

Analysing India's Judicial Backlog: A Data-Centric Approach to AI Integration in District Courts

Sanjay Subramanyan Sivakumar, Thanmay Mohana

Student, Deakin University, Burwood VIC - Australia

Date of Submission: 28-05-2025

Date of Acceptance: 08-06-2025

ABSTRACT: This paper investigates India's judicial backlog by analysing district-level pendency data, revealing acute caseload pressures, infrastructure gaps, and judge vacancies particularly in five states. Drawing from global case studies, it examines the role of Artificial Intelligence (AI) in enhancing judicial efficiency and proposes a locally hosted Retrieval-Augmented Generation (RAG) model tailored for Indian district courts. The study underscores the benefits of AI for faster case retrieval and workload management, while also cautioning that ethical safeguards must be in place to ensure transparency, accountability, and data sovereignty in judicial applications.

KEYWORDS: Judicial Pendency, Indian District Courts, Artificial Intelligence, Retrieval-Augmented Generation (RAG), Data Analysis, Ethical AI.

I. INTRODUCTION

India's judiciary is an essential pillar of its democratic framework, functioning independently from both the legislature and the executive. The country follows a bicameral parliamentary system, and its judicial structure is organised in a three-tier hierarchy comprising the Supreme Court at the apex, followed by High Courts at the state level, and District and Subordinate Courts at the grassroots level.

This research paper begins with a detailed analysis of the pendency landscape, highlighting patterns in civil and criminal caseloads, categorisation based on time duration, and disparities across states. It then addresses the major challenges currently faced by the judiciary. Drawing on international case studies, the paper explores how Artificial Intelligence (AI) has been implemented in other judicial systems to enhance efficiency and effectiveness.

The paper also evaluates India's initial AI initiatives and proposes the adoption of a Retrieval-Augmented Generation (RAG) model, hosted on secure local infrastructure, as a practical solution to support judges in document retrieval, case summarisation, and real-time decision-making.

II. DATA ANALYSIS

The dataset used for this analysis was obtained from the National Data and Analytics Platform (NDAP), a flagship initiative by NITI Aayog aimed at enhancing public access to government data. The analysis specifically draws from the "District and Taluk Court Cases" CSV file, which provides granular, district- and taluk-level insights into the number of pending court cases segmented by time intervals. (National Judicial Data Grid, 2024)

The dataset includes detailed classifications of criminal and civil cases, as well as information on the procedural stage at which the cases are pending (e.g., evidence, argument, judgment).

In total, the dataset comprises 2,206 rows and 27 columns, representing 59,562 data points, offering a comprehensive view of judicial pendency across India.

The analysis of case pendency in Indian district courts reveals a deeply entrenched backlog, particularly in the stages of evidence, argument, and judgment. As of the latest available data, over 30 million cases remain pending in these final stages. Alarming, approximately one-third of these cases (10 million) are criminal in nature and have been pending for more than 10 years, underscoring systemic delays in adjudicating serious offences.

The total number of pending cases across all courts stands at 91,144,772, of which 78,658,020 cases form the original pendency. The

remainder arises from appeals, applications, and execution proceedings. Among the additional categories, pending application cases account for approximately 6.63 million cases (57.95%),

followed by execution cases at 3 million (26.21%), and appeal cases at 1.81 million (15.84%) (see Figure 1).

