REGULATING ARTIFICIAL INTELLIGENCE IN COURTS: DESIGNING A LEGAL FRAMEWORK FOR JUDICIAL AI IN INDIA

Mishika Bhargava, Research Scholar, Raffles University

Dr. Pranay Kumar Aditya, (Senior Associate Professor & Dean School of Law), Raffles University

ABSTRACT

The integration of Artificial Intelligence (AI) into the Indian judicial system represents a paradigm shift from traditional digitisation to "intelligent" governance. This research paper explores the design of a comprehensive legal framework for Judicial AI in India, particularly in light of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, and the e-Courts Project Phase III.

While AI tools like SUPACE (Supreme Court Portal for Assistance in Court Efficiency) and SUVAS (translation software) have enhanced administrative efficiency and legal research, their deployment raises critical concerns regarding algorithmic bias, "black-box" opacity, and the potential erosion of judicial reasoning. This paper evaluates the current "assistive-only" mandate of the Supreme Court, which ensures that AI augments but does not replace human adjudication. It proposes a multi-tiered regulatory architecture—rooted in the seven "sutras" of trust, fairness, and accountability—to govern AI procurement, data privacy under the Digital Personal Data Protection Act (DPDPA), and the ethical use of generative AI in courtrooms.

The study further discusses the necessity of a techno-legal framework that mandates "explainability by design" and establishes a dedicated Judicial AI Oversight Committee. By balancing innovation with constitutional safeguards under Article 21, the proposed framework aims to mitigate risks of automated inequity while leveraging AI to unclog India's massive case backlogs.

Keywords: Judicial AI, e-Courts Phase III, Legal Framework, India, Algorithmic Accountability, SUPACE, Artificial Intelligence Regulation, Judicial Independence, Techno-legal Governance.

INTRODUCTION

The Indian Judiciary, often described as the backbone of the world's largest democracy, is currently standing at a digital crossroads. For decades, the system has grappled with the "Sisyphean labor" of a massive case pendency—currently exceeding 50 million cases across various tiers of the hierarchy. While the first two phases of the e-Courts Project successfully laid the hardware foundation through computerization and digital filing, the dawn of 2026 has ushered in a more profound transformation: the transition from Digitization to Algorithmic Adjudication.[1]

The integration of Artificial Intelligence (AI) into the legal fabric of India is no longer a matter of "if" but "how." With the rollout of e-Courts Project Phase III, the judiciary has begun deploying sophisticated "Intelligent Systems" designed to go beyond mere data storage. Tools such as SUPACE (Supreme Court Portal for Assistance in Court Efficiency) and SUVAS (Supreme Court Vidhik Anuvaad Software) represent the first wave of this revolution, aiming to augment human intelligence by automating legal research and linguistic translation. However, as AI evolves from a clerical assistant to a cognitive collaborator, it brings forth unprecedented legal and ethical quandaries.[2]

The primary tension lies between Efficiency and Equity. While AI promises to drastically reduce the "logjam" of litigation, its "black-box" nature threatens the foundational principles of Natural Justice. Can an algorithm truly understand the "spirit of the law"? How does a machine account for the socio-economic vulnerabilities of a litigant in a rural district court? In the absence of a dedicated regulatory framework, the unchecked use of AI risks codifying historical biases, eroding judicial transparency, and infringing upon the Right to a Fair Trial guaranteed under Article 21 of the Constitution.

This research paper, titled *"Regulating Artificial Intelligence in Courts: Designing a Legal Framework for Judicial AI in India,[3]"* argues that the current "light-touch" regulatory approach is insufficient for the unique complexities of the Indian courtroom. It posits that for AI to be a "force multiplier" for justice, it must be bound by a Sutra-based architecture—a set of seven

[1] Justice D.Y. Chandrachud, "Technology and the Changing Face of Indian Justice" 58 *JILI* 1 (2021).
[2] *Maneka Gandhi* v. *Union of India*, AIR 1978 SC 597.
[3] Sandra Wachter, Brent Mittelstadt, et.al., "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" 7 *International Data Privacy Law* 76-99 (2017)

ethical and legal pillars that prioritize human agency, accountability, and inclusivity.

