Research Partners

DAKSH

digital futures lab

UNDP

AI For Justice

Ethical, Fair and Robust Adoption in India's Courts

February 2026

UNDP is the leading United Nations organisation fighting to end the injustice of poverty, inequality, and climate change. Working with our broad network of experts and partners in 170 countries, we help nations to build integrated, lasting solutions for people and planet.

In India, we have been working for over seven decades in almost all areas of human development. Together with the Government of India, we work towards eradicating poverty, reducing inequalities, protecting the planet, enhancing community resilience, and accelerating sustainable development for all.

With projects and programmes in every state and union territory in India, UNDP works with national and subnational governments, and diverse development actors to deliver people-centric results, particularly for the most vulnerable and marginalised communities.

Research Partners:

DAKSH is a civil society organisation that has been working in the area of judicial reforms since 2015. We are placed at the intersection of law, policy and technology and follow an evidence-based approach to tackling the problem of pendency of cases.

Digital Futures Lab is an independent, interdisciplinary research studio that studies the complex interplay between technology and society in India and the Majority World. Through evidence-based research, systematic foresight, and public engagement, we work to realise pathways toward equitable, safe and sustainable digital futures.

This report was produced by DAKSH and Digital Futures Lab, commissioned by the United Nations Development Programme (UNDP) under its project 'Strengthening Rule of Law, Access to Justice pathways to accelerate India's Sustainable Development'. Views expressed in this publication are those of the authors and do not necessarily represent those of the United Nations, including UNDP, or the UN member States.

Date of publication: February 2026

Acknowledgements

This report has been prepared with the support and contributions of multiple institutions and individuals. We express our gratitude to UNDP for supporting this project. A special thanks to Nusrat Khan, Nalinaksha Singh and Thridhara Pathipati of UNDP for their guidance and support.

Our sincere thanks to the judicial officers, court staff, legal-tech founders and representatives, think tank researchers and lawyers who shared insights on current judicial processes and AI initiatives.

We appreciate the thoughtful review and comments received from domain experts in law, ethics, and public policy, particularly members of the Advisory Panel for this project:

- **Dr. Kamel EL HILALI,** AI & the Rule of Law Specialist, UNESCO Paris
- **Nehaa Chaudhari,** Partner, Ikigai Law
- **Justice Rajiv Shakdher (Retd),** Senior Advocate and former Chief Justice Himachal Pradesh
- **Prof. (Dr.) Sudhir Krishnaswamy,** National Law School of India University (NLSIU)

We thank reviewers of both the report and the assessment tools:

- **Aayushi Vishnoi**
- **Gurjot Singh,** Anand and Anand
- **Harish Narasappa,** DAKSH
- **Lakshmidevi Somnath,** Anand and Anand
- **Dr. Nandana Sengupta,** School of Public Policy, IIT Delhi
- **Pallavi Sondhi,** Ikigai Law
- **Peerapat Chokesusattanaskul,** Chulalongkorn University
- **Prashant Reddy T**
- **Sukriti,** Centre for Communication Governance, National Law University Delhi
- **Vibhav Mittal,** Anand and Anand
- **Vidhi Udayshankar,** Ikigai Law
- **Vinayak Hegde,** Brainy Yak Consulting LLP

We are grateful to all interview participants and reviewers for their time and expertise. Their contributions helped us better understand the practical realities, challenges and opportunities in integrating AI into the Indian judiciary.

Interview participants: Dinesh NG (Karnataka High Court), Bommi Reddy Meghana Vardhan (Abdul Latif Jameel Poverty Action Lab), Mousomi Panda (Research Associate, Aapti Institute), Nithiyanandam Yogeswaran (Takshashila Institution), Shashank Mohan (Centre for Communication Governance, National Law University Delhi), Siddarth Raman (XKDR Forum), Sushmita Vishwanath (Aapti Institute), Tanisha Singh (Nyaay AI), Vikas Mahendra (TEReS). Several interview participants who shared their experiences and insights generously prefer to remain anonymous.

We thank our colleagues and interns, including Ranjit Jacob, for their advice and research support.

Any errors or omissions are solely those of the authors.

Contributing authors:

DAKSH: Leah Verghese, Smita Mutt, Lakshmi Menon
Digital Futures Lab: Dona Mathew, Urvashi Aneja
Editing: Shreya Ramnath
Design: Meher Rajpal
Cover Illustration: Harshita Kesarwani
Cover Image Source: Jamillah Knowles & Digit / https://betterimagesofai.org / https://creativecommons.org/licenses/by/4.0/
The image has been edited to include additional design elements and to adjust saturation.

Table of Contents

Executive Summary

Artificial Intelligence (AI) is increasingly being used in the Indian legal sector, including within courts, often via experimental or "shadow" forms of use. While properly governed AI systems can improve access to justice, support the Sustainable Development Goals (SDGs), and strengthen institutional performance, their adoption must be underpinned by clear governance frameworks, transparency, and safeguards rooted in rights-based principles.

