



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

“A Literature Survey On The Role Of Artificial Intelligence In Transforming The Indian Judicial System: Models, Applications, And Ethical Implications”

Authors:

- **Naresh Kumar Yalla**, Research Assistant, Department of Computer Science Engineering, JNTUH University College of Engineering Science and Technology, Hyderabad
- **Dr.V. Kamakshi Prasad**, Sr. Professor, Department of Computer Science Engineering, JNTUH University College of Engineering Science and Technology, Hyderabad

Abstract: This literature survey explores the transformative potential of Artificial Intelligence (AI) within the Indian judicial system. With an overwhelming backlog of cases, linguistic diversity, and procedural complexities, India's judiciary presents a compelling opportunity for AI-driven interventions. The survey critically examines contemporary research focusing on the integration of AI into legal processes—particularly in the domains of legal document analytics, natural language processing (NLP), multilingual summarization, bias detection, semantic reasoning, translation, and domain-specific large language models. Through a systematic review of academic contributions, this study establishes a comprehensive foundation for future AI innovations tailored to enhance judicial efficiency, transparency, and accessibility. The findings aim to support a vision where AI not only augments the judiciary's functionality but also advances the broader goal of timely and equitable justice delivery for the common citizen.

1. Introduction

India's judicial system is one of the most burdened in the world, grappling with an unprecedented backlog of over 45 million cases across various courts. Delays due to procedural inefficiencies, linguistic diversity, voluminous documentation, and limited legal aid access hinder the timely delivery of justice. Against this backdrop, Artificial Intelligence (AI)—encompassing machine learning, natural language

processing (NLP), knowledge graphs, and large language models (LLMs)—has emerged as a promising enabler of reform.

Recent advancements underscore AI's potential in automating repetitive legal tasks, summarizing complex judgments, identifying legal entities, retrieving precedents, and supporting judges in case prioritization. The adoption of AI in the judiciary is not merely a matter of technological upgrade but a transformative shift toward improving institutional functionality and access to justice. This survey systematically reviews recent scholarly and technical contributions that explore AI applications in the Indian legal domain. Emphasis is placed on domain-adapted legal language models, multilingual summarization efforts, semantic bias detection, human-AI collaboration frameworks, and ethical considerations.

By analyzing these developments, this paper seeks to provide a roadmap for leveraging AI technologies to strengthen judicial infrastructure and facilitate a more responsive and citizen-friendly legal system.

2. Advancements in Legal AI and Indian Judicial Systems: A Comprehensive Survey

Recent strides in applying large language models (LLMs) to the legal domain have catalyzed the development of jurisdiction-aware architectures, robust evaluation benchmarks, and innovative knowledge integration techniques.

A notable study titled "Analyzing the Effectiveness of AI in the Indian Judicial System"¹² explores how artificial intelligence can enhance the operational efficiency of India's judiciary. Conducted by a group of law students leveraging data and documents from the National Judicial Data Grid (NJDG), this research undertakes a comprehensive data science approach to address delays and administrative bottlenecks in the legal system. The primary objective of the study is to evaluate the role of AI applications in mitigating systemic inefficiencies through legal analytics, intelligent case filtering, and automated document summarization. The paper highlights how AI tools can effectively manage repetitive administrative tasks, minimize manual errors, and support judges and legal practitioners in case prioritization. The study contributes significant quantitative findings, notably indicating that AI deployment in pilot scenarios can potentially reduce case pendency by up to 25%.

A significant leap in this trajectory is InLegalLLaMA (Ghosh et al., 2024)¹³, which introduced a domain-specialized Indian legal model fine-tuned from LLaMA-2 and augmented with a retrieval-augmented generation (RAG) framework to automate petition drafting. This builds upon foundational efforts like LegalBERT (Chalkidis et al., 2020)⁵, which adapted BERT to European legal corpora, subsequently inspiring Indian contextual adaptations such as InLegalBERT and InCaseLawBERT (Agarwal et al., 2022)¹², both tailored for tasks including statute identification and judgment prediction using Indian case law. Together, these innovations signify a maturing ecosystem of legal AI, advancing both generalizable and India-centric legal NLP capabilities.

To rigorously evaluate such models, IL-TUR (Joshi et al., 2023)⁴ and LegalBench (Guha et al., 2023)¹⁰ have emerged as critical benchmarks assessing multilingual legal reasoning and comprehension. Training of such models is further enhanced by resources like the Indian Legal Document Corpus (Malik et al., n.d.)⁶, a curated dataset of over 35,000 annotated court judgments.