By analyzing the Information Technology (Intermediary Guidelines) Amendment Rules, 2026, and drawing comparisons with global standards like the EU AI Act, this study proposes a tailored "Techno-Legal" roadmap[4]. The goal is to design a framework where technology serves the law, ensuring that the quest for a speedier judiciary does not result in a "mechanical" version of justice that lacks a constitutional soul.[5]

THE SEVEN SUTRAS: A REGULATORY BLUEPRINT FOR JUDICIAL AI IN INDIA

The integration of Artificial Intelligence (AI) into the Indian Judiciary is no longer a futuristic speculation but a present-day administrative reality. From the Supreme Court's translation engine SUVAS to the research-enhancing capabilities of SUPACE, the transition toward "Intelligent Courts" is well underway. However, as the e-Courts Project Phase III matures in 2026, the shift from administrative automation to substantive judicial assistance necessitates a robust ethical and legal anchor. The "Seven Sutras" of Judicial AI Regulation provide this framework, ensuring that the digitisation of justice does not come at the cost of constitutional morality.[6]

1. Manava-Centricity: The Human-in-the-Loop Imperative

The first and most vital sutra is *Manava-Centricity*, which mandates that AI remain a tool of assistance rather than a source of adjudication. In the Indian context, the "Right to a Fair Trial" under Article 21 inherently implies a right to a human judge. An algorithm, however sophisticated, lacks the "judicial conscience" required to weigh the nuances of equity, mercy, and societal impact. This principle ensures that while AI may summarize 10,000 pages of evidence in seconds, the final decision—the *ratio decidendi*—must be an act of human intellect. By keeping a "Human-in-the-Loop" (HITL), we prevent the rise of "Robo-Judges"

[4] Supreme Court of India, "Manual on SUVAS: AI-Assisted Translation Tool," available at: https://sci.gov.in (last visited Feb. 27, 2026).
[5] Ministry of Electronics & IT, *Report of the Committee of Experts on Non-Personal Data Governance Framework* (Government of India, 2020).
[6] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026: These rules updated the 2021 guidelines to include specific compliance standards for Large Language Models (LLMs) used in public service delivery.

that could potentially treat justice as a mere mathematical optimization problem.[7]

2. Paradarshita: The End of the "Black Box"

One of the most significant legal hurdles in deploying AI is the "Black Box" problem—the inability of developers to explain exactly how an AI reached a specific conclusion. The sutra of *Paradarshita* (Transparency) mandates Explainable AI (XAI). For a judicial decision to be valid, it must be reasoned. If a court uses an AI tool to calculate alimony or bail risks, the parties involved have a right to know the underlying logic. A legal framework for India must require that any AI used in courts provides a "traceability log," allowing lawyers to scrutinize the weights and parameters used by the algorithm, thereby upholding the principles of natural justice.[8]

3. Nishpakshata: Auditing for Algorithmic Bias

Algorithms are only as neutral as the data they are fed. In a diverse nation like India, there is a grave risk that historical biases related to caste, gender, religion, or socio-economic status could be codified into judicial software. *Nishpakshata* (Non-Discrimination) requires rigorous, mandatory pre-deployment auditing. If an AI trained on decades of arrest records inadvertently flags certain marginalized communities as "high risk," it perpetuates systemic injustice under the guise of technological objectivity. This sutra necessitates a "Bias Impact Assessment" for every tool before it enters the courtroom ecosystem.[9]

4. Suraksha: Data Sovereignty and Privacy

The Indian judiciary handles the most sensitive personal data of its citizens. The sutra of *Suraksha* (Security) aligns judicial AI with the Digital Personal Data Protection Act (DPDPA), 2023.[10] AI models often require massive datasets for training, but this cannot justify the unauthorized "scraping" of private litigation files. Judicial AI must be built on "privacy-by-design" principles, ensuring that data used for training is anonymized and that cloud

[7] e-Courts Project Phase III: A flagship initiative of the Government of India (2023–2027) with a budget of ₹7,210 crore, focusing on a "paperless" and "AI-integrated" judiciary.
[8] SUPACE (Supreme Court Portal for Assistance in Court Efficiency): Launched in 2021, it is a tool designed to collect and analyze data to assist judges in legal research.
[9] SUVAS (Supreme Court Vidhik Anuvaad Software): An AI-driven machine-assisted translation tool developed to translate judicial documents into regional languages.
[10] Digital Personal Data Protection Act (DPDPA), 2023: India's primary legislation governing the processing of digital personal data, emphasizing the rights of "Data Principals."