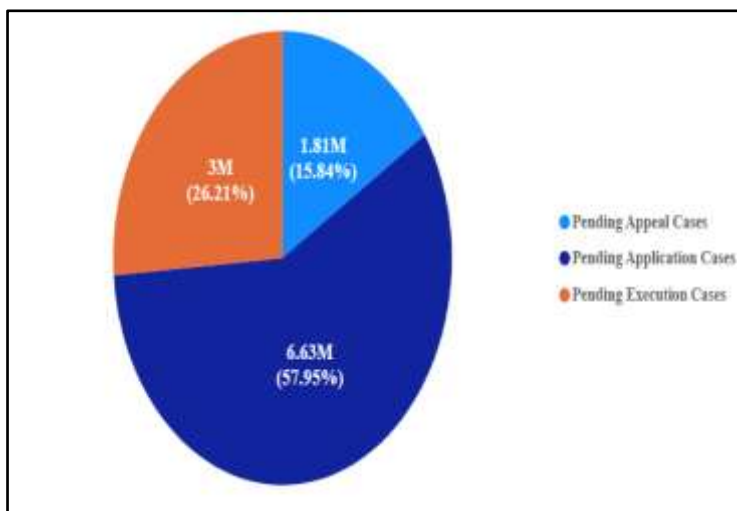


Figure 1:

Among the Indian states, a disproportionate burden of judicial pendency is concentrated in just five jurisdictions. The top five states with the highest number of pending cases are Uttar Pradesh, Maharashtra, Bihar, West Bengal, and Rajasthan, which together account for approximately 52.8 million pending cases. This figure represents nearly 58% of the total pending

caseload across all district courts nationwide, highlighting a significant regional concentration of judicial workload.

The pending cases are broadly categorized into two primary types: criminal cases and civil cases. A closer examination of the top contributors in each category reveals distinct patterns.

Table 1: Categorical Case Pending Count

Pending Cases (Top 5 Contributors)			
Criminal Cases		Civil Cases	
State	Case Count (in Millions)	State	Case Count (in Millions)
Uttar Pradesh	10.1	Uttar Pradesh	1.86
Maharashtra	3.7	Maharashtra	1.68
Bihar	3.0	Karnataka	0.99
West Bengal	2.5	Tamil Nadu	0.74
Rajasthan	1.8	West Bengal	0.62

In terms of criminal case pendency, Uttar Pradesh leads by a wide margin, with 10.1 million cases, followed by Maharashtra (3.7 million), Bihar (3.0 million), West Bengal (2.5 million), and Rajasthan (1.8 million).

For civil cases, the distribution shifts slightly. Uttar Pradesh emerges as the top contributor with 1.86 million pending civil cases, followed by Maharashtra (1.68 million). Other significant contributors include Karnataka (0.99

million), Tamil Nadu (0.74 million), and West Bengal (0.62 million).

To better understand the extent of delays in the Indian judiciary, the pending cases have been categorized into seven distinct time bins: 0 to 1 year, 1 to 3 years, 3 to 5 years, 5 to 10 years, 10 to 20 years, 20 to 30 years, and over 30 years. This distribution provides crucial insight into the ageing profile of judicial backlogs across states and case types.