The Marino et al. (2023)³ paper introduces LEGAL-TransformerOverBERT (LEGAL-ToBERT), a hierarchical transformer model designed to classify rhetorical roles in legal documents. The model stacks a transformer encoder over a legal-domain-specific BERT model, effectively capturing relationships between sentences. Evaluations on Italian and English legal judgment datasets demonstrate that LEGAL-ToBERT outperforms standalone LEGAL-BERT models, offering improved accuracy in rhetorical role classification.

Parallel advancements in knowledge integration explore diverse approaches—ranging from natural language transformation of knowledge graph triples (Agarwal et al., 2021)¹² and direct embedding of structured triples into T5 architectures (Moiseev et al., 2022)⁸, to adapter-based domain alignment strategies (Diao et al., 2023)⁹. Additionally, task-specific research efforts—such as precedent case retrieval using event extraction (Joshi et al., 2023)¹⁰ and rhetorical role prediction via transformers (Bhattacharya et al., Belfathi et al., n.d.)²—have deepened contextual understanding of legal discourse.

In their work “MILDSum: A Novel Benchmark Dataset for Multilingual Summarization of Indian Legal Case Judgments”, Debtanu Datta et al. from IIT Kharagpur introduced the first multilingual dataset focused on Indian legal judgments. The dataset includes 3,122 case judgments from the Supreme Court and various High Courts, along with professionally written summaries in both English and Hindi. This resource aims to support the development and evaluation of cross-lingual summarization models, helping make complex legal texts more accessible to people across different language backgrounds in India. The study highlights the need for better summarization models in the legal domain and establishes a foundation for future work in multilingual legal AI.

Several recent studies have explored the intersection of artificial intelligence (AI) and the Indian judicial system. A notable one, again, is "Analyzing the Effectiveness of AI in the Indian Judicial System"¹², investigating how AI can improve operational efficiency in the judiciary. Another significant contribution is the BetterCall project¹⁴, which presents an AI-based legal assistant utilizing NLP for legal query processing, precedent retrieval, and automated drafting.

Further advancing the field, Human-Centered AI for Indian Legal Text Analytics by Sudipto Ghosh et al.¹⁵ proposes interpretable AI models incorporating human oversight to ensure fairness, traceability, and domain-specific accountability. Similarly, MILDSum (arXiv:2310.18600)¹ introduces a multilingual legal summarization dataset aimed at enhancing accessibility. Another critical work, Artificial Intelligence in Judicial Adjudication – Bias Classification¹⁶, presents a bias audit framework using LSTM and attention-based models to detect semantic bias in judicial texts.

The study InLegalLLaMA (arXiv.org)¹³ by Sudipto Ghosh et al. contributes by fine-tuning a large language model specifically on Indian legal data. The practical application of conversational AI is illustrated through SWATI AI¹⁷, a chatbot designed to assist domestic abuse victims using NLP-trained architectures. Addressing the need for language inclusivity, MILPaC¹⁴ proposes a legal translation benchmark covering ten Indian languages. The theoretical underpinnings of AI in judicial decision-making are reviewed by Dr. Talat Anjum¹, who explores ethical and institutional frameworks.

Chatlaw is an innovative AI-driven legal assistant that integrates a Mixture-of-Experts (MoE) large language model with a multi-agent system to enhance the reliability and accuracy of legal consulting services. By constructing a high-quality legal dataset through artificial screening and incorporating knowledge graphs, Chatlaw effectively addresses the hallucination issues commonly associated with large language models. The system emulates real law firm workflows by implementing Standardized Operating Procedures (SOPs) involving four independent intelligent agents responsible for information gathering, material research, legal advice, and consultation report writing. This structured approach ensures efficient and accurate information processing, significantly improving the quality of legal services and client satisfaction. Notably, Chatlaw outperforms GPT-4 in both the Lawbench and the Unified Qualification Exam for Legal Professionals, demonstrating its robust capability for legal consultation (Cui et al., 2023)⁶.

Jain et al. (2021)¹¹ present an extractive summarization method tailored for Indian legal judgment documents, addressing the challenges posed by their length and unstructured nature. Their approach employs domain-specific contextual embeddings, derived from pre-trained legal models, to represent individual sentences, which are then classified using a multilayer perceptron (MLP) to assess their summary worthiness. This method was evaluated in the AILA 2021 shared tasks at FIRE, where it achieved notable success—securing the 2nd position in Task 2(a) and the highest ROUGE-F1 scores across all metrics in Task 2(b). While effective for summarization, the model's performance in rhetorical role labeling (Task 1) was less competitive. The work highlights the potential of combining domain-adapted embeddings with MLP classifiers to automate the summarization of extensive legal documents, ultimately aiding legal professionals and end users by saving time and improving document accessibility.