[11]infrastructures hosting these systems are sovereign and immune to foreign surveillance or commercial exploitation.[12]

THE PROBLEM OF "BLACK-BOX" JURISPRUDENCE: NAVIGATING ALGORITHMIC BIAS AND CONSTITUTIONAL SAFEGUARDS

The integration of Artificial Intelligence into the Indian judicial machinery is often marketed as a panacea for the chronic ailment of "justice delayed." However, beneath the veneer of neutral, data-driven efficiency lies a profound jurisprudential crisis: the "Black-Box" problem. In the context of the Indian legal system, where the Rule of Law is predicated on reasoned decisions and the protection of marginalized identities, the opaque nature of algorithms presents a direct challenge to the Basic Structure of the Constitution.

1. The Anatomy of the "Black Box"

A "Black-Box" system refers to an AI model whose internal logic is hidden from the user. In traditional adjudication, a judge's thought process is externalized through a written judgment, which is then subject to appeal and public scrutiny. In contrast, deep-learning algorithms—such as those potentially used for bail risk assessment or sentencing recommendations—process input data through millions of neurons and hidden layers. The output is delivered without a corresponding "reasoned order."[13]

From a legal standpoint, this opacity violates the **Principle of Natural Justice**, specifically the *Audi Alteram Partem* rule. If a litigant is denied bail based on an algorithmic score, yet neither the judge nor the litigant understands the specific weightage given to various parameters, the right to a fair hearing becomes illusory. The "Black Box" replaces judicial deliberation with a "mathematical fiat," creating a system where the machine's conclusion is beyond the reach of human cross-examination.[14]

---

[11] Explainable AI (XAI): A set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms.
[12] Department of Justice, *Detailed Project Report (DPR) for e-Courts Project Phase III* (Ministry of Law and Justice, 2023).
[13] Pasquale, Frank, *The Black Box Society: The Secret Algorithms That Control Money and Information* 12-14 (Harvard University Press, 2015).
[14] *S.P. Kapoor* v. *State of Himachal Pradesh*, AIR 1981 SC 2181 (Discussing the necessity of reasoned orders in administrative and judicial functions).

2. Algorithmic Bias in the Indian Context: Caste and Socio-Economic Variables

The most dangerous myth of AI is its supposed "objectivity." Algorithms are mirrors of the data they consume. In India, the historical data available to train judicial AI is deeply contaminated by systemic biases related to caste (*varna/jati*), religion, and socio-economic status.

For instance, if an AI is trained on historical arrest records or "history-sheeter" databases, it will inevitably reflect the over-policing of Scheduled Castes (SC), Scheduled Tribes (ST), and Denotified Tribes (DNTs). If the training data shows that a disproportionate number of people from a specific pin code or community have been arrested, the AI will label individuals from those backgrounds as "high-risk," regardless of their personal conduct.[15]This creates a "feedback loop of injustice" where historical marginalization is rebranded as "data-driven risk."

Furthermore, socio-economic variables such as "stability of employment" or "educational qualification"—often used by AI to predict recidivism—act as proxies for privilege. A landless laborer from a rural district in Chhattisgarh would inherently receive a lower "reliability score" than a salaried professional in Bengaluru, even if the nature of their alleged offense is identical. Without a legal mandate for **Algorithmic Neutrality**, Judicial AI risks becoming a tool for "high-tech casteism."[16]

3. The Conflict: Automated Efficiency vs. Article 21

The primary driver for Judicial AI in India is efficiency—specifically the goal of clearing the 50 million case backlog.[17] However, this quest for speed creates a structural conflict with Article 21 of the Constitution, which guarantees the Right to Life and Personal Liberty. The Supreme Court has repeatedly held that "procedure established by law" must be "just, fair, and reasonable."[18]

Automated efficiency often prioritizes "throughput" over "due process." For example, the use of AI in "Traffic Virtual Courts" for automated fine imposition might be acceptable for petty

[15] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* 84-90 (Crown, 2016).
[16] *Report of the Committee of Experts on Artificial Intelligence*, NITI Aayog (Government of India, June 2018).
[17] *Maneka Gandhi* v. *Union of India*, AIR 1978 SC 597.
[18] National Judicial Data Grid, "Statistics of Case Pendency," available at: https://njdg.ecourts.gov.in (last visited Feb. 27, 2026).

offenses. However, as these systems move toward more substantive areas—such as determining the "rarest of rare" status in capital offenses or assessing custodial requirements—the lack of "judicial empathy" becomes a constitutional infirmity.