Criminal Case Pendency

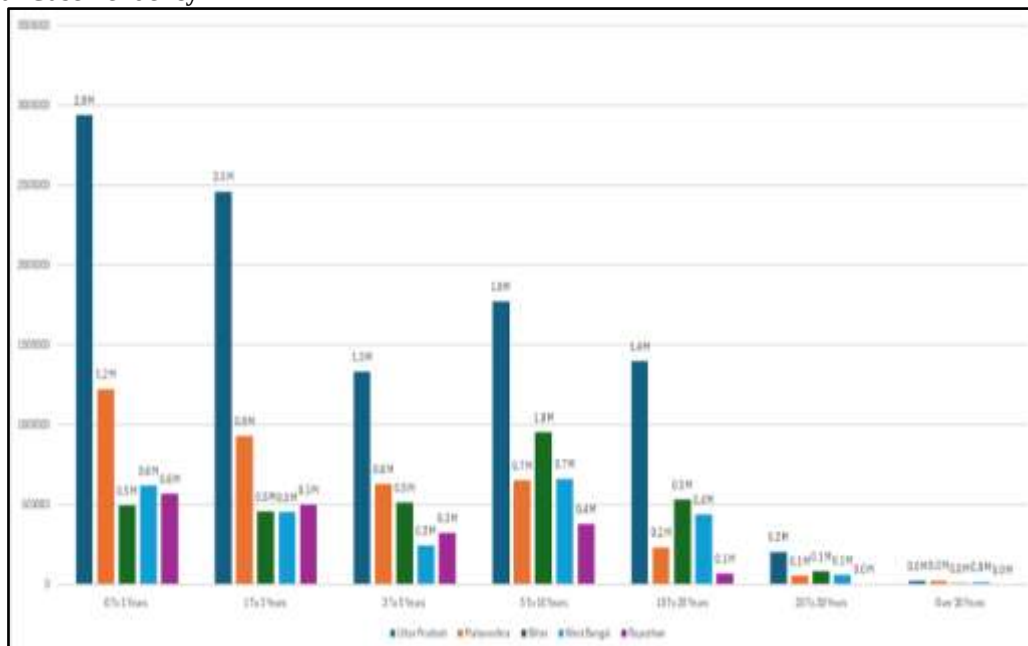


Figure 2:

As illustrated in the first chart, Uttar Pradesh consistently leads in the number of pending criminal cases across nearly all time brackets, reflecting the state's persistently high criminal case volume. Specifically, Uttar Pradesh accounts for:

- 2.9 million cases pending for 0 to 1 year.
- 2.5 million cases for 1 to 3 years.
- 1.3 million cases for 3 to 5 years.
- 1.8 million cases for 5 to 10 years.

This trend continues until the over 30 years bracket, where Maharashtra marginally surpasses Uttar Pradesh, indicating Maharashtra's deeper historical accumulation of unresolved criminal cases.

Importantly, approximately 10.4% of all criminal cases are pending for more than 10 years, which includes the 10 to 20, 20 to 30, and over 30 years categories. Within this long-pending segment, a staggering 94% of cases which are pending for more than 30 years come from just four states: West Bengal, Uttar Pradesh, Maharashtra, and Bihar.

Civil Case Pendency

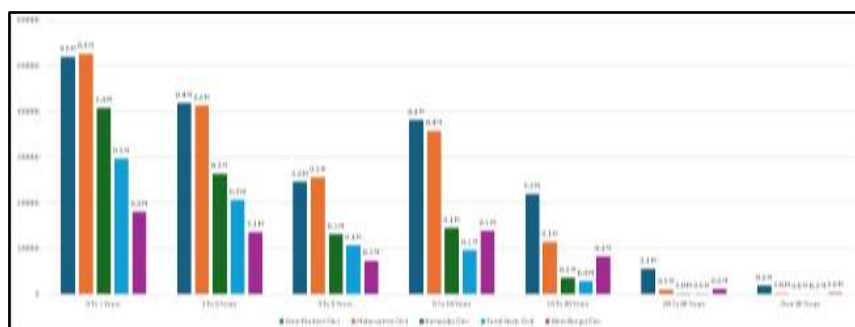


Figure 3:

In the second chart, a similar bin-wise distribution is observed for civil cases. The data indicates a more competitive pattern between Uttar Pradesh and Maharashtra, which exhibit similar volumes of civil pendency in the initial time brackets.

However, Uttar Pradesh significantly surpasses Maharashtra in the 10 to 20 years, 20 to 30 years and Over 30 years categories, suggesting a more pronounced accumulation of long-pending civil litigation in that state.

Overall, 7.5% of civil cases are pending for more than 10 years, again inclusive of those exceeding three decades. Consistent with criminal pendency, the civil cases which are pending for over 30 years, among them 93.4% of the cases comes from four states - Uttar Pradesh, Maharashtra, West Bengal and Bihar.

III. CHALLENGES

The Indian judicial system continues to face a series of structural and operational challenges, particularly at the district court level where the majority of litigation is concentrated. One of the most pressing issues is the sheer volume of pending cases, which has placed an unsustainable burden on judicial resources and led to systemic inefficiencies.