Kalamkar et al. (2022)¹² address the unique challenges of Named Entity Recognition (NER) in Indian legal texts, which often contain domain-specific entities not typically found in general NER tasks. They introduce a comprehensive corpus comprising 14,444 annotated sentences and 2,126 judgment preambles from Indian court judgments, encompassing 46,545 labeled entities across 14 legal-specific categories such as PETITIONER, RESPONDENT, STATUTE, PROVISION, and PRECEDENT. To facilitate accurate extraction of these entities, the authors develop a transformer-based baseline model complemented by rule-based post-processing techniques that consider document-level context and coreference resolution. This work not only provides a valuable resource for legal NLP applications in

India but also lays the groundwork for advanced tasks like relation extraction and knowledge graph construction. The annotated dataset and baseline models are publicly accessible via GitHub, promoting further research and development in the field.

Bindal et al. (2023)¹⁸ address the challenges of summarizing landmark judgments in the Indian legal system, where the common law framework relies heavily on precedent and citation. Given the voluminous and complex nature of court rulings, especially in India where structures vary significantly compared to jurisdictions like the US, UK, or Canada, manual summarization proves to be both laborious and domain-intensive. To overcome these limitations, the authors propose CB-JSumm, an unsupervised extractive summarization algorithm that leverages contextual references from citing judgments to generate summaries of landmark cases. Their method emphasizes the importance of citations in understanding the legal and jurisprudential value of judgments, using them as a guide to highlight key arguments and issues. Evaluated on two datasets curated from Indian court rulings, the approach shows promising results, reinforcing the relevance of citation-based techniques in the development of effective legal summarization tools. This work contributes to ongoing efforts in legal NLP by providing a scalable alternative to manual summarization, tailored specifically for the Indian legal domain.

The paper titled “*ALDA: Cognitive Assistant for Legal Document Analytics*” by Karuna P. Joshi et al.²⁰ introduces a cognitive assistant designed to address the growing challenge of analyzing vast corpora of digitized legal documents, including contracts, compliance guidelines, regulations, and court case archives. As legal texts increasingly shift to digital formats—spanning public, private, and judicial sectors—there is a pressing need to automate the labor-intensive and manual processes currently used for monitoring and ensuring regulatory compliance. ALDA (Analytical Legal Document Assistant) leverages cognitive technologies such as semantic web, text mining, machine learning, and graph analytics to transform unstructured legal documents into machine-readable, graph-based data structures. The system aims to enable users—ranging from legal professionals to government officials—to query legal texts using natural language and receive structured, inferential responses. By identifying applicable laws, statutes, and case precedents in response to user queries, ALDA functions as a legal question-answering system that significantly enhances the efficiency of legal decision-making. It has the potential to offer real-time compliance insights, especially in domains like IT outsourcing and cloud computing, where legal adherence is critical. This approach not only reduces the time and cost associated with legal document review but also facilitates proactive compliance monitoring and improved access to legal knowledge across sectors.

Lastly, the 2023 IEEE paper Chatbot Frameworks in Legal Applications compares rule-based, retrieval-based, and generative AI systems for automated legal support. Collectively, these studies form a comprehensive foundation for future research in building intelligent, ethical, and inclusive AI systems tailored for the Indian legal ecosystem.

3. Comparative Analysis

Each paper highlighted in this survey focuses on a unique aspect of AI in the judiciary, offering a multidimensional approach to legal innovation. From document summarization in **MILDSum** and legal assistance tools such as **BetterCall** to domain-specific models like **InLegalLLaMA**, these works showcase AI's potential to reshape legal practice in India.. In contrast, the **Bias Detection** study is particularly concerned with identifying and mitigating systemic prejudices in judicial language. Collectively, these studies affirm the need for domain-specific datasets, ethically designed systems, and inclusive language capabilities to ensure effective AI deployment in the Indian legal ecosystem.

