1000 Acts Reference	Act Summary
Court Data Visualizer	Data Analytics Engine	Courts & FTSC Statistics	Interactive Charts
Tele-Law District Analyzer	Judicial Dataset Processing	Demographic Case Data	Visual Dashboard
Moot Court Problem Generator	LLaMA Generative AI	Practice Case Scenarios	Case Document
Legal Quiz (English)	Rule Engine + AI	Knowledge Assessment	Score Output
Telugu Legal Quiz	Multilingual Engine	Vernacular Learning	Score Output
Word Search Game	Interactive Logic Engine	Legal Term Reinforcement	Game Grid

Legal Checklist Generator	AI Prompt-Based Generator	Task-Oriented Legal Checklists	Interactive Checklist
Legal Term Matching	Logic-Based Matching Engine	Concept Reinforcement	Score Feedback
Spot the Risk Word	Clause Pattern Detection + AI	Identify Risky Clauses	Risk Highlight
Legal Path Builder	Workflow Sequencing Engine	Legal Process Simulation	Ordered Steps

5. FUNCTIONAL MODULES

5.1 Contract Analysis Module (RAG)

This module enables users to upload contractual documents (PDF/DOCX/TXT). Using RAG pipelines, the system:

- Extracts and chunks clauses,
- Embeds segments into vector space,
- Retrieves relevant context for user queries,
- Generates structured insights, summaries, and red flags.

Outputs include clause breakdowns, compliance concerns, summary highlights, and semantic Q&A responses.

5.2 Contract Generation Module

The generator uses prompt-driven templates together with LLM inference to develop legally binding contracts. The system creates valid contract drafts in DOCX and PDF formats when users enter their requirements which include parties and jurisdictions and obligations and termination clauses. The module includes standard contractual agreements which include NDAs and rental contracts and service contracts.

5.3 Multilingual Legal Chatbot (English and Telugu)

The proposed architecture includes chatbots capable of:

- Explaining legal concepts,
- Translating legal terms,
- Answering user legal queries,
- Generating simplified legal explanations.

The Telugu component bridges accessibility gaps for vernacular users, supporting transliteration and contextual translation for region-specific usage.

5.4 Notary Assistance Module

This module serves notarization workflows via:

- AP Notary Officer Registry: A searchable directory of certified notary officers in Andhra Pradesh.
- Notarization Generator: An affidavit/document generator using LLaMA 3.1 70B instruct models to produce legally formatted notarizable affidavits and declarations.
- Process Literacy: Guides on notary requirements, jurisdiction rules, and signature workflows.

5.5 Redaction & Privacy Module

The PII redaction unit detects and masks sensitive textual information which includes names and addresses and financial identifiers and contact details. The purpose of this system is to allow document sharing between parties while maintaining privacy which is appropriate for both legal document submissions and third-party assessment processes.

5.6 Legal Knowledge Repository (Top 1000 Acts)

The proposed architecture contains a reference set of frequently cited Indian legal acts, supporting:

- Search-based lookup,
- Act summaries,
- Citation context,
- Link-outs to verdicts and legal commentary sources.

This repository enhances legal research productivity and facilitates educational usage.

5.7 Judicial Informatics and Data Analytics

The Informatics module displays judicial data through three established datasets which currently provide support to its operations.

The Courts-in-India dataset provides State and Union Territory court counts by year from 2014 through 2023.

The Fast Track Special Courts (FTSC) dataset provides information on active FTSC courts operating throughout various State and Union Territory locations from 2020 through 2023. The District-wise Tele-Law dataset tracks Tele-case registration according to gender and caste category and CSC center and district. The inclusion of demographically granular variables (gender and caste) aligns with ongoing legal inclusion research.

Feature	JurisAIPocket	Harvey AI	Lexis+ AI	ChatGPT
Free Access	Yes	No (Enterprise)	No (Subscription)	Limited
Mobile-First Design	Yes	No	No	Partial
Multilingual (English + Telugu)	Yes	Limited	No	Partial
RAG Contract Inspection	Yes	Yes	Yes	Limited
AI Agreement Builder	Yes	Yes	Yes	Yes
Notary Officer Directory (India)	Yes	No	No	No
Judicial Data Visualization	Yes	Limited	Limited	No
Legal Gamification Modules	Yes	No	No	No
Privacy Redaction	Yes	Limited	Yes	No
Ad-Free Platform	Yes	N/A	N/A	N/A

5.7.1. Datasets and Sources

6.8 5.7.1.1. *Courts in India Dataset:* This dataset contains yearly values (2014–2023) of judicial infrastructure units by State/UT. It enables longitudinal trends on institutional capacity.