At the lower tiers of judicial performance, states such as Bihar, Rajasthan, Jharkhand, Uttarakhand, Uttar Pradesh, and West Bengal have consistently remained at the bottom of the performance index with only minor positional changes. For instance, Uttar Pradesh, which previously ranked lowest, has moved up slightly, exchanging positions with West Bengal. These fluctuations reflect the evolving impact of governance reforms, state-level administrative focus, and socio-economic conditions influencing judicial outcomes.

Despite the Law Commission of India's 1987 recommendation of 50 judges per million population, this benchmark remains unmet across all states and union territories. The current sanctioned strength of judges at the district level stands at 25,771, which translates to an average of 18 judges per one million (10 lakh) population. However, only 20,478 judges are currently in position to serve India's 1.4 billion population, resulting in an operational average of 15 judges per million people.

This shortage severely compromises judicial capacity. On average, each district judge is handling 4,450 cases. The situation in Uttar Pradesh is particularly critical, where the sanctioned strength of 3,700 judges must manage a pendency of 24 million cases. This equates to approximately 6,480 cases per judge, assuming full staffing. However, as noted in the India Justice Report 2025, the judicial vacancy rate in Uttar Pradesh is estimated at 25% to 35%, translating to a shortfall of 925 to 1,295 judges. As a result, the effective caseload per serving judge is estimated to range between 8,650 and 9,980 cases, substantially higher than the national average.

Moreover, the issue is compounded by regional concentration, with five states accounting for over 58% of all pending cases in district courts. This uneven distribution exacerbates disparities in judicial access and case resolution efficiency. (Indian Justice Report, 2025)

The crisis is further deepened by chronic underinvestment in the judicial infrastructure. In 2022–23, no state allocated more than 1% of its total budget to the judiciary. This persistent underfunding affects not only physical infrastructure but also staffing, digital systems, and case management capabilities. (Indian Justice Report, 2025)

Long-pending cases present additional administrative complexities. Cases pending for over a decade require extensive documentation screening, which consumes disproportionate judicial time and diverts attention from newer filings. Notably, 61% of civil cases and 57% of criminal cases have been filed within the past three years, yet must compete for attention with legacy cases that often involve voluminous records and procedural backlogs. (Indian Justice Report, 2025)

The continued presence of high vacancy rates, low case clearance ratios, and mounting arrears affirms that the formal court system remains under substantial pressure. Without urgent structural reforms and focused investment in judicial capacity, the system risks becoming further overwhelmed, compromising access to timely and equitable justice.

IV. AI SOLUTION

Artificial Intelligence (AI) has become a transformative force across multiple industries and sectors, and the judicial domain is no exception. As judicial systems across the world face mounting case backlogs, administrative inefficiencies, and resource constraints, AI is increasingly being explored as a tool to augment human decision-making, automate routine processes, and improve access to justice. This section presents a series of international case studies that illustrate how different countries, including India, have integrated AI into their judicial frameworks. It also outlines the observed benefits of these implementations and highlights relevant insights that may inform AI integration within Indian district courts. (Bennett Moses et al., 2024)

Case Study Review: Global Implementations of AI in the Judiciary

India has initiated multiple AI-based solutions within its judiciary. The Supreme Court Portal for Assistance in Court Efficiency (SUPACE) is designed to assist judges by summarising case files, extracting relevant facts, and retrieving legal precedents. Although it does not participate in decision-making, it significantly reduces the time and cognitive load involved in legal research. Another innovation is the Supreme Court VidhikAnuvaad Software (SUVAS), which translates judgments from English into vernacular Indian languages. This tool enhances linguistic inclusivity and accessibility across India's multilingual population. (Rashmi Singh Rana et al., 2023)

In Canada, the British Columbia Civil Resolution Tribunal (CRT) represents the world's

first fully online tribunal. It uses AI-guided pathways and expert systems to help citizens resolve small claims and housing disputes independently, minimising the need for courtroom proceedings. Estonia has taken a unique approach by piloting an AI judge to handle low-value administrative claims. Though the system operates within a narrow domain and remains under human supervision, it reflects Estonia's commitment to digital governance. (Bennett Moses et al., 2024)