4. Limitations from the current Suvery

While the integration of Artificial Intelligence into legal systems presents transformative potential, current research and deployment efforts are constrained by significant limitations. Most existing models rely on datasets drawn primarily from the Supreme Court and a select few High Courts, excluding district-level judgments and not on complete files including First Information Reports (FIRs), and procedural records, thereby restricting representational depth and generalizability. Moreover, legal documents are largely unstructured and not machine-readable, posing additional challenges to automated processing. A critical shortcoming is the lack of multilingual support, as lower courts often operate in regional languages such as Hindi, Telugu, Tamil, and Bengali. This language gap impedes the usability and inclusivity of AI tools at grassroots levels. Furthermore, existing models such as InLegalLLaMA and BetterCall lack explainability, operating as opaque black boxes with minimal transparency in their decision-making processes. Such opaqueness undermines trust and limits practical adoption in sensitive legal contexts. Ethical and legal risks also abound, stemming from inherent biases in training data and the generation of potentially misleading or hallucinated summaries. The absence of fairness-aware training mechanisms and audit trails exacerbates these concerns. Compounding these issues is the lack of real-world deployment; most systems are evaluated solely in academic or controlled environments without validation in actual courtrooms, tribunals, or legal aid settings. Minimal collaboration with legal stakeholders—including judges, advocates, and regulatory bodies—further isolates AI development from real-world legal workflows, reducing the practical relevance of these tools. Evaluation frameworks predominantly rely on generic natural language processing (NLP) metrics such as BLEU and ROUGE, which fail to measure legal fidelity, statutory alignment, or precedential soundness. Additionally, research in this domain disproportionately emphasizes civil and constitutional law, leaving critical areas such as family law, criminal justice, labor disputes, and public grievance redressal underrepresented. Lastly, the absence of a comprehensive regulatory or policy framework leaves questions of accountability, liability, and ethical deployment unaddressed, posing significant barriers to trustworthy and equitable AI integration in the judiciary.

5. Proposed Novel Research Area:

The proposed improvements and recommendations outlined above represent a novel and underexplored research area within the Indian legal and AI ecosystem. Despite growing interest in legal tech, comprehensive research integrating standardized legal datasets, multilingual legal models, explainable AI, and bias mitigation tailored to India's diverse judicial landscape remains largely untouched. Critical gaps persist in deploying AI tools in real-world legal settings like Lok Adalats and consumer forums, as well as in developing domain-specific evaluation metrics that go beyond generic NLP benchmarks. Furthermore, underrepresented legal areas—such as family law and public grievances—lack focused AI intervention. Regulatory frameworks to govern AI deployment in legal contexts are also nascent. Hence, this area offers significant scope for pioneering research that is contextually relevant, ethically grounded, and socially impactful. The full file reading of end to end file is also something which is persisting

6. Conclusion

The integration of Generative AI into the Indian legal ecosystem holds transformative potential to enhance efficiency, accessibility, and transparency in justice delivery. However, the existing body of research reveals critical gaps that hinder real-world deployment and adoption. Key limitations include narrow dataset representation, lack of multilingual support, absence of explainability, ethical concerns, and insufficient legal stakeholder involvement. Furthermore, current evaluation practices fail to capture legal reasoning or statutory accuracy, and there is a notable lack of regulatory frameworks guiding AI use in judicial contexts.

To bridge these gaps, future research must prioritize the development of standardized, multilingual, and explainable legal AI systems, designed in collaboration with legal professionals and tailored for diverse legal domains. Establishing robust legal evaluation metrics, piloting AI deployment in quasi-judicial bodies, and creating a national policy framework are essential steps toward responsible AI integration. By addressing these challenges systematically, India can pioneer an inclusive, transparent, and technology-driven justice system empowered by Generative AI.

7. References

1. Anjum, T. (2024). *Exploring the role and impact of artificial intelligence in the judiciary*. International Journal of Dynamic Educational Research Society, 5(1), 14. (Assistant Professor, Institute of Legal Studies & Research, Mangalayatan University, Beswan, Aligarh)
2. Belfathi, A., Hernandez, N., & Monceaux, L. (2023, June). *Enhancing pre-trained language models with sentence position embeddings for rhetorical roles recognition in legal opinions*. In *Proceedings of the 6th Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL)*, co-located with the 19th ICAIL, Braga, Portugal. Universidad do Minho. <https://hal.science/hal-0413812>