6.9 5.7.1.2. *FTSC Functional Courts Dataset:* Captures the deployment of Fast-Track Special Courts, assisting stakeholders in analyzing regional legal reforms in tackling sexual offenses and related case backlogs.

6.10 5.7.1.3. *District-wise Tele-Law Cases Dataset:* Provides district-level attributes including CSC counts, gender distributions, caste category breakdowns (General/OBC/SC/ST), and total registrations. This dataset is valuable for legal inclusion studies.

6.11 5.7.1.4. *AP Notary Officer Dataset:* A registry of commissioned notary officers including jurisdictional metadata. Supports authentication workflows and notarization service navigation.

7. 8 Interactive Legal Micro-Modules: In addition to core AI modules, JurisAIPocket integrates lightweight interactive micro-modules designed for reinforcement learning and task simulation.

Legal Checklist Generator: Uses prompt-based AI synthesis to create customizable task-oriented legal checklists (e.g., company registration, tenancy process, filing procedures). Users can track completion through interactive checkboxes.

Legal Term Matching Module: Implements rule-based matching logic combined with structured definitions to reinforce conceptual clarity in legal terminology.

Spot the Risk Word: Uses clause-pattern recognition and AI-assisted keyword detection to identify potentially risky or ambiguous contractual language, enabling users to practice risk awareness.

Legal Path Builder: Simulates procedural sequencing tasks where users arrange legal steps in correct workflow order (e.g., civil suit filing stages), supporting process literacy.

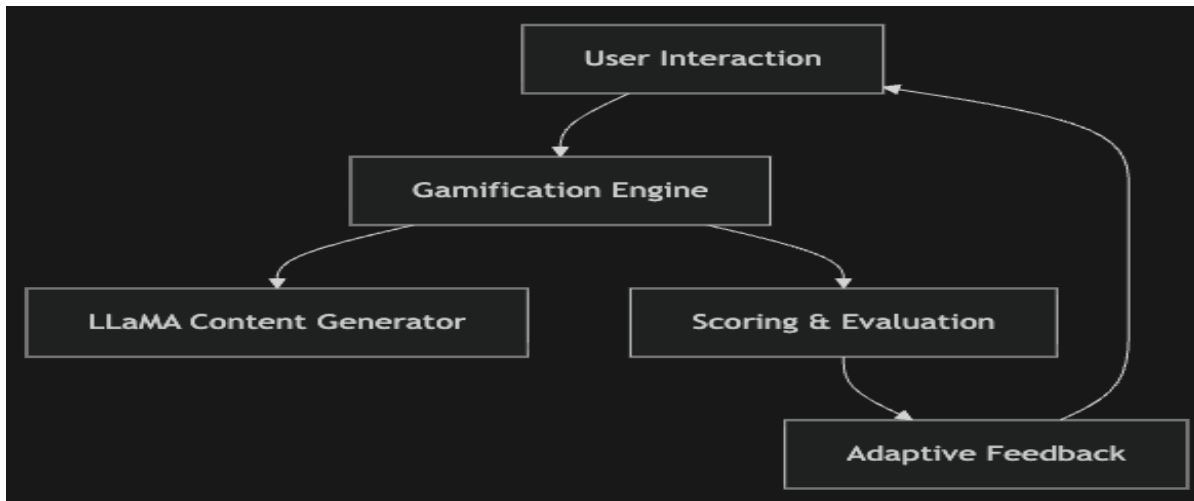
Word Search Reinforcement Module: An interactive puzzle engine designed to reinforce commonly used legal terms through gamified discovery tasks.

Telugu Legal Quiz Module: Provides vernacular-language legal assessment questions to improve regional accessibility and legal literacy. These micro-modules enhance experiential learning while maintaining lightweight computational overhead suitable for mobile environments.

6. EDUCATIONAL GAMIFICATION FRAMEWORK

AI-Driven Legal Pedagogy Framework

JurisAIPocket introduces a gamified legal learning architecture integrating AI-generated moot court problems, clause-risk identification tasks, structured quizzes, legal workflow sequencing, and interactive terminology exercises. This framework supports experiential and self-paced legal education, particularly for law students preparing for moot courts and competitive examinations.