The conflict is best illustrated by the tension between Speed and Substance. An AI can process a thousand bail applications in a minute, but it cannot "hear" the silent plea of a first-time offender who has no legal representation. If efficiency is pursued at the cost of individual liberty, the judiciary risks transforming from a "temple of justice" into a "factory of disposal."[19]The Bharatiya Sakshya Adhiniyam, 2023 now recognizes electronic records as primary evidence, but it does not yet provide a mechanism to challenge the *logic* of the AI that analyzes such evidence.[20]

The solution to "Black-Box" jurisprudence is not the rejection of technology, but its constitutionalization. India requires a Techno-Legal Framework that mandates "Explainability by Design" (XAI). This means that no AI tool should be deployed in the Indian judiciary unless it can provide a human-readable explanation of its output. Furthermore, mandatory "Bias Audits" must be conducted by independent bodies to ensure that algorithms do not perpetuate caste-based or religious prejudices.

Justice must not only be done but must also be *seen* to be done. In the age of AI, this means justice must be *calculated* in a way that is transparent, fair, and ultimately, human-centric.

DESIGNING THE LEGAL FRAMEWORK: INSTITUTIONAL AND PROCEDURAL SAFEGUARDS

The integration of Artificial Intelligence into the Indian Judiciary necessitates a shift from "policy-led" transformation to "statutory-backed" regulation. As we move through 2026, the success of the e-Courts Project Phase III—supported by a massive allocation of ₹7,210 crore—depends on a framework that balances the rapid deployment of "Intelligent Systems" with the immutable principles of the Indian Constitution. This section outlines the institutional and procedural pillars of the proposed framework, focusing on oversight, vendor accountability, and the evidentiary status of AI outputs.

[19] Justice D.Y. Chandrachud, "The Future of Justice: Technology and the Citizen," 60 *JILI* 15-20 (2023).
[20] The Bharatiya Sakshya Adhiniyam, 2023 (Act 45 of 2023), s. 63.

The cornerstone of judicial AI governance is the establishment of a specialized Judicial AI Oversight Committee (JAOC). Moving beyond the existing administrative e-Committee, the JAOC serves as a techno-legal regulator. Its composition must be multidisciplinary: chaired by a sitting Supreme Court Judge and including the Chief Justice of a High Court, the MeitY Secretary, and two independent "Ethicists-in-Residence" from academia. This structure ensures that technical advancement does not outpace judicial wisdom. The Committee's primary power includes the authority to "veto" any AI pilot project that fails to meet the Seven Sutras, ensuring that tools like predictive sentencing do not enter the courtroom without rigorous scrutiny.

Under the 2026 regulatory regime, private legal-tech vendors can no longer operate in a "regulatory vacuum." Every AI tool procured by the Department of Justice must undergo Mandatory Certification by the JAOC. These standards mandate "Explainability by Design," requiring vendors to provide the committee with a technical "traceability log" of the model's decision-making process. This prevents the "Black-Box" dilemma where proprietary secrets override public justice. Vendors must also commit to biennial "Bias Audits" to ensure their algorithms have not developed discriminatory patterns against specific demographics or socio-economic groups.

The Bharatiya Sakshya Adhiniyam (BSA), 2023, represents a landmark shift by recognizing electronic and digital records as primary evidence. However, Section 61 and Section 63 of the BSA present a "doctrinal gap" when applied to autonomous AI outputs. Traditionally, electronic evidence requires a human "certificate" to verify its authenticity. In the case of AI-generated analytics or deepfake-detected forensic reports, the law must evolve to allow "Machine-generated Certification." The proposed framework suggests an amendment to Section 63 to recognize certificates issued by JAOC-certified AI systems, provided they are accompanied by a digital signature and a verifiable chain of custody on a judicial blockchain.

To avoid stifling innovation, the framework adopts a Risk-Based Classification similar to the EU AI Act but tailored for India. "Low-Risk" tools, such as the SUVAS translation software, follow a simplified compliance path. "Medium-Risk" tools, like SUPACE for legal research, require transparency disclosures. "High-Risk" tools—those impacting personal liberty, such as AI-assisted bail risk scoring—are subject to the highest level of scrutiny, including mandatory human review of every output. This tiered approach ensures that administrative efficiency is not hampered by the same red tape required for sensitive judicial functions.