3. Marino, G., Licari, D., Bushipaka, P., Comandé, G., & Cucinotta, T. (2023, June). *Automatic rhetorical roles classification for legal documents using LEGAL-TransformerOverBERT*. In *ASAIL 2023, ICAIL*, Braga, Portugal (pp. 28–36). CEUR Workshop Proceedings, Vol. 3441. <https://ceur-ws.org/Vol-3441/paper4.pdf>
4. Gupta, A., Furniturewala, S., Kumari, V., & Sharma, Y. (2023, May). *Rhetorical role labeling of legal documents using transformers and graph neural networks*. arXiv. <https://doi.org/10.48550/arXiv.2305.04100>
5. Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). *LEGAL-BERT: The muppets straight out of law school*. arXiv. <https://arxiv.org/abs/2010.02559>
6. Cui, J., Ning, M., Li, Z., Chen, B., Yan, Y., Li, H., Ling, B., Tian, Y., & Yuan, L. (2023). *ChatLaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model*. arXiv. <https://arxiv.org/abs/2306.16092>
7. Datta, D., Soni, S., Mukherjee, R., & Ghosh, S. (2023). *MILDSum: A novel benchmark dataset for multilingual summarization of Indian legal case judgments*. arXiv. <https://arxiv.org/abs/2310.18600>
8. Fei, Z., Shen, X., Zhu, D., Zhou, F., Han, Z., Zhang, S., Chen, K., Shen, Z., & Ge, J. (2023). *LawBench: Benchmarking legal knowledge of large language models*. arXiv. <https://arxiv.org/abs/2309.16289>
9. Ghosh, S., Verma, D., Ganesan, B., Bindal, P., Kumar, V., & Bhatnagar, V. (2024, March 16). *Human centered AI for Indian legal text analytics [Preprint]*. arXiv. <https://doi.org/10.48550/arXiv.2403.10944>
10. Guha, N., Nyarko, J., Ho, D. E., Ré, C., Chilton, A., Narayana, A., Chohlas-Wood, A., Peters, A., Waldon, B., Rockmore, D. N., Zambrano, D., Talisman, D., Hoque, E., Surani, F., Fagan, F., Sarfaty, G., Dickinson, G. M., Porat, H., Hegland, J., ... (2023). *LegalBench: A collaboratively built benchmark for measuring legal reasoning in large language models*. arXiv. <https://arxiv.org/abs/2308.11462>
11. Marrivagu, S., & Rao, A. S. L. (2024). *Artificial intelligence for legal chatbot*. Journal of Emerging Science. <https://jespublication.com/uploads/2024-V15I9025.pdf>
12. Kalamkar, P., Agarwal, A., Tiwari, A., Gupta, S., Karn, S., & Raghavan, V. (2022). *Named entity recognition in Indian court judgments*. In *Proceedings of the Natural Legal Language Processing Workshop 2022* (pp. 184–193). Association for Computational Linguistics. <https://aclanthology.org/2022.nllp-1.15>
13. Nigam, S. K., Patnaik, B. D., Mishra, S., Shallum, N., Ghosh, K., & Bhattacharya, A. (2024). *NyayaAnumana & INLegalLlama: The largest Indian legal judgment prediction dataset and specialized language model for enhanced decision analysis*. arXiv. <https://arxiv.org/abs/2412.08385>

14. Patnaik, B. D., Nigam, S. K., Mishra, S., Shallum, N., Ghosh, K., & Bhattacharya, A. (2023). *MILPaC: A novel benchmark for evaluating translation of legal documents in Indian languages*. arXiv. <https://arxiv.org/pdf/2310.09765>
15. Pardhi, B., Koli, S., Khanzode, V., & Raut, A. S. (2024). *LEGALBOT – AI law advisor chatbot*. International Journal of Novel Research and Development, 9(4). <https://www.ijnrd.org/papers/IJNRD2404730.pdf>
16. ScienceDirect. (2024). *Bias classification in legal systems*. Heliyon. <https://doi.org/10.1016/j.heliyon.2024.e062157>
17. Vidhi Centre for Legal Policy. (2020). *JALDI Report: Justice, Access, and Lowering Delays in India*. <https://vidhilegalpolicy.in/research/jaldi-report-2020/>
18. Bindal, P., Kumar, V., Bhatnagar, V., Sirohi, P., & Siwal, A. (2023). *Citation-based summarization of landmark judgments*. In *Proceedings of the 20th International Conference on Natural Language Processing (ICON 2023)* (pp. 447–457). Association for Computational Linguistics. <https://aclanthology.org/2023.icon-1.56>
19. Jain, D., Dutta Borah, M., & Biswas, A. (2021). *Summarization of Indian legal judgement documents via ensembling of contextual embedding based MLP models*. In *FIRE 2021 – Working Notes Track* (pp. 553–561). CEUR-WS.org. <https://ceur-ws.org/Vol-3159/T2-6.pdf>
20. Joshi, K. P., Gupta, A., Mittal, S., Pearce, C., Joshi, A., & Finin, T. (2016, September). *ALDA: Cognitive assistant for legal document analytics*.
21. Awasekar, D. D., & Lobo, L. M. R. J. (2025, January). *SWATI AI: Advancing AI models from rule-based frameworks to NLP-driven prescriptive analytics to assist victims of domestic abuse*. In *Proceedings of the 1st International Conference on AIML-Applications for Engineering & Technology (ICAET)*, Pune, India.