

# Insaafai: A Retrieval-Assisted Multilingual Legal Information and Document-Analysis Prototype for Indian Criminal Law

Siddhi Vishwakarma<sup>1</sup> and Riddhi Vishwakarma<sup>2</sup>

Department of Information Technology<sup>1,2</sup>

Kes' B K Shroff College of Arts & M H Shroff College of Commerce, Mumbai, India

vsiddhi549@gmail.com<sup>1</sup> and vriddhi549@gmail.com<sup>2</sup>

**Abstract:** *To Indian Criminal-Law Information Remains Difficult Due To Complex Legal Terminology, Lengthy Statutory Documents, And The Coexistence Of References To The Indian Penal Code (Ipc) And The Bharatiya Nyaya Sanhita (Bns), 2023. This Paper Presents Insaafai, A Retrieval-Assisted Web-Based System Developed To Make Criminal-Law Provisions More Accessible For Educational And Legal-Information Exploration. The Proposed System Integrates Deterministic Section Lookup, Sentence-Bert (Sbert)-Based Semantic Retrieval, Keyword Fallback Search, Pdf/Txt Document Extraction, Multilingual Presentation, And Optional Generative Explanations. The Implemented Knowledge Base Consists Of 575 Structured Ipc/Bns Records Organized Across 23 Chapters. Exact Section Queries Are Handled Through Deterministic Retrieval, Whereas Natural-Language Queries Are Converted Into Semantic Embeddings And Ranked Using Cosine Similarity To Retrieve Contextually Relevant Legal Provisions. Retrieved Statutory Information Is Subsequently Used As Grounded Context For Document Analysis And Conversational Assistance, Helping Reduce Dependence On Unrestricted Model-Generated Responses. Since The Current Implementation Does Not Include An Expert-Labelled Benchmark Or Formal User Study, The Study Focuses On System Architecture And Retrieval Functionality Rather Than Claims Of Legal Or Predictive Accuracy. A Future Evaluation Framework Incorporating Recall@K, Mean Reciprocal Rank (Mrr), Ndcg, Expert Legal Verification, Multilingual Assessment, Privacy Testing, And Overconfidence Analysis Is Proposed. Thstudy demonstrates the potential of retrieval-assisted AI to improve access toTo Indian Criminal-Law Information While Emphasizing The Importance Of Authoritative Legal Sources, Expert Validation, Privacy Safeguards, And Human Oversight.*

**Keywords:** Insaafai, Legal Artificial Intelligence, Indian Penal Code, Bharatiya Nyaya Sanhita, Sentence-BERT, Semantic Search, Legal Information Retrieval

## I. INTRODUCTION

Legal Information Retrieval Is Inherently Complex Because The Relevance Of A Legal Document Depends Not Only On Matching Keywords But Also On Legal Concepts, Factual Context, Jurisdiction, Procedural Setting, And Relationships Among Statutes And Judicial Decisions. Conventional Keyword-Based Search Methods May Therefore Fail When The Terminology Used By A User Differs From The Language Of The Relevant Statutory Provision. Studies Of Legal Information Retrieval (Ir) Have Similarly Identified Specialized Terminology, Complex Document Structures, And Extensive Cross-Referencing As Major Challenges To Effective Legal Search [2]. The Indian Criminal-Law System Introduces An Additional Information-Retrieval Challenge Because Of The Transition From The Indian Penal Code (Ipc), 1860 To The Bharatiya Nyaya Sanhita (Bns), 2023. Although The Bns Now Provides The Principal Framework For Substantive Criminal Law, A Large Body Of Earlier Judgments, Academic Literature, And Educational Material



Continues To Reference Ipc Provisions [3]. Consequently, Legal-Information Systems Must Support Both Precise Section-Level Retrieval And Meaningful Correspondence Between Historical Ipc Provisions And Current Bns Provisions.

Exact Section-Based Search Is Highly Effective When Users Already Know The Relevant Provision Number. However, It Becomes Less Useful When A Legal Issue Is Expressed In Ordinary Or Descriptive Language. Sentence-Bert (Sbert) Helps Address This Limitation By Transforming Textual Queries And Statutory Provisions Into Dense Vector Representations That Can Be Compared Efficiently Using Cosine Similarity [4]. This Enables Retrieval Based On Semantic Similarity Rather Than Exact Word Overlap. Nevertheless, Research In Legal-Domain Natural Language Processing Indicates That General-Purpose Language Representations May Not Fully Capture Specialized Legal Meaning Without Appropriate Domain Adaptation And Task-Specific Evaluation [7]. Legal Retrieval Benchmarks, Including The Fire 2019 Aila Shared Task, Further Demonstrate The Importance Of Evaluating Statute And Precedent Retrieval Using Curated Legal Scenarios And Established Ranked-Retrieval Measures [5].