The framework codifies a new procedural right: the "Right to Human Review." Whenever an AI tool is used to assist a judge, the final order must explicitly state the extent of the AI's involvement. Litigants have the right to challenge the AI's "reasoning" as if it were a witness's testimony. If a lawyer suspects algorithmic bias in a case-scheduling tool or a sentencing recommendation, they can move the court for a "Technical Review" by the JAOC. This ensures that the machine remains a subordinate assistant and that the judge remains the ultimate arbiter of truth under Article 21.

Given the sensitive nature of litigation data, the framework mandates that all Judicial AI training and processing occur within a dedicated, sovereign "Judicial Cloud." Under the Digital Personal Data Protection Act (DPDPA), 2023, the Judiciary acts as a "Significant Data Fiduciary." The proposed framework prohibits the use of third-party commercial LLMs for processing active case files unless those models are hosted on local, government-controlled servers. This prevents "data leakage" where sensitive legal strategies or personal information could be used by private corporations to train commercial models.

The legal framework clarifies the "Liability Gap" for AI errors. If an AI-assisted tool "hallucinates" a precedent that leads to a wrongful judicial order, the framework establishes a Dual-Liability Model. For errors caused by technical "bugs" or un-audited biases, the certified vendor is held liable under a specialized Product Liability clause. However, if a judge or lawyer relies on an AI output without verification, they remain professionally accountable. This ensures that the "Human-in-the-Loop" isn't just a metaphor but a legal obligation with real consequences for negligence.

No framework is effective without competent executors. The proposed plan allocates a portion of the Phase III budget for the mandatory training of judges and court staff at the National Judicial Academy. This "Capacity Building" sutra aims to create a new cadre of "Techno-Legal" officers who can interpret AI audits and identify potential "deepfakes" in evidence. By 2026, every High Court is envisioned to have a "Technical Registrar" who assists the bench in understanding the technical nuances of AI-generated evidence and reports.

To maintain public trust, the framework adopts an "Open-Algorithm" Policy for all public-facing judicial tools. While proprietary code can remain protected, the "Logic-flow" and "Training-data origins" of tools like SUPACE must be public documents. This allows civil society, legal researchers, and the bar to act as external auditors. Transparency isn't just a

technical requirement; it is a "democratic imperative" that ensures the digital transformation of the judiciary remains inclusive and accountable to the citizens it serves.

Ultimately, the legal framework for Judicial AI in India is not merely about managing technology; it is about protecting the "Soul of the Law" in the age of Silicon. By institutionalizing the JAOC, enforcing Mandatory Certification, and updating the BSA 2023, India can lead the Global South in responsible AI adoption. The framework ensures that while the "wheels of justice" may turn faster through algorithms, they continue to stay on the path of constitutional morality.

CONCLUSION

The integration of Artificial Intelligence into the Indian Judiciary represents the most significant procedural shift since the adoption of the Constitution in 1950. As this research has demonstrated, the transition from the "e-Courts" era of digital storage to the "Intelligent Courts" era of algorithmic assistance is inevitable but fraught with constitutional risks. The pendulum of judicial reform is currently swinging toward rapid innovation to address the staggering backlog of cases; however, without the "Seven Sutras" of regulation, this momentum risks overshooting the mark of justice. The "Black-Box" nature of modern AI and the inherent biases in historical legal data pose a direct threat to the **Right to a Fair Trial** under **Article 21**. Therefore, the goal for 2030 is not merely a "faster" judiciary, but a "smarter" one that remains fundamentally human-centric.

To achieve this balance, the primary policy recommendation is the immediate statutory formalization of the **Judicial AI Oversight Committee (JAOC)**. This body must move beyond administrative coordination to become the ethical gatekeeper of judicial technology. Secondly, the **Bharatiya Sakshya Adhiniyam (BSA), 2023**, must be amended to provide a clear framework for the cross-examination of algorithmic logic. The law must evolve to recognize that in an AI-driven courtroom, the "witness" is often a line of code, and the "testimony" is a statistical probability. By mandating "Explainability by Design," the Indian legal system can ensure that every automated output is accompanied by a reasoned justification, thereby preserving the "Rule of Law" over the "Rule of Algorithms."