China has developed Smart Courts and Internet Courts in major cities such as Hangzhou, Beijing, and Guangzhou. These courts incorporate AI technologies for sentencing recommendations, automated document drafting, and blockchain-based evidence authentication. Brazil's Supreme Federal Court uses Victor AI to identify cases involving constitutional matters, allowing judges to focus on high-priority cases. (Bennett Moses et al., 2024)

Germany has partnered with IBM to implement AI systems that assist in the classification, structuring, and summarisation of complex legal documents, particularly in commercial and civil law. Similarly, Colombia has deployed IBM Watson-based solutions that support natural language processing for judicial documentation and rule extraction, contributing to faster case turnaround. In the United States, tools like the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) are used for risk assessment in bail and parole decisions, although concerns have been raised regarding algorithmic bias and transparency. In the United Kingdom, AI is primarily used in the private legal sector for predictive analytics and litigation forecasting, with limited adoption in public courts. (Schindler, 2025)

Benefits and Lessons for Indian District Courts

The implementation of AI across these jurisdictions has yielded several practical benefits that India can draw upon. Firstly, AI systems have demonstrated a clear capacity to reduce clerical burden by automating repetitive tasks such as document review, case classification, and information retrieval. This has allowed judicial officers to focus more on substantive decision-making, thereby improving overall productivity.

Secondly, AI translation tools like SUVAS in India and similar language-access features in other countries have made legal processes more inclusive and citizen-friendly, especially for non-English-speaking populations. In the Indian context, where judicial accessibility

remains a challenge, such solutions hold significant promise.

Thirdly, countries like Canada and China illustrate how AI can support alternative dispute resolution and enhance the speed and efficiency of handling low-complexity cases. The Civil Resolution Tribunal in Canada, for instance, provides a template for online-first, AI-assisted resolution of disputes that could be replicated in India's overburdened district courts.

Lastly, by streamlining preliminary processes such as triage, case routing, and legal research, AI tools have shown potential in reducing delays, managing judicial workload, and enhancing consistency in judgments. These advantages suggest that AI, when ethically deployed, could play a meaningful role in alleviating the severe pendency crisis in India's lower judiciary.

V. RECOMMENDATION

The integration of Artificial Intelligence (AI) into the judicial ecosystem must be grounded in principles of sovereignty, transparency, efficiency, and data security. Based on the evaluation of current infrastructural constraints and comparative international practices, the following recommendation is proposed: the deployment of a Retrieval-Augmented Generation (RAG) model through a locally hosted infrastructure, rather than a cloud-based platform.

A locally hosted RAG model would ensure that the entire AI ecosystem, including data processing, knowledge base storage, and retrieval pipelines, remains within the territorial boundaries of the nation. This approach would significantly strengthen national data sovereignty and reduce vulnerabilities associated with third-party cloud services. The centralized knowledge base would function both as a secure repository for judicial documents and as the training and reference source for the RAG-based AI system. By eliminating dependencies on external cloud providers, the judiciary can minimize risks of data leakage, cyber threats, and unauthorized access to sensitive case materials.

Functionally, a RAG model would enhance judicial productivity by enabling instant retrieval of case histories, precedents, and document summaries. For example, in appellate cases that have undergone numerous hearings over two to three years, it can be time-consuming for a judge to manually review 10 to 20 hearing transcripts and related submissions. A RAG-based AI assistant could retrieve and summarise these materials instantly and also suggest similar resolved cases for benchmarking judicial decisions.

This utility is particularly valuable in pre-hearing preparations, offering judges a comprehensive and efficient review mechanism.

Furthermore, in a live courtroom setting, the RAG model can support real-time decision-making. For instance, if the opposing counsel repeatedly requests time extensions for evidence collection, the AI system can retrieve historical data regarding similar requests within the same case and provide a pattern-based report to the presiding judge. Such functionality could enable data-driven judicial discretion, reducing procedural delays and ensuring consistency in rulings.