Against This Background, Insaafai Was Developed As An Educational Legal-Information Prototype Integrating Structured Ipc/Bns Data, Deterministic Section Lookup, Semantic Retrieval, Document Text Extraction, Multilingual Presentation, And Optional Generative Explanations Within A Unified Web-Based Environment. The System Is Intended To Support Exploratory Access To Legal Information Rather Than Provide Autonomous Legal Advice Or Predict Judicial Outcomes.

Accordingly, The Central Research Question Of This Study Is: How Can A Lightweight Web-Based System Integrate Structured Statutory Lookup, Semantic Retrieval, Multilingual Presentation, And Document Analysis While Maintaining Source Traceability, Uncertainty Disclosure, And Appropriate Human Oversight?

This Study Makes Four Principal Contributions: (I) An Implementation-Grounded Description Of The Insaafai Prototype; (Ii) A Formal Description Of Its Hybrid Exact, Keyword, And Semantic Retrieval Pipeline; (Iii) Measurable Observations Regarding The Implemented Legal Corpus And Application Functionality; And (Iv) A Validation And Governance Framework That Clearly Distinguishes Demonstrated System Capabilities From Performance Claims Requiring Future Empirical And Expert Evaluation.

## **II. REVIEW OF LITERATURE**

### **A. Legal Information Retrieval**

Legal Information Retrieval (LIR) Focuses On Identifying Statutes, Judicial Decisions, Legal Passages, And Other Authoritative Materials Relevant To A Legal Query. Unlike Conventional Information Retrieval, Legal Ir Must Account For Specialized Terminology, Lengthy Documents, Contextual Relationships, And Cross-References Between Legal Authorities. Sansone And Sperli [2] Classify Contemporary Legal Ir Approaches Across Natural Language Processing (Nlp), Machine Learning, And Knowledge-Extraction Techniques, While Highlighting Persistent Challenges Related To Long-Document Processing, Domain-Specific Language, Explainability, And Reliable Evaluation. In The Indian Context, The Aila Benchmark Required Systems To Retrieve Relevant Statutes And Prior Cases From Factual Legal Scenarios, Providing An Important Framework For Evaluating Narrative-To-Statute Retrieval [5]. Collectively, These Studies Support The Use Of Hybrid Retrieval Approaches That Combine Lexical And Semantic Techniques While Emphasizing The Need For Systematic Evaluation Rather Than Assuming Relevance From Isolated Examples.

### **B. Dense Embeddings And Domain Adaptation**

Semantic Retrieval Methods Address Situations In Which A User's Wording Differs From The Terminology Used In Statutory Provisions. Sentence-Bert (Sbert) Employs Siamese And Triplet Network Architectures To Generate Dense Sentence Embeddings That Enable Efficient Semantic Similarity Comparison [4]. Insaafai Employs The All-MiniLM-L6-V2 Sentence-Transformer To Perform Semantic Retrieval Over Its Relatively Small Statutory Corpus. This Provides A Computationally Efficient Solution For A Prototype; However, The Model Cannot Be Assumed To Fully Capture Specialized Indian Criminal-Law Concepts. Research on Legal-BERT demonstrates that domain-specific pre-training



and adaptation can influence performance on legal NLP tasks [7]. Similarly, the BEIR benchmark shows substantial variation in retrieval performance across different domains and datasets, Reinforcing The Importance Of Task-Specific And Zero-Shot Evaluation [8]. These findings justify future comparison of general-purpose and legal-domain embedding models for Indian Statutory Retrieval.

### **C. Retrieval-Augmented Generation**

Retrieval-Augmented Generation (RAG) Combines Information Retrieval With Generative Language Models By Conditioning Generated Responses On Externally Retrieved Information [9]. In Legal-Information Systems, This Approach Can Improve Source Traceability And Contextual Grounding By Retrieving Potentially Relevant Statutory Provisions Before Producing An Explanation. However, Successful Retrieval Does Not Establish That A Provision Is Legally Applicable, Nor Does Fluent Generation Guarantee Legally Correct Reasoning. Research On Legal Retrieval-And-Generation Tasks Has Demonstrated That Even Capable Language Models May Produce Unsupported Or Inaccurate Legal Content [10]. Therefore, Legal Rag Systems Should Evaluate Retrieval Relevance, Citation Correctness, Factual Grounding, And Generated Explanations Separately, Rather Than Relying Solely On The Apparent Quality Or Fluency Of The Final Response.

### **D. Responsible AI, Privacy, And Multilingual Access**

Responsible Deployment Is Particularly Important For AI-Assisted Legal-Information Systems. The Nist Ai Risk Management Framework Organizes Trustworthy Ai Practices Around Governance, Mapping, Measurement, And Risk Management, Emphasizing Reliability, Safety, Transparency, Privacy, And Fairness [6]. Its Generative Ai Profile Additionally Identifies Risks Associated With Confabulation, Information-Integrity Failures, And Excessive Human Reliance On Ai-Generated Outputs [11]. Privacy Is Also Significant Because Uploaded Legal Documents May Contain Personally Identifiable Or Sensitive Information. Indian Implementations Must Therefore Consider Applicable Requirements Under The Digital Personal Data Protection Act, 2023, Together With Relevant Rules And Data-Handling Safeguards [12].

Multilingual Accessibility Is Equally Important In The Indian Legal Environment. Machine-Translation Systems Such As Indictrans2 Provide Broad Coverage Across Indian Languages And Demonstrate The Feasibility Of Multilingual NLP Applications [13]. Nevertheless, General Translation Performance Does Not Guarantee Preservation Of Statutory Terminology, Legal Interpretation, Or Contextual Meaning. Multilingual Legal Outputs Should Therefore Undergo Evaluation By Bilingual Legal Reviewers To Ensure Semantic And Legal Consistency Across Languages.

## **III. RESEARCH METHODOLOGY**

### **A. Research Design And Evidence Boundary**

This Study Adopts A Design-And-Implementation Case-Study Approach To Examine The Architecture And Functionality Of The Insaafai Prototype. The Unit Of Analysis Comprises The Supplied System Artifacts, Including The Backend And Frontend Components, Structured Legal Dataset, Cached Embeddings, Project Documentation, And Bundled Test Documents [1]. The Methodology Distinguishes Between Directly Observed Implementation Features, Findings Supported By Existing Literature, And Capabilities Requiring Future Empirical Validation.

The Present Study Does Not Claim Improvements In Legal Literacy, Prediction Accuracy, Professional Productivity, or legal correctness because the available implementation does not contain a controlled user study or an expert-labelled benchmark dataset. Consequently, The Reported Findings Are Limited To Verifiable System Characteristics And Retrieval Behaviour.

### **B. Corpus And Data Representation**

The Primary Structured Legal Resource Used By Insaafai Is [ipc\\_data.json](#). Direct Inspection Of The Dataset Identified 575 Structured Records Distributed Across 23 Chapter Labels. Each Record Contains An Ipc Section Identifier, Title,



Description, Chapter Information, And Corresponding Bns-Related Fields. The Project Archive Additionally Contains Statutory Documents And A Limited Collection Of Court-Decision PDFs Used For Development And Testing Purposes [1]. The Presence Of A Legal Document Or Ipc-Bns Correspondence Within The Project Archive Is Not Considered Sufficient Evidence Of Its Authoritative Legal Status. Prior To Deployment, Statutory Provisions And Ipc-Bns Mappings Should Therefore Be Validated Against Official And Version-Controlled Legal Sources, With Appropriate Documentation Of Amendments, Effective Dates, Exceptions, And Provenance [3].

### C. System Architecture

Insaafai Follows A Lightweight Client-Server Web Architecture. The Backend Is Implemented Using Flask And Exposes Application Programming Interfaces Through An /Api Prefix, While The Frontend Is Developed Using Html, CSS, and JavaScript. Project Documentation Also Specifies Firebase-Based Functionality For Authentication And User-Data Management [1].

During Initialization, The Backend Loads The Structured Legal Records And Either Loads Previously Cached Embeddings Or Generates The Required Semantic Representations. Users Can Provide Input In The Form Of A Section Number, Keyword, Natural-Language Case Description, Conversational Query, Or Pdf/Txt Document. The System Validates The Submitted Input, Extracts Textual Content Where Required, Retrieves Potentially Relevant Ipc/Bns Provisions, And Returns Structured Legal Metadata. For Generative Functions, Retrieved Statutory Provisions Are Supplied To The Language Model As Contextual Evidence For Producing An Explanation.

### D. Hybrid Retrieval Procedure

Insaafai Combines Deterministic Section Matching, Semantic Retrieval, And Keyword-Based Fallback Searching. When A Query Contains A Recognizable Ipc, Bns, Or General Section-Number Pattern, The System First Attempts An Exact Match. This Approach Provides Reliable Retrieval When The User Already Knows The Relevant Statutory Section.

For Descriptive Or Natural-Language Queries, Semantic Similarity Between The User Query  $Q$  And A Statutory Record  $D_i$  Is Calculated Using Cosine Similarity Between Their Sbert Embedding Vectors:  $\text{Similarity}(Q, D_i) = [E(Q) \cdot E(D_i)] / [\|E(Q)\| \times \|E(D_i)\|]$

Where:  $E(Q)$  = Embedding Vector Representing The User's Query,

$E(D_i)$  = Embedding Vector Representing The  $i$ th Statutory Record,

$E(Q) \cdot E(D_i)$  = Dot Product Of The Two Vectors, And

$\|E(Q)\|$  And  $\|E(D_i)\|$  = Magnitudes Of The Corresponding Vectors.

A Higher Cosine-Similarity Value Indicates Greater Semantic Similarity Between The Query And Statutory Record. The Implementation Retains Semantic Candidates Above A Threshold Of 0.25, Or 0.35 Following An Exact Match, While Similarity Scores Above 0.45 Are Displayed By The Interface As "High Confidence." If Semantic Retrieval Does Not Produce An Adequate Result, A Title-Keyword Fallback Is Applied With A Fixed Score Of 0.20 [1]. These Values Represent Retrieval Heuristics And Similarity Thresholds Rather Than Calibrated Probabilities Of Legal Correctness.

### E. Document-Analysis Procedure

For Uploaded Pdf Documents, The System Extracts Textual Content From Individual Pages And Examines An Initial Portion Of The Extracted Text To Identify Potentially Relevant Statutory Provisions. A Truncated Representation Of The Document, Together With The Retrieved Ipc/Bns Context, Is Subsequently Supplied To The Optional Generative Component. The Generated Response May Contain The Document Type, Summary, Identified Entities, Relevant Statutory Provisions, Legal Reasoning, Implications, Suggested Actions, And Model-Generated Risk Or Probability Statements [1]. Such Probability Or Risk Outputs Are Not Treated As Statistically Validated Predictions In This Study. For Safer Deployment, The Interface Should Present Supporting Statutory Passages, Clearly Distinguish Retrieval Similarity Scores From Legal Conclusions, And Avoid Probability-Based Terminology Unless Appropriate Calibration And Expert Validation Have Been Performed.



#### F. Proposed Evaluation Protocol

Future Empirical Validation Should Employ A Version-Controlled Benchmark Consisting Of Anonymized Legal Questions And Factual Scenarios. At Least Two Qualified Legal Annotators Should Independently Identify Relevant Statutory Provisions, Cite Authoritative Sources, And Document Disagreements. The Performance Of Exact Lookup, Keyword-Based Retrieval, Dense Semantic Retrieval, And The Complete Hybrid Retrieval System Should Then Be Compared Under Identical Experimental Conditions. Retrieval Effectiveness Should Be Evaluated Using Recall@K, Mean Reciprocal Rank (Mrr), And Normalized Discounted Cumulative Gain (Ndcg). Confidence Intervals Obtained Through Bootstrap Resampling Should Accompany Aggregate Performance Measures. Document Extraction, Multilingual Output, Generated Explanations, Privacy, Safety, And System Latency Should Be Assessed Independently. This Component-Wise Evaluation Prevents Strong Performance In One Subsystem From Concealing Weaknesses In Another.

Table I. Proposed Validation Matrix

| Component             | Primary Measure                                       | Purpose                                                                  |
|-----------------------|-------------------------------------------------------|--------------------------------------------------------------------------|
| Section Retrieval     | Recall@1/3, Mrr, Ndcg                                 | Evaluates Whether Relevant Provisions Are Retrieved And Correctly Ranked |
| Exact Lookup          | Accuracy And Latency                                  | Verifies Deterministic Section Retrieval And Response Efficiency         |
| Text Extraction       | Word/Character Recall                                 | Identifies Missing, Incomplete, Or Corrupted Document Text               |
| Generated Explanation | Expert Factuality And Citation Precision              | Evaluates Factual Grounding And Source Consistency                       |
| Multilingual Output   | Adequacy And Terminology Consistency                  | Determines Whether Legal Meaning Is Preserved Across Languages           |
| Safety And Privacy    | Unsupported-Advice, Overconfidence, And Leakage Rates | Evaluates High-Risk System Failures And Privacy Concerns                 |

## IV. RESULTS

#### A. Verified Implementation Findings

Static Inspection Of The Supplied Insaafai Implementation Confirmed The Major Functional And Data Characteristics Summarized In Table Ii. The Structured Legal Dataset Contains 575 Records Distributed Across 23 Chapter Labels, With Title Information And Corresponding Bns Mapping Fields Represented Throughout The Inspected Dataset. The Backend Configuration Specifies All-Minilm-L6-V2 As The Sentence-Embedding Model Used For Semantic Retrieval. The Implementation Further Confirms A Three-Stage Retrieval Strategy Consisting Of Exact Section Matching, Semantic Similarity Retrieval, And Keyword-Based Fallback Search. Document-Processing Routes Accept Pdf And Txt Files, While The User-Facing Application Provides Functions For Statutory Search, Text Summarization, Document Analysis, And Conversational Interaction. These Observations Establish The Implemented Scope Of The Prototype But Should Not Be Interpreted As Evidence Of Legal Accuracy, Retrieval Effectiveness, Or Professional Reliability.



Table II. Observed Characteristics Of The Supplied Prototype

| Finding                      | Observed Value Or Behavior               | Evidence                   |
|------------------------------|------------------------------------------|----------------------------|
| Structured Legal Records     | 575                                      | Ipc_data.Json Inspection   |
| Chapter Labels Represented   | 23                                       | Unique-Value Count         |
| Records Containing Titles    | 575                                      | Dataset Field Inspection   |
| Bns Mapping Fields Populated | 575                                      | Dataset Field Inspection   |
| Embedding Model              | All-Minilm-L6-V2                         | Backend Configuration      |
| Retrieval Stages             | Exact → Semantic → Keyword Fallback      | Backend Retrieval Logic    |
| Accepted Document Types      | PDF and TXT                              | Upload-Route Inspection    |
| Main User Functions          | Search, Summary, Document Analysis, Chat | Routes And Interface Files |

## B. Functional Interpretation

Analysis Of The Retrieval Workflow Identified Three Principal Query-Handling Conditions. First, When The Input Contains A Recognized Ipc Or Bns Section Number, The System Performs Deterministic Lookup To Locate The Corresponding Structured Record. This Provides A Direct Retrieval Mechanism For Users Who Already Know The Relevant Statutory Provision.

Second, When The User Submits A Natural-Language Description Or Factual Scenario, The Query Is Transformed Into An Embedding And Compared With Stored Statutory Embeddings Using Cosine Similarity. Candidate Provisions Are Then Ranked According To Semantic Similarity. Third, When Deterministic And Semantic Retrieval Do Not Adequately Resolve A Query, The System Can Apply A Title-Based Keyword Fallback, Providing An Additional Retrieval Pathway. This Hybrid Sequence Combines The Precision Of Exact Section Matching With The Flexibility Of Semantic Search. The Document-Analysis Component Further Uses Retrieved Ipc/Bns Records As Contextual Information For Generative Processing. Consequently, Generated Explanations Can Be Connected To Locally Retrieved Statutory Content Rather Than Relying Entirely On The Language Model's Internal Knowledge. However, Retrieval Relevance Represents Computational Similarity Rather Than A Determination Of Legal Applicability, And The Highest-Ranked Provision Should Therefore Not Automatically Be Interpreted As The Legally Correct Provision.

## C. Absence Of Predictive Evidence

Inspection Of The Supplied Project Materials Did Not Identify An Expert-Labelled Retrieval Benchmark, Confusion Matrix, Calibration Experiment, Controlled User Study, Or Comparative Performance Evaluation. Therefore, The Current Implementation Does Not Provide Sufficient Empirical Evidence To Report Numerical Claims Concerning Legal Accuracy, Prediction Accuracy, Retrieval Effectiveness, Or Improvements In User Performance.

Similarly, Model-Generated Probability Values And Interface Labels Such As “High Confidence” Represent Implementation-Level Outputs Rather Than Statistically Calibrated Estimates Of Legal Correctness. They Should Not Be Interpreted As Probabilities That A Particular Legal Provision, Argument, Or Outcome Is Correct.

This Distinction Is An Important Result Of The Implementation Analysis. It Establishes Insaafai As A Functional Retrieval-Assisted Legal-Information Prototype Rather Than A Validated Legal Decision Or Case-Outcome Prediction System. Quantitative Performance Claims Should Therefore Be Reserved For Future Experiments Using Expert-Labelled Datasets, Established Retrieval Metrics, Controlled Comparisons, And Appropriate Statistical Validation.



## **V. DISCUSSION**

### **A. Principal Interpretation**

The Principal Contribution Of Insaafai Lies In Its Integrated Architecture For Retrieval-Assisted Access To Indian Criminal-Law Information. The Prototype Demonstrates How A Locally Indexed And Structured Legal Corpus Can Support Exact Statutory Lookup, Natural-Language Semantic Retrieval, Document-Assisted Analysis, And Multilingual Presentation Through A Unified Web-Based Service Layer. Rather Than Depending Entirely On Generative Ai, The System Introduces An Explicit Retrieval Stage In Which Potentially Relevant Statutory Provisions Are Identified Before An Explanation Is Generated.

This Architecture Offers An Important Advantage In Terms Of Traceability And Auditability, Since Retrieved Provisions Can Be Presented To Users As Supporting Context For Generated Responses. Such An Approach Is Consistent With Established Research In Legal Information Retrieval And Retrieval-Augmented Generation [2], [5], [9]. However, Architectural Feasibility Should Be Distinguished From Demonstrated Effectiveness. The Current Implementation Establishes That These Components Can Be Integrated Successfully, While Their Retrieval Accuracy, Legal Reliability, And Usefulness To End Users Remain Subjects For Empirical Evaluation.

### **B. Strengths**

A Major Strength Of The Prototype Is Its Hybrid Retrieval Architecture, Which Separates Deterministic Section Lookup From Semantic Similarity Matching And Keyword-Based Fallback. Exact Retrieval Provides Predictable Behaviour For Known Statutory Sections, Whereas Semantic Retrieval Enables Users To Search Through Ordinary-Language Descriptions Without Requiring Precise Legal Terminology.

The System Also Caches Corpus Embeddings, Reducing The Need To Recompute Representations For Every Request And Making Semantic Retrieval More Practical For A Relatively Small Legal Corpus. Support For Both Short Textual Queries And Document-Scale Pdf/Txt Inputs Extends The System Beyond Conventional Section Browsing. Furthermore, Structured Response Fields And The Inclusion Of Ipc/Bns Metadata Improve Inspectability Compared With A Purely Generative Chatbot In Which The Origin Of Legal Information May Be Unclear. The Educational Framing Of Insaafai And Its Use Of Non-Binding Legal-Information Language Also Provide Appropriate Initial Safeguards For A Prototype.

### **C. Limitations**

The Most Significant Limitation Is The Absence Of Expert-Labelled Evaluation Data. Without A Validated Benchmark, It Is Not Possible To Determine Whether Retrieved Provisions Are Consistently Relevant Or Legally Applicable. The Use Of A General-Purpose Sentence-Embedding Model Also Creates The Possibility That Legally Distinct Situations Containing Similar Vocabulary May Receive High Semantic-Similarity Scores.

The Current Retrieval Thresholds Are Heuristically Defined Rather Than Empirically Calibrated, Meaning That Interface Confidence Labels Should Not Be Interpreted As Probabilities Of Legal Correctness. Similarly, Static Ipc-Bns Mappings May Become Incomplete Or Outdated Unless Statutory Amendments, Effective Dates, Exceptions, And Source Provenance Are Systematically Maintained.

Document Processing Introduces Additional Limitations. PDF Text Extraction May Perform Poorly On Scanned Documents, Multi-Column Layouts, Images, Or Low-Quality Files, Potentially Affecting Downstream Retrieval. Multilingual Translation May Also Modify Subtle Legal Terminology Or Contextual Meaning. Generative Explanations Remain Vulnerable To Incorrect Citations, Omitted Statutory Exceptions, Unsupported Interpretations, And Excessive Confidence.

Privacy Represents Another Significant Concern Because Uploaded Legal Documents May Contain Personal, Confidential, Or Sensitive Information. Finally, The Limited And Uncontrolled Collection Of Development Documents Available In The Current Project Cannot Support Broad Conclusions About System Performance Across Different Legal Cases Or Document Types.



**D. Recommended Improvements**

Future Development Should Prioritize Both Retrieval Performance And Responsible Deployment. Each Statutory Provision Displayed By The System Should Be Linked To An Authoritative And Version-Controlled Legal Source, Together With Amendment History And Effective-Date Metadata. Retrieval Experiments Should Compare Lexical Methods Such As Bm25, Dense Embedding Retrieval, And Hybrid Approaches Using The Same Expert-Labelled Benchmark.

Similarity Thresholds Should Be Optimized On A Held-Out Validation Dataset Rather Than Selected Heuristically. Generated Explanations Should Provide Citation-Level Supporting Evidence, Allowing Users To Distinguish Retrieved Statutory Text From Model-Generated Interpretation. Uncalibrated Probability Or Confidence Terminology Should Also Be Replaced With Transparent Uncertainty Statements Until Appropriate Calibration Has Been Demonstrated.

Document Handling Should Incorporate OCR-Quality Assessment, File Validation, Encryption, Controlled Data Retention, And Secure Deletion Mechanisms. Multilingual Outputs Should Undergo Bilingual Legal Review To Determine Whether Statutory Terminology And Legal Meaning Remain Consistent Across Languages. High-Risk Or Case-Specific Queries Should Additionally Provide An Appropriate Escalation Pathway To Qualified Legal Professionals. Behavioral Testing Can Complement Conventional Benchmark Metrics By Systematically Examining Predictable Failure Categories Such As Unsupported Citations, Misleading Confidence, Incomplete Retrieval, And Incorrect Statutory Correspondence [14].

**E. Practical Implications**

For Students And Learners, Insaafai Has Potential As A Guided Statute-Discovery Environment In Which Relevant Ipc/Bns Provisions Can Be Located From Either Section Numbers Or Ordinary-Language Descriptions. Its Educational Value Would Be Strengthened When Generated Explanations Are Accompanied By Clearly Identifiable Primary Statutory Text.

For Researchers And Developers, The Prototype Provides A Useful Experimental Platform For Studying Hybrid Legal Retrieval, Multilingual Interfaces, Document-Grounded Analysis, And Retrieval-Assisted Generation In The Indian Legal Context. It May Also Support Future Comparative Studies Of Lexical, Semantic, And Legal-Domain Retrieval Models.

Nevertheless, The Present Prototype Should Not Be Used To Determine Legal Rights, Recommend Litigation Strategies, Predict Judicial Outcomes, Or Substitute For Professional Legal Advice. In a high-stakes legal-information environment, Human Review Should Be Treated As An Integral Component Of System Operation Rather Than Merely An Optional Final Verification Step [6], [11].

**VI. CONCLUSION AND FUTURE SCOPE****A. Conclusion**

This Study Presented Insaafai, A Retrieval-Assisted And Multilingual Web-Based Prototype Designed To Improve Educational Access To Indian Criminal-Law Information. The Implemented System Integrates 575 Structured Ipc/Bns Records, Deterministic Section-Level Lookup, Sentence-Bert (Sbert)-Based Semantic Retrieval, Keyword Fallback Search, Pdf/Txt Document Extraction, And Optional Generative Explanations Within A Unified Architecture. By Supporting Both Statutory Identifiers And Ordinary-Language Queries, The System Provides A Practical Mechanism For Exploring Legal Provisions Without Requiring Users To Know Precise Legal Terminology.

The Study Demonstrates The Architectural Feasibility And Implemented Functionality Of The Proposed Approach Rather Than Claiming Validated Legal Correctness Or Case-Outcome Prediction Accuracy. The Combination Of Deterministic And Semantic Retrieval Improves The Flexibility And Traceability Of Legal-Information Access, While Retrieved Statutory Provisions Provide Contextual Grounding For Generative Explanations. Nevertheless, Similarity Scores And Generated Responses Cannot Substitute For Expert Legal Interpretation.



Overall, Insaafai Provides A Foundation For Developing More Accessible, Transparent, And Retrieval-Grounded Legal-Information Systems. Its Present Role Is Best Understood As An Educational And Research Prototype In Which Ai Assists Users In Locating And Understanding Potentially Relevant Legal Information While Preserving The Essential Role Of Qualified Human Judgment.

### **B. Future Scope**

Future Development Should Focus On Transforming The Prototype Into A Systematically Evaluated And Responsibly Governed Legal-Information Platform. The Statutory Knowledge Base Should Be Connected To Authoritative And Version-Controlled Legal Sources, With Mechanisms For Tracking Amendments, Effective Dates, Provision Mappings, And Source Provenance.

A Comprehensive Expert-Labelled Benchmark Should Be Developed To Evaluate Exact, Lexical, Semantic, And Hybrid Retrieval Approaches Using Metrics Such As Recall@K, Mrr, And Ndcg. Future Research Should Also Investigate Legal-Domain Embedding Models And Compare Them With The Currently Implemented General-Purpose Sbert Model. Generated Explanations Should Undergo Citation-Level Factuality And Grounding Evaluation, While Multilingual Outputs Should Be Reviewed By Bilingual Legal Experts To Ensure Preservation Of Statutory Terminology And Legal Meaning. Additional Work Is Required In Privacy-Preserving Document Processing, Secure Data Handling, Ocr Support, Uncertainty Calibration, Behavioral Safety Testing, And Continuous Risk Monitoring.

With Systematic Validation, Authoritative Data Management, Appropriate Privacy Safeguards, And Explicit Human Oversight, The Proposed Architecture Could Evolve Into A Valuable Platform For Legal Education, Statutory Research, Multilingual Legal-Information Discovery, And Responsible AI-Assisted Legal Exploration.