Understanding AI in the court context

AI systems operate through data, models and outputs, each of which raises concerns relating to accuracy, fairness and privacy. As custodians of highly sensitive information, courts must exercise caution in sharing data with external actors who do not operate under the same ethical mandates or levels of public responsibility. Risks such as bias, hallucinations, lack of transparency, and limitations in underlying datasets pose specific threats to judicial legitimacy.

Current landscape in India

India's eCourts digitisation efforts have been substantial, led by the Supreme Court eCommittee, High Courts, National Informatics Centre (NIC) and the Department of Justice (DoJ), with AI committees only beginning to emerge. Despite these efforts, AI deployment remains ad hoc and poorly documented. Translation and transcription are the most visible use cases, typically implemented through pilots that offer limited transparency around vendors, data handling, or validation methods. Many courts continue to lack technical capacity and impact-assessment processes, making it difficult to evaluate whether AI integration is improving judicial efficiency. Furthermore, the absence of digitised legacy records and accurate metadata leads to data gaps, increasing the risk that models built on available datasets reflect these gaps.

AI adoption is largely driven by individual champions, making it vulnerable to disruption due to judicial transfers or retirements. Additionally, it unfolds in a context of already stretched administrative capacity. While concerns around localisation, security and accuracy are common, the discourse on ethical considerations and mechanisms for post-deployment monitoring remains underdeveloped. Taken together, these gaps risk eroding public trust in courts and raise concerns about the legitimacy of the judiciary as AI use expands.

Global outlook

In countries around the world, from Singapore to the USA, various models of governance of AI use in courts have emerged. These are generally:

Cautionary:
Policies that either ban specific use cases (like recommendatory systems) or impose rigorous safeguards and oversight requirements for higher risk contexts.

Neutral:
Policies that lay down guidelines for individual court officers and judges but do not take an explicit prohibitive stance and generally follow a "wait-and-see" approach.

Enthusiastic:
Policies that encourage or facilitate the piloting and use of AI across the judiciary.

Sources of harm

The use of AI in courts carries the potential for violations of rights. Such harms may arise not only from training data and model design, but also from the specific institutional and procedural contexts of the legal system such as:

- the nature of the judicial process (e.g., translation vs judgment drafting),
- type of case or proceeding (e.g., privacy issues in sensitive case types), and
- characteristics of litigants (e.g., vulnerable witnesses).

Proposed frameworks

Given the absence of established processes and the potential for harm, this report proposes a reproducible, rights-based and risk-sensitive framework for AI adoption in the Indian judiciary. Tailored to the institutional realities of courts in India, this framework comprises four assessments:

Institutional Readiness: evaluates human resources, infrastructure and compliance preparedness before any adoption.

Risk Assessment: identifies potential harms and safeguards at the specific use case level.

Technical Assessment: examines vendor capabilities, data governance, transparency, security and accountability at the tool level.

Ongoing/Continuous Assessment: monitors real-world impacts, success metrics and emergent risks throughout deployment.

In addition to the frameworks proposed for the formal adoption of AI, this report also provides indicative contractual clauses for inclusion in agreements and MOUs with vendors, along with general good practices for judicial officers and court staff who may rely on generative AI for their day-to-day work.

Safeguards and good practices

Overall, we recognise the following good practices for courts considering the use of AI tools for judicial work:

- Domain expert consultation during AI tool research & Development
- Clear public disclosure of AI uses and vendors
- Human-in-the-loop for all substantive outputs
- Mandatory judicial or registry validation before operational use
- Data minimisation and privacy-by-design
- Opt-out opportunities for litigants
- Dedicated complaint and redress mechanisms
- Cybersecurity and data-security oversight
- Bias and fairness audits
- User-based audit logs to enable accountability and review

Conclusion

AI can support a more efficient, transparent and accessible justice system, but only if its adoption is guided by structured, rights-protective frameworks. The proposed assessments and safeguards are designed to ensure that AI strengthens rather than compromises judicial independence, fairness and public trust. However, assessments alone are insufficient to prevent unreliable or discriminatory outcomes. Complementary measures, including the establishment of technical cadres within courts, use case registries, and sandboxing frameworks, are essential to enable transparency, accountability, and responsible AI integration in India's judiciary.

List of Abbreviations & Key Terms

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
CERT-in	Indian Computer Emergency Response Team
CIO	Chief Information Officer
CIS	Case Information System
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions
CrPC/BNSS	Code of Criminal Procedure/Bharatiya Nagarik Suraksha Sanhita (BNSS), 2023
CTO	Chief Technology Officer
DOJ	Department of Justice
DPDPA/Rules	Digital Personal Data Protection Act, 2023 and associated Rules, 2025
EOI	Expression of Interest
HCCC	High Court Computer Committees
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social & Cultural Rights
iDEX	Innovations for Defence Excellence
IPR/IP	Intellectual Property Rights/Intellectual Property

IT Act	Information Technology Act, 2000
KPI	Key Performance Indicator
LLMs	Large Language Models
MeitY	Ministry of Electronics and Information Technology
MOU	Memorandum of Understanding
NJDG	National Judicial Data Grid
NCSC	National Centre for State Courts
NIC	National Informatics Centre
NIST AI RMF	National Institute of Standards and Technology Artificial Intelligence Risk Management Framework
NLP	Natural Language Processing
PPP	Public-Private Partnerships
R&D	Research & Development
RFP	Request for Proposals
SDG	Sustainable Development Goal
SUPACE	Supreme Court Portal for Assistance in Court's Efficiency
SUVAS	Supreme Court Vidhik Anuvaad Software
TERES	Technology Enabled Resolution
UDHR	Universal Declaration of Human Rights
UNESCO	United Nations Educational, Scientific and Cultural Organization