7. System Architecture

7.1 Model Architecture

The system integrates:

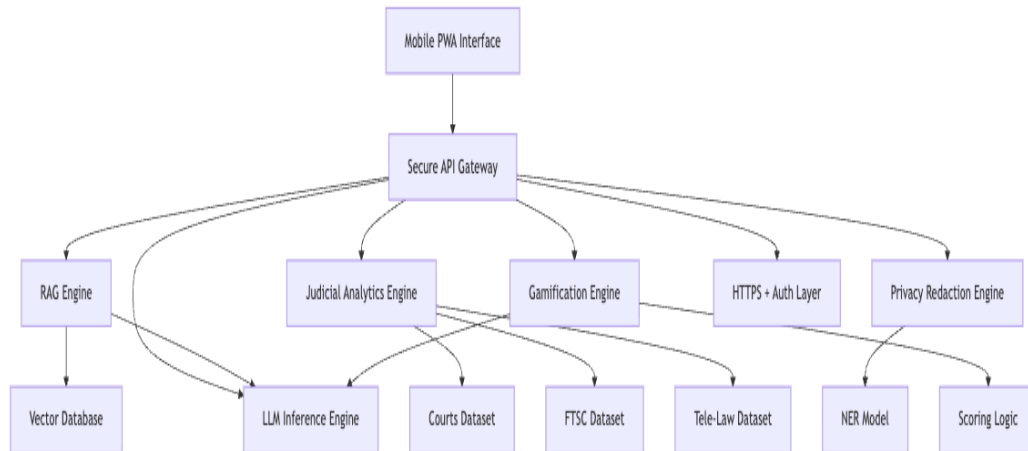
- LLaMA 3.1 70B Instruct Model for text generation,
- Embedding models for semantic retrieval,
- Classification and NER models for redaction tasks.

7.2 RAG Pipeline

The retrieval pipeline consists of:



1. Document preprocessing and OCR (if needed),
2. Clause-based segmentation,
3. Embedding vectorization,
4. Vector search retrieval,
5. Contextual prompt assembly,
6. LLM response generation.



7.3 LLM Integration

The inference backend connects to a distributed LLM environment supporting high-token context for contract handling, enabling long document comprehension.

7.4 Frontend and Deployment Architecture

The proposed architecture is implemented using React with Vite-based Progressive Web Application (PWA) architecture. The mobile-first interface minimizes bundle size and ensures fast rendering performance. Cloudflare DNS and caching layers are integrated to reduce latency and improve global accessibility. The application operates without advertisements, intrusive tracking, or background data mining, reinforcing trust and lightweight deployment. The backend integrates LLaMA inference endpoints and vector retrieval systems via secure API gateways.

8. EVALUATION

8.1 Experimental Setup: The system was evaluated using 250 real-world legal documents including contracts, affidavits, rental agreements, employment agreements, and service agreements. User-based evaluation was conducted with 50 participants comprising law students, junior advocates, and academic researchers.

Performance metrics included response latency, page loading speed, document handling capacity, and usability scoring.

8.2 Performance Evaluation: The system demonstrated efficient performance under standard mobile network conditions:

- **Average LLM response time:** 1.8–2.0 seconds
- **Average page load time (PWA interface):** ~700 milliseconds
- **Document processing capacity tested:** 250 uploaded legal documents
- **Cache-assisted repeated query latency:** Reduced by approximately 35%

The use of vector retrieval grounding and optimized API calls contributed to consistent low-latency responses even for multi-clause contracts.

8.3 Accuracy and Functional Validation: Manual validation was conducted on generated agreements and contract inspection outputs across the 250-document test set. Human reviewers evaluated:

- Clause completeness

- Format correctness
- Logical consistency
- Contextual relevance

The majority of outputs were assessed as structurally coherent and contextually appropriate, with minor refinements required primarily in jurisdiction-specific phrasing.

8.4 Usability Evaluation

User-based testing (N = 50) yielded:

- **Average usability rating:** 4.9 / 5
- High satisfaction in multilingual accessibility
- Strong positive feedback for gamified learning modules
- Reported reduction in manual drafting effort

Participants particularly highlighted the mobile-first interface, fast loading speeds, and simplified interaction flow as key strengths.

8.5 Deployment and Device Compatibility Testing

The system was extensively tested on multiple Android devices to ensure cross-version compatibility and performance stability. Primary development and stress testing were conducted using a OnePlus Nord CE5 device. In addition, compatibility validation was performed across 10 Android devices spanning different Android versions and browser environments (Chrome, Edge, Firefox).

The Progressive Web Application (PWA) architecture was evaluated for offline functionality using service worker caching mechanisms. Once installed, the application remained fully accessible without active internet connectivity, except for cloud-based LLM inference tasks. Core UI components, static resources, educational modules, and cached content were verified to function reliably in offline mode.