Key Features and Safeguards

- **Data Sovereignty:** Hosting the RAG model on domestic servers ensures that all case-related data remains within national jurisdiction, reducing exposure to transnational data privacy risks.
- **Explainability and Transparency:** The RAG model retrieves only from the knowledge base curated and sanctioned by the judiciary. This ensures outputs are based on fact and provides transparency to judicial decisions.
- **Institutional Control:** All operational workflows, including access rights, system updates, and data modifications, would remain under the exclusive purview of the judicial administration. This preserves the independence of the judiciary.
- **Minimized Security Risk:** The local infrastructure eliminates reliance on cloud encryption protocols and reduces the need for masking, tokenization, or other data anonymization techniques. This also simplifies compliance with domestic data protection laws.
- **Bias Mitigation:** As the RAG model is trained solely on authenticated and verified legal texts, it avoids the propagation of societal, political, or ideological biases often observed in general-purpose AI models.

This locally hosted RAG framework not only enhances judicial efficiency but also upholds the core principles of data protection, institutional integrity, and public trust. As India explores AI-enabled reforms for its overburdened district court system, such a model provides a scalable and secure foundation for digital transformation.

VI. CONCLUSION

This research paper highlights the urgent need for judicial reform in India. Current data shows more than 91 million pending cases, with

one-third of long-pending criminal cases extending beyond ten years. A significant portion of this backlog is concentrated in just five states. This issue is further worsened by judicial vacancies, inadequate infrastructure, and low budget allocations, all of which impact the timely delivery of justice. The gap between the number of sanctioned and serving judges, along with increasing caseloads, places unsustainable pressure on the system.

Artificial Intelligence, particularly Retrieval-Augmented Generation (RAG) models hosted on secure local servers, offers a practical solution to these challenges. These tools can assist judges by quickly retrieving case histories, identifying similar past rulings, and improving overall decision-making. International case studies show the potential of AI to improve efficiency, reduce delays, and expand access to justice.

However, the benefits of AI must be balanced with caution. Its implementation should follow a responsible approach that includes strong ethical oversight. Ensuring data privacy, transparency, and institutional control is essential to maintaining public trust and judicial independence.

In conclusion, AI has the potential to transform India's district courts and reduce case backlogs. But to achieve this, its use must be carefully planned, ethically governed, and aligned with the principles of justice, fairness, and accountability.

REFERENCES

- [1]. Bennett Moses, L., Legg, M., Silove, J., Zalnieriute, M., & Mahmood, S. (2024, August 20). AI Decision-Making and the Courts: a guide for Judges, Tribunal Members and Court Administrators (2023 update) - Australasian Institute of Judicial Administration. Australasian Institute of Judicial Administration. <https://aija.org.au/publications/ai-decision-making-and-the-courts-a-guide-for-judges-tribunal-members-and-court-administrators-2023-update/>
- [2]. Indian Justice Report. (2025, April). India Justice Report 2025. indiajusticereport.org. <https://indiajusticereport.org/>
- [3]. National Judicial Data Grid. (2024, July 10). National Data and Analytics Platform. [Ndap.niti.gov.in](https://ndap.niti.gov.in/). <https://ndap.niti.gov.in/>
- [4]. Rashmi Singh Rana, Singh, S., Aggarwal, M., & Badoni, M. (2023). Unveiling the Future: Exploring AI Applications in the Indian Judicial System. 2023 11th International Conference on Intelligent Systems and Embedded Design (ISED). <https://doi.org/10.1109/ised59382.2023.10444600>
- [5]. Schindler, E. (2025). Judicial systems are turning to AI to help manage vast quantities of data and expedite case resolution. [Ibm.com. https://www.ibm.com/case-studies/blog/judicial-systems-are-turning-to-ai-to-help-manage-its-vast-quantities-of-data-and-expedite-case-resolution](https://www.ibm.com/case-studies/blog/judicial-systems-are-turning-to-ai-to-help-manage-its-vast-quantities-of-data-and-expedite-case-resolution)