### **ACKNOWLEDGMENT**

The Authors Gratefully Acknowledge The Guidance And Institutional Support Provided By [Department Of Information Technology, Kes' B K Shroff College Of Arts & M H Shroff College Of Commerce,] During The Development Of This Research. The Authors Also Sincerely Acknowledge Each Other's Collaborative Contributions, Technical Support, Research Inputs, And Continuous Cooperation Throughout The Design, Development, Analysis, And Preparation Of The Insaafai Project And This Manuscript. The Collective Efforts And Constructive Discussions Among The Authors Significantly Contributed To The Successful Completion Of This Work.

### **Declarations**

**Conflict Of Interest:** The Authors Declare That There Are No Conflicts Of Interest Associated With This Research.

**Funding:** The Authors Received No External Funding For The Development Of This Project Or Preparation Of This Manuscript.

**Ethics And Data:** This Study Involved The Design And Implementation Of A Software Prototype And Did Not Include A Documented Experiment Involving Human Participants. Any Future Research Involving Users, Personal Legal Records, Or Sensitive Information Should Follow Applicable Requirements For Ethical Approval, Informed Consent, Anonymization, Secure Data Handling, And Data-Protection Compliance.

**Author Contributions:** All Authors Contributed Collaboratively To The Conceptualization, System Development, Research Analysis, Interpretation Of Findings, And Preparation Of The Manuscript. All Authors Reviewed The Final Manuscript And Approved It For Submission.

### **REFERENCES**

- [1] Author-Supplied Project Archive, "Insaafai Ai Legal Assistant," Source Code, Structured Data, Documentation, And Bundled Development Documents, 2026.
- [2] C. Sansone And G. Sperli, "Legal Information Retrieval Systems: State-Of-The-Art And Open Issues," Information Systems, Vol. 106, Art. No. 101967, 2022, Doi: 10.1016/J.Is.2021.101967.



- [3] Government Of India, "The Bharatiya Nyaya Sanhita, 2023," Act No. 45 Of 2023, Dec. 25, 2023. [Online]. Available: <https://www.indiacode.nic.in/bitstream/123456789/20062/1/A2023-45.pdf>
- [4] N. Reimers And I. Gurevych, "Sentence-Bert: Sentence Embeddings Using Siamese Bert-Networks," In Proc. EMNLP-IJCNLP, 2019, Pp. 3982-3992, Doi: 10.18653/V1/D19-1410.
- [5] P. Bhattacharya Et Al., "Fire 2019 Aila Track: Artificial Intelligence For Legal Assistance," In Proc. Fire 2019, 2019, Pp. 4-6, Doi: 10.1145/3368567.3368587.
- [6] E. Tabassi, "Artificial Intelligence Risk Management Framework (Ai Rmf 1.0)," Nist Ai 100-1, National Institute Of Standards And Technology, 2023, Doi: 10.6028/Nist.Ai.100-1.
- [7] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, And I. Androutsopoulos, "Legal-Bert: The Muppets Straight Out Of Law School," In Findings Of Emnlp, 2020, Pp. 2898-2904, Doi: 10.18653/V1/2020.Findings-Emnlp.261.
- [8] N. Thakur Et Al., "Beir: A Heterogeneous Benchmark For Zero-Shot Evaluation Of Information Retrieval Models," In Proc. Neurips Datasets And Benchmarks, 2021.
- [9] P. Lewis Et Al., "Retrieval-Augmented Generation For Knowledge-Intensive Nlp Tasks," In Advances In Neural Information Processing Systems, Vol. 33, 2020, Pp. 9459-9474.
- [10] A. B. Hou Et Al., "Clerc: A Dataset For Legal Case Retrieval And Retrieval-Augmented Analysis Generation," In Proc. Iclr, 2025. [Online]. Available: <https://arxiv.org/abs/2406.17186>
- [11] National Institute Of Standards And Technology, "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile," Nist Ai 600-1, 2024, Doi: 10.6028/Nist.Ai.600-1.
- [12] Government Of India, "The Digital Personal Data Protection Act, 2023," Act No. 22 Of 2023, Aug. 11, 2023. [Online]. Available: <https://www.meity.gov.in/static/uploads/2024/02/2f5050ecf4a1eec6a14a5e1d01e5dc35.pdf>
- [13] V. Gala Et Al., "Indictrans2: Towards High-Quality And Accessible Machine Translation Models For All 22 Scheduled Indian Languages," Transactions On Machine Learning Research, 2023. [Online]. Available: <https://openreview.net/forum?id=Vft4yuzaya>
- [14] M. T. Ribeiro, T. Wu, C. Guestrin, And S. Singh, "Beyond Accuracy: Behavioral Testing Of Nlp Models With Checklist," In Proc. Acl, 2020, Pp. 4902-4912, Doi: 10.18653/V1/2020.Acl-Main.442.