These tests confirm the feasibility of deploying AI-assisted legal systems in constrained mobile environments, including regions with intermittent connectivity.

9. RESULTS AND DISCUSSION

The integration of multiple legal capabilities into a single platform provides significant utility for real-world legal workflows. RAG-based contract analysis reduced manual review time, while multilingual interfaces facilitated comprehension among vernacular users. Judicial informatics modules aided research and policy analysis by exposing structural legal datasets with demographic depth. The notarization generator and officer directory streamlined an often fragmented process involving physical verification and localized knowledge. Collectively, these results demonstrate a scalable model for future AI-enabled legal ecosystems.

10. LIMITATIONS

Limitations of this study include:

- Dependency on LLM correctness for generative outputs,
- Need for jurisdiction-specific legal validation,
- Lack of predictive analytics for case outcomes (future scope),
- Limited multi-language support beyond Telugu.

11. ETHICAL AND REGULATORY CONSIDERATIONS

JurisAIPocket is designed with responsible AI deployment principles in mind. The system incorporates safeguards to mitigate hallucination risks associated with large language models by combining retrieval-based grounding mechanisms within contract analysis workflows.

The proposed architecture does not store sensitive user-uploaded legal documents beyond session scope unless explicitly permitted, minimizing data persistence risks. Privacy redaction modules assist users in masking personally identifiable information prior to document sharing.

The system is positioned as an assistive legal intelligence tool and does not substitute licensed legal counsel. Generated contracts and advisory responses require professional validation prior to execution.

Public judicial datasets are sourced from government repositories, ensuring transparency and reproducibility. No demographic prediction or profiling is performed.

The system avoids intrusive advertising, background tracking, and unauthorized data monetization practices, reinforcing ethical digital deployment standards.

All communications between client and server are secured using HTTPS encryption. API endpoints are protected through authentication layers with controlled Cross-Origin Resource Sharing (CORS) policies. The platform follows a zero-hour data retention policy for user-uploaded documents unless explicitly authorized, ensuring privacy-first operation.

12. CONCLUSION AND FUTURE WORK

JurisAIPocket contributes to Legal Informatics by demonstrating an integrated AI-driven ecosystem supporting contract processing, notarization workflows, legal research, and judicial analytics. Future work includes:

- Expanding multilingual support to Hindi and regional languages,
- Integrating legal judgment summarization on National Judicial Data Grid (NJDG) cases,
- Adding compliance auditing for company law and taxation,
- Integrating reinforcement learning from human feedback (RLHF) for improved factual grounding.

13. Acknowledgements

The authors acknowledge the institutional support of Sri Padmavati Mahila Visvavidyalayam (SPMVV) and the contributors of judiciary public datasets enabling informatics research.

REFERENCES

1. Susskind, R. "The Future of Law." Oxford University Press, 2020.
2. Katz, D. M. et al. "Legal Informatics: Technology and the Future of Legal Practice." *Annual Review of Law & Social Science*, 2017.
3. Bommarito, M. J. et al. "LexNLP: Natural Language Processing for Legal Text." *Journal of Open Source Software*, 2018.
4. National eCourts Mission Mode Project, Government of India.
5. Department of Justice, Tele-Law Programme Data, Government of India.
6. Ministry of Law and Justice, Notary Rules and Commissioning Data, Government of India.
7. Touvron, H. et al. "LLaMA: Open and Efficient Foundation Language Models." Meta AI Research, 2023.
8. Harvey AI Official Documentation (2023–2024 Whitepaper).
9. Lexis+ AI Product Documentation (LexisNexis, 2023).
10. Thomson Reuters AI Legal Research Reports.
11. Chalkidis, I., et al. (2020). "LEGAL-BERT: The Muppets straight out of Law School." *Findings of EMNLP 2020*.
12. Hendrycks, D., et al. (2021). "Measuring Massive Multitask Language Understanding (MMLU)." *ICLR 2021*.
13. Zhong, H., et al. (2020). "How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence." *ACL 2020*.
14. Lewis, P., et al. (2020). "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." *NeurIPS 2020*.

15. Ji, Z., et al. (2023). "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys*.
16. McGinnis, J., & Pearce, R. (2014). "The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers." *Fordham Law Review*.
17. Biørn-Hansen, A., et al. (2019). "Progressive Web Apps: The Possible Web-Native Unifier for Mobile Development." *ACM Web Conference Companion*.
18. Guha, N., et al. (2023). "LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in LLMs." *NeurIPS 2023 Workshop*.