Key Terms

AI agents	An AI agent is a software system that acts autonomously (with minimal human direction or intervention) to carry out a specified task. It interacts with its environment to gather data, is capable of reasoning, and executes assigned tasks based on given parameters.
AI wrappers	An AI wrapper is a software layer that sits on top of an AI model, like an LLM, and facilitates user interaction with the model. This could be a chatbot or a UI/UX interface that uses APIs or another data exchange protocol to send prompts to the underlying model.
Black box	Black box is a metaphor for a model or system in which the inputs and outputs are clearly visible to users, but the internal logic, reasoning, or processes are not understandable.
Data minimisation	Data minimisation is a principle under which only limited and necessary data elements are captured, processed, or shared for a specified purpose. Non-essential data is not collected, and data is only kept for as long as necessary for the underlying purpose before it is deleted or masked.
Datasheet	A datasheet contains information about a dataset, including its structure, purpose, elements, and potential limitations or biases. It is considered a crucial piece of documentation and offers users and developers necessary details to build trust in the dataset. It is an increasingly expected element under ethical and responsible AI.
Encryption	Encryption is a cybersecurity process by which information is scrambled to protect it from unauthorised users and can only be re-composed with a secure key. End-to-end encryption (E2EE) is a variation where the message is only visible on an authorised sender and receiver's devices and is not accessible to the platform, internet service providers, regulators, etc.
Gen AI/Generative AI	Generative AI is a variant of AI models that generate new content based on training from a large dataset of similar material; this may generate text, images, audio, etc. and stands in contrast to task-oriented AI.

Hallucinations	Hallucinations are fabricated content provided by a model due to the probabilistic nature of generative AI. Hallcuinations may arise due to limited or biased training data or the model's training to provide an answer rather than admitting a prompt is beyond its scope of knowledge base.
Human-in-the-loop	Human-in-the-loop is a process/protocol by which human users review and refine AI-generated outputs by providing specialised knowledge or addressing bias. This improves the accuracy and relevance of the outputs as well as model performance over time.
Impact assessment	Impact assessments are structured processes to interrogate the future consequence of a given intervention and decide whether or not to go ahead in the present. They are intended to anticipate a wide range of effects and may be conducted with varying purposes (such as an ethical impact assessment vs. an environmental impact assessment).
ISO 42001	ISO 42001 is an international standard for an Artificial Intelligence Management System (AIMS) with guidance to establish, implement, maintain, and continuously improve AI systems to use them responsibly and effectively.
Logs	Logs are detailed records of user- or system-generated events.
Model cards	Model cards are files that are provided alongside models that contain information about their training datasets, parameters, intended uses, limitations or biases and evaluation results[1].
Model drift	A model trained on historical data may "drift", degrade, or become less relevant or accurate to current patterns over time, leading to prediction errors; models require continuous retraining and effective monitoring to remain up to date.
[Model] Overfitting	Overfitting occurs when a model is able to perform well on training data or data it has been exposed to but is unable to generalise its logic and yield appropriate results when exposed to new data. This occurs when the model is equally responsive to relevant and irrelevant data (i.e., noise).

Reinforcement learning	Reinforcement learning is the process of training machines to make decisions by trial and error. An agent interacts with its environment and receives feedback in the form of rewards and penalties, maximising performance for greater rewards.
Risk assessment	Risk assessments for AI involve identification and analysis of potential deviations, harms, or unintended consequences, forecasting areas from which they may accrue, and outlining how they will be managed across the lifecycle of AI.
Safety stop mechanisms	Safety stop mechanisms are ways to quickly halt functioning of AI systems when triggered, isolate harmful elements/processes and correct them before allowing them to resume. They can include kill switches for emergency intervention, blinding techniques to prevent learning from sensitive variables, and boxing techniques to isolate systems from external networks.
SOC 2	System and Organisation Control 2 is a data security compliance framework and audit process that reviews how cloud-based service providers maintain and protect customer data; it can be conducted by a third-party to provide assurance to a client and build trust that the software has adequate controls.
Shadow AI	Shadow AI is the use of AI by employees without the knowledge or authorisation of their supervisors or employers.
[Model] Underfitting	Underfitting occurs when a model is too simplistic and is neither able to perform well on training data or new data. It does not adequately identify logical patterns due to poor training and is not fit-for-purpose.
Unsupervised learning	Unsupervised learning occurs when a model is trained by detecting patterns in unlabelled data.

List of Abbreviations & Key Terms • Endnotes

1. Model Cards. Huggingface. https://huggingface.co/docs/hub/en/model-cards

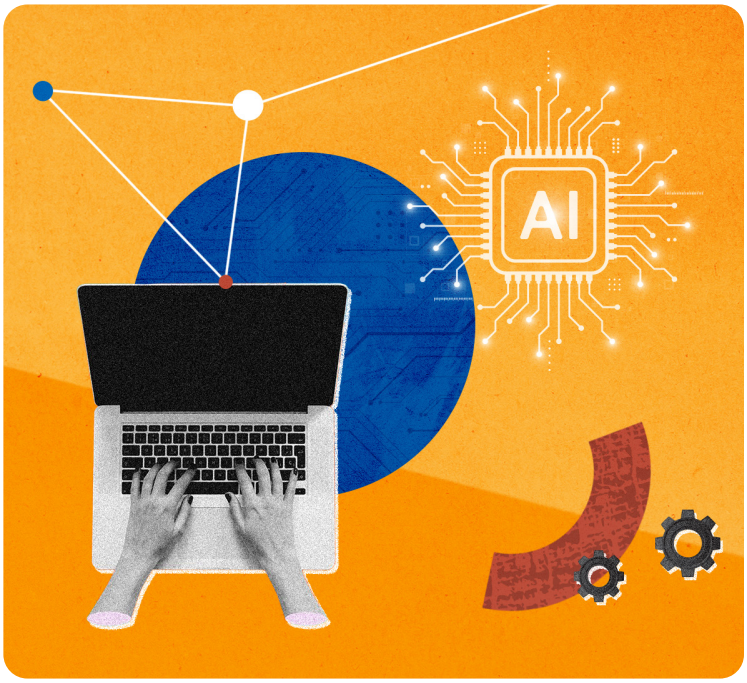
1

Introduction

In July 2025, Harvey AI announced plans to open its first office in Bengaluru, India's technology capital[2]. Given its role as a global AI platform for law firms, Harvey AI's expansion in India naturally aligns with the country's position as the world's second-largest legal-tech market[3]. Across the legal ecosystem, AI is increasingly viewed as a "force for good"[4], with AI-enabled tools being rolled out not only at the enterprise level but also in courts. In October 2025, the Kerala High Court directed all courts in the state to adopt Adalat AI, an automated transcription tool, for recording witness depositions[5].

AI's promise of convenience, time saving, and enhanced support is particularly compelling in a system as overburdened as the Indian judiciary. As of 15 December 2025, nearly five and a half crore cases were pending in India's district and High Courts, strengthening the case for AI use in courts[6]. As in other workplaces, the use of "shadow AI"[7] is also rising in courts[8]. A 2024 global UNESCO survey revealed that 44% of judicial operators had used AI tools for work-related tasks[9]. In the absence of expertise, training or guidelines for use, many judges and court staff are experimenting with primarily free tools, often unaware of the associated risks, such as the potential for AI tools to hallucinate precedents.

In November 2025, the Centre for Research and Planning at the Supreme Court of India published a white paper on AI and the judiciary, highlighting the need for responsible adoption of AI in courts[10]. The paper outlines potential judicial use cases as well as significant risks, raising rights concerns, including threats to due process, equality before the law, and the right to a fair trial. It also proposes guidelines for AI use among court users, lawyers and law clerks. This caution is warranted: when AI influences judicial adjudication, it can affect individual rights and, consequently, erode the credibility of the judicial system. The rule of law is compromised if judicial decision-making is not genuinely independent, if proceedings are not fair to all parties, or if emergent technologies impose discriminatory burdens on citizens trying to understand and engage with judicial institutions. Indeed, many of these risks are yet to be fully understood and have implications for institutional credibility and natural justice principles. The use of black box AI tools, whose internal decision-making processes cannot be readily examined, may further undermine the transparency of courts,



contravening Sustainable Development Goal 16, which calls on states to "provide access to justice for all" and "build effective, accountable and inclusive institutions at all levels"[11].

While jurisdictions around the world are adopting guidelines for the use of AI in the judiciary[12], designated a high-risk category by the EU AI Act[13], the regulatory landscape in India remains underdeveloped. In relation to the storage, processing, and sharing of litigant data collected by courts, the Digital Personal Data Protection Act is being enforced in a phased manner following the notification of the Rules in November 2025[14]. Beyond the recent Supreme Court white paper, the Kerala High Court is the only High Court to have issued a bare-bones policy for AI use in courts[15]. While the white paper prescribes punitive action for law clerks found using generative AI applications, it does not address the consequences or liabilities arising from court applications, which could have broader implications for litigants[16]. As India's legal-tech market expands and courts increasingly experiment with AI in judicial processes, it is an opportune time to develop decision-making models for designing and adopting AI systems in courts.

1.1 Structure of the report

This report contains the following sections:

- Brief on approach and methodology (Chapter 1.3)
- The current state of AI adoption in India, with a brief about the eCourts project, existing models for AI adoption, their limitations, and the contextual subtleties that impact AI adoption processes (Chapter 2)
- A global judicial AI policy outlook and prominent types of use cases (Chapter 3.1)
- The rationale for the proposed rights and risk-based approach to AI adoption (Chapter 3.2)
- Explanations for the assessment tools: risk assessment, institutional readiness, technical and ongoing/ continuous monitoring (Chapter 4)
- Annexures:
    1. Indicative Contractual Clauses; and
    2. Court User Guidance for AI

1.2 What is AI and how does it work?

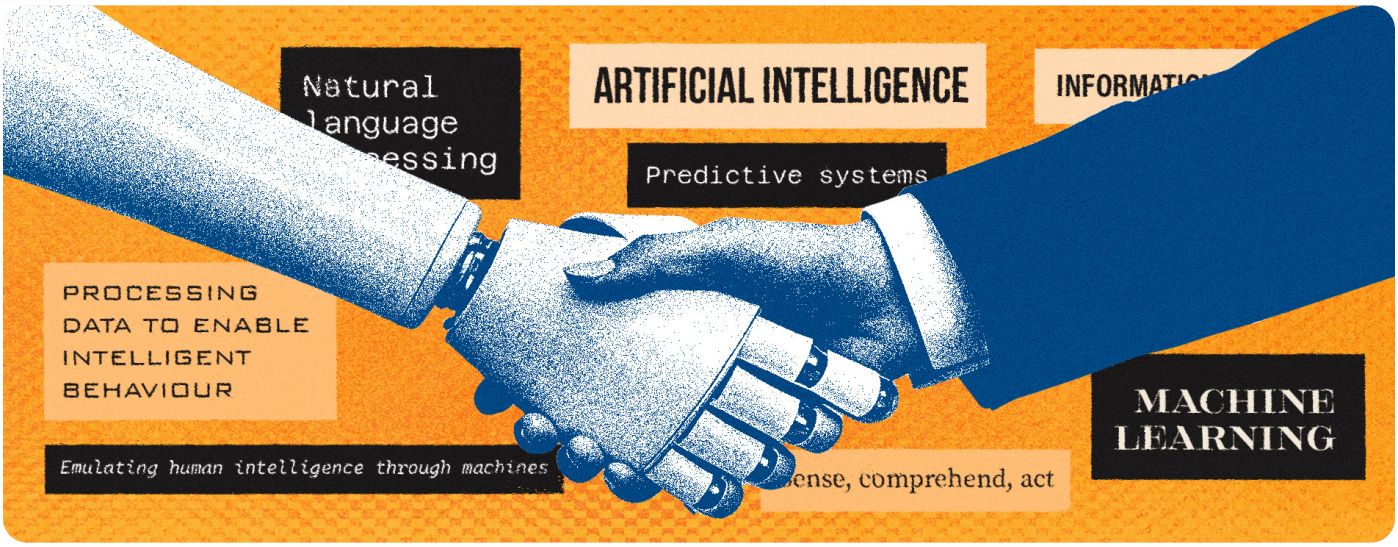
Given the fast-evolving nature of the field, public sector definitions of AI are intentionally broad in scope, covering a wide range of systems from "basic rule-based algorithms"[17] to "complex learning systems"[18]. They emphasise AI systems' ability to mimic human intelligence, both in the kinds of cognitive tasks they perform and in the diversity of outputs they generate, including text, images, audio, video, predictions, recommendations, or task outcomes performed by AI agents.

Definitions of AI:

- "AI is a constellation of technologies that enable machines to act with higher levels of intelligence and emulate the human capabilities of sense, comprehend and act. Thus, computer vision and audio processing can actively perceive the world around them by acquiring and processing images, sound and speech. The natural language processing and inference engines can enable AI systems to analyse and understand the information collected. An AI system can also take action through technologies such as expert systems and inference engines or undertake actions in the physical world. (NITI Aayog, 2018: 12)[19]

- "... the ability of the software systems to carry out tasks that usually require human intelligence: vision, speech, language, knowledge, and search" (World Bank, 2021: 2)[20]

- Computational systems that can "process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control" (UNESCO 2024: 7)[21]



An AI system rests on three building blocks: data, model, and output.[22]

Regardless of medium, AI models are trained on large volumes of **data**, which they analyse to identify relationships or patterns. Processes such as cleaning, labelling, and enrichment of datasets are essential to facilitate accuracy, broader applicability, and innovation. To ensure outputs that are accurate, fair, and unbiased, training data must be sufficiently representative, capturing socio-cultural, linguistic, demographic, and geographic diversity.

Limitations of judicial data for training AI

In the Indian judicial context, case-level data is often paper-based or marked by significant quality gaps[23]. Despite initiatives such as IndiaCode[24], there is still no single, authoritative database that consolidates all sources of law, from primary legislations and judgments, to delegated legislation and subsequent amendments and repeal, in a consistent, machine-readable format[25].

In 2018, the Ministry of Electronics and Information Technology (MeitY), drawing on literature identifying "big data" as a strategic resource in AI development, called on the government to create "high quality datasets" to facilitate individual and collaborative solution-building across domains[26]. As courts play a role in providing such datasets, it is essential to ensure that they adequately represent the diversity of litigants, disputes, processes, and vocabulary that appear before the judiciary. Doing so reduces the risks of biased or inaccurate AI outputs.

During **model** training and development, the system iteratively evaluates possible logical pathways and reinforces those that yield the most accurate results. In simpler systems, developers can set parameters and adjust the model when outputs require correction. However, once techniques such as unsupervised or reinforcement learning[27] are introduced, the model adapts to new data and refines its internal logic without explicit instructions. This dynamic and iterative learning process is a key reason AI requires procurement, governance, or regulatory approaches distinct from those used for conventional software.

To ensure robust **outputs**, models can be evaluated for factors such as accuracy, efficiency, and fairness. The metrics and thresholds applied (e.g., <1% error rate in transcription) must be grounded in the specific context of use. Outputs may be delivered through a chat interface, where human interaction can refine or iterate results, or may feed directly into task execution, such as in automated case-scheduling tools. Assistive systems and those that function autonomously require different protocols for acceptable use and oversight.

Notably, while simple or rule-based AI systems may be developed independently, more complex systems may sometimes be developed as "AI wrappers", applications built on top of existing commercial or open-source AI models[28]. Solution providers choosing among pre-built elements, models, or Application Programming Interfaces (APIs) available online typically weigh factors such as cost, performance, and ease of integration. Courts may wish to examine issues such as the representativeness of training data, the ethical development of models, or the uptime of APIs. However, when underlying core AI models are maintained by large global companies such as Meta, Google, OpenAI, or Anthropic, key design choices, training processes, and guardrails are often proprietary, limiting transparency and ethical scrutiny[29].

1.3 Research Approach and Methods

Although some guidelines for AI use in the judiciary have adopted a risk-based approach[30], this approach may prove inadequate when anticipated risks fail to capture broader or systemic rights violations. Risk-based approaches in AI governance have been criticised for overlooking various categories of risks, creating a false sense of finality while obscuring the centrality of rights[31]. For instance, a scheduling tool designed to reduce case disposal times by prioritising matters based on factors such as a litigant's gender, age, or profession may inadvertently disadvantage those not belonging to these groups. Even so, risk assessments remain valuable tools to anticipate and mitigate foreseeable harms arising from the use of AI systems. Recognising both the limitations and the value of such frameworks, this report proposes an approach for AI governance in the Indian judiciary that is anchored in both rights and risks.

The UNESCO Global Toolkit on AI and the Rule of Law for the Judiciary adopts a human rights approach to AI governance, identifying specific rights that may be affected by AI deployment[32]. In its recent report on AI procurement and deployment, the United Nations (UN) likewise emphasises the centrality of human rights protection to regulatory developments in AI[33]. In addition, the UNESCO Guidelines for the use of AI systems in Courts and Tribunals[34] and the Recommendations on the Ethics of AI[35] highlight several key principles relevant to our current work, including but not limited to:

- **Proportionality and do no harm:** AI technologies adopted by the judiciary should mitigate any potential harms to human rights, fundamental freedoms, and the environment.
- **Fairness and non-discrimination:** AI technologies in judicial processes must adhere to equality before the law and not discriminate on the basis of class, caste, gender, religion or sexual orientation.
- **Safety and security:** Safety risks and system vulnerabilities must be addressed throughout the AI lifecycle.
- **Privacy and data protection:** Given the sensitive nature of court data, data management protocols must be established before data is shared with third parties.
- **Human oversight and accountability:** Where there is a possibility of harm, there must be clear pathways for demonstrating human responsibility across the system's lifecycle.

- **Transparency:** Transparency is an essential prerequisite to the protection of human rights and fundamental freedoms, but it must be balanced against considerations of safety and privacy.

- **Explainability:** The availability of intelligible insights into the working of an AI system. A system is explainable if its inputs, outputs and functioning can be traced by humans. Explanations must go beyond generic or static explanations and take into account varying information and contextual needs. They should be offered in plain language and supported by supplemental resources (text, visual, audio, video), accessible without technical expertise, and available in multiple Indian languages.

- **Accuracy and reliability:** Adopt AI systems that perform consistently and correctly across differing inputs, contexts, and situations.

- **Awareness and literacy:** In judicial contexts, responsible AI adoption hinges on capacity-building within the ecosystem to evaluate and recognise both the potential and limitations of AI.

- **Judicial independence:** Given the role of courts in upholding the rule of law and the Constitution, ultimate decision-making must remain human-led to safeguard judicial independence.

1.3.1 Methodology

The research for this report employed a combination of of desk research and stakeholder interviews. The literature review examined AI-use guidance documents in other jurisdictions, AI risk assessment frameworks, and AI procurement guidelines for the public sector. The authors conducted semi-structured interviews with sitting and retired judges, members of the registry, representatives from civic-tech and legal-tech companies, civil-society actors, and academics. Interviewees have been anonymised in all citations. The authors also visited courts where some of these tools are in use. A draft of the report and proposed framework was circulated for feedback, and comments received have been incorporated into this version.

1.3.2 Limitations

While the research team has made every effort to capture the current state of AI adoption in Indian courts and develop a risk assessment framework within six months, the study has certain limitations:

- The framework has been tested internally for different use cases but will require refinement when applied across court levels and geographies.
- The assessment tools are not proposed as stand-alone pathways to realise responsible AI adoption, they need to be accompanied by more research and action on a whole-of-system approach, including considering formal procurement processes

01

Introduction • Endnotes

2. The Economic Times. (2025, July 11). Harvey AI bets big on Indian legal market, announces Bengaluru office opening. https://legal.economictimes.indiatimes.com/news/industry/harvey-ai-bets-big-on-indian-legal-market-announces-bengaluru-office-opening/122380058
3. Ibid; Definition of legal-tech: For the purpose of this report, legal-tech refers to any digital technology that developed for the legal industry, including case management tools, workflow automation, and online dispute resolution.
4. Grand View Horizon. (n.d.). India Legal Technology Market Size & Outlook, 2025-2030. https://www.grandviewresearch.com/horizon/outlook/legal-technology-market/india
5. Pratap, G. (2025, October 3). Kerala High Court mandates all courts in State to adopt AI tool to record witness depositions. Bar and Bench. https://www.barandbench.com/news/kerala-high-court-mandates-all-courts-in-state-to-adopt-ai-tool-to-record-witness-depositions
6. High Courts of India (n.d.). National Judicial Data Grid. https://njdg.ecourts.gov.in/njdg_v3/.
7. The use of AI by employees without the knowledge or authorisation of their supervisors or employers.
8. Pollard, J. (2025, June 13). Small Purchases, Big Risks: Shadow AI Use in Government. Forrester (Featured Blogs). https://www.forrester.com/blogs/small-purchases-big-risks-shadow-ai-use-in-government/.
9. UNESCO. (2024). UNESCO Global Judges' Initiative: survey on the use of AI systems by judicial operators. https://unesdoc.unesco.org/ark:/48223/pf0000389786.
10. Centre for Research and Planning, Supreme Court of India. (2025). White Paper on Artificial Intelligence and Judiciary. https://cdn.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/uploads/2025/11/2025112244.pdf
11. United Nations (n.d.). Sustainable Development Goal 16. https://sdgs.un.org/goals/goal16.
12. E.g., Superior Council of the Judiciary. (2024, December 16). Agreement PCSJA24-12243, "By which guidelines are adopted for the respectful, responsible, safe and ethical use and exploitation of artificial intelligence in the Judicial Branch". https://rm.coe.int/colombia-guidelines-for-the-use-of-artificial-intelligence-in-the-judi/1680b53484; Singapore Judiciary. (2024). Guide on the use of generative artifical intelligence tools by court users. https://www.judiciary.gov.sg/docs/default-source/news-and-resources-docs/guide-on-the-use-of-generative-ai-tools-by-court-users.pdf?sfvrsn=3900c814_1; Courts and Tribunals Judiciary. (2025, October). Artificial Intelligence (AI): Guidance for Judicial Office Holders. (n.d.). https://www.judiciary.uk/wp-content/uploads/2025/10/Artificial-Intelligence-AI-Guidance-for-Judicial-Office-Holders-2.pdf; Arizona Judicial Branch. (2024, April). Arizona Code of Judicial Administration, Part 1: Judicial Branch Administration Chapter 5: Automation Section 1-509: Use of Generative Artificial Intelligence Technology and Large Language Models. https://www.azcourts.gov/Portals/0/0/admcode/pdfcurrentcode/1-509%20Use%20of%20AI%20Tech%20and%20LLMs%2001_2025.pdf?ver=acMF-P2SER0dArzTQohBjQ%3D%3D; Supreme Court of New South Wales. (2024). Guidelines For New South Wales Judges In Respect Of Use Of Generative AI. https://supremecourt.nsw.gov.au/documents/About-the-Court/policies/Guidelines_Gen_AI.pdf; Conselho Nacional de Justica. (2025). RESOLUTION No. 615/2025, of March 11, 2025: Establishes guidelines for the development, use, and governance of artificial intelligence solutions within the Judiciary. https://rm.coe.int/resolution-6152025/1680b51b66.
13. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal L 2024/1689 (2024). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
14. Ministry of Electronics and Information Technology. (2025, November 13). Digital Personal Data Protection Rules, 2025. The Gazette of India: Extraordinary. https://www.meity.gov.in/static/uploads/2025/11/53450e6e5dc0bfa85ebd78686cadad39.pdf
15. High Court of Kerala. (2025, September). Policy Regarding The Use of Artificial Intelligence (AI) Tools in District Judiciary. https://images.assettype.com/theleaflet/2025-07-22/mt4bw6n7/Kerala_HC_AI_Guidelines.pdf
16. (Centre for Research and Planning, Supreme Court of India, 2025).
17. Hickok, M. (2024). From Trustworthy AI Principles to Public Procurement Practices. De Gruyter. p. 22
18. (Centre for Research and Planning, Supreme Court of India, 2025).
19. NITI Aayog. (2018). National Strategy for Artificial Intelligence. https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf
20. World Bank Group (n.d.). Artificial Intelligence in the Public Sector: Summary Note (English). https://documents1.worldbank.org/curated/en/746721616045333426/pdf/Artificial-Intelligence-in-the-Public-Sector-Summary-Note.pdf
21. UNESCO (n.d.). Document for consultation: Draft UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals. https://unesdoc.

unesco.org/ark:/48223/pf0000390781
22. (Hickok, 2024, p.24); Kaplan, S., Uusitali, H., Lensu, L. (2024). A unified and practical user-centric framework for explainable artificial intelligence. Knowledge-Based Systems, Volume 283, 111107. https://www.sciencedirect.com/science/article/pii/S0950705123008572
23. Damle, D. and Anand, T. (2020). Problems with the e-Courts data. NIPFP Working Paper Series. https://www.nipfp.org.in/media/medialibrary/2020/07/WP_314__2020.pdf
24. India Code. Digital Repository of Laws - A system of laws for communication. https://www.indiacode.nic.in/
25. DAKSH (2021, January). Whitepaper series on Next Generation Justice Platform, Paper 5: Single Source For Laws. https://dakshindia.org/wp-content/uploads/2021/03/FINAL-Daksh-Paper-March-3-.pdf
26. MEITY (2019). REPORT OF COMMITTEE - A ON PLATFORMS AND DATA ON ARTIFICIAL INTELLIGENCE. https://www.meity.gov.in/static/uploads/2024/02/11ab.pdf
27. "Unsupervised learning occurs when a model is trained by detecting patterns in unlabeled data. Reinforcement learning is the process of training AI systems to make decisions by allowing them to "interact with their environment to maximise cumulative rewards". Ghasemi, M., Ebrahimi, D. (2024). Introduction to Reinforcement Learning. https://arxiv.org/pdf/2408.07712
28. Ellen, L. (April 28, 2025). When OpenAI Isn't Always the Answer: Enterprise Risks Behind Wrapper-Based AI Agents. Towards Data Science. https://towardsdatascience.com/when-openai-isnt-always-the-answer-enterprise-risks-behind-wrapper-based-ai-agents/
29. Hopkins, A., Struckman, I., Klyman, K., Silbey, S. (2025). Recourse, Repair, Reparation, & Prevention: A Stakeholder Analysis of AI Supply Chains. FAccT '25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency. https://dl.acm.org/doi/10.1145/3715275.3732017
30. Rangone, N., & Megale, L. (2025). Risks Without Rights? The EU AI Act's Approach to AI in Law and Rule-Making. European Journal of Risk Regulation, 16(3), 1082–1097. doi:10.1017/err.2025.13; De Souza, S.P. (2024, November 28). AI and the Indian Judiciary: The Need for a Rights-based Approach [HTML version], The Hindu Centre for Politics and Public Policy. https://www.thehinducentre.com/incoming/ai-and-the-indian-judiciary-the-need-for-a-rights-based-approach-html-version/article68917505.ece.
31. (Rangone, N., & Megale, L. 2025); Malgieri, G., & Santos, C. (2025). Assessing the (severity of) impacts on fundamental rights. Computer Law & Security Review, 56, 106113. https://doi.org/10.1016/j.clsr.2025.106113
32. UNESCO. (2025). Global Toolkit on AI and the Rule of Law for the Judiciary. https://unesdoc.unesco.org/ark:/48223/pf0000387331
33. OHCHR. (2025). Artificial Intelligence Procurement And Deployment Ensuring Alignment With The Guiding Principles On Business And Human Rights (A/HRC/59/53). https://www.ohchr.org/sites/default/files/2025-06/a-hrc-59-53-executive-summary-eng.pdf
34. UNESCO. (2024). Draft UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals. https://unesdoc.unesco.org/ark:/48223/pf0000390781
35. UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000380455.

2

Current state: AI use in Indian courts

Key Takeaways

- **While digital infrastructure is robust, AI governance in courts remains underdeveloped.** The eCourts project has created a nationwide digital backbone, through CIS, NIC infrastructure, and the NJDG, enabling access to standardised case data across courts. However, this infrastructure was designed for digitisation, not for governing AI use, leaving gaps as AI tools begin to rely on this data.

- **Though growing, AI use in courts remains fragmented and opaque.** AI adoption has largely focused on translation and transcription, supported by human review, alongside limited use of summarisation and research tools. Beyond visible uses, many AI applications operate with little public disclosure, making it difficult for litigants to know when AI affects court processes.

- **"Pilots" dominate AI adoption, often without structure or exit criteria.** Courts use the term "pilot" broadly, covering proofs-of-concept, co-development, and limited deployment. Many pilots lack clear objectives, timelines, evaluation metrics, or feedback loops, resulting in prolonged experimentation without evidence of impact.

- **Multiple actors build AI tools, but coordination is weak.** AI tools are developed through in-house teams, NIC partnerships, academic collaborations, and private vendors. Courts increasingly prefer legal-tech startups for performance and customisation but lack consistent standards to evaluate accuracy, bias, or rights impacts.

- **Adoption processes prioritise security over rights and accountability.** While courts emphasise data localisation, encryption, and human oversight, existing processes rarely assess algorithmic bias, systemic harms, or long-term accountability. Informal vendor engagement and individual decision-making further weaken transparency.

- **Institutional constraints heighten risk.** Limited technical capacity, reliance on individual champions, frequent transfers, and informal use of generative AI by staff increase the risk of automation bias, errors, and erosion of trust, underscoring the need for a structured, rights-based AI framework for courts.

2.1 How courts digitised: The eCourts project

Technology adoption in Indian courts has been largely driven by the eCourts project, a Mission Mode Project funded by the Union Government that enabled the digitisation of district courts[36]. Launched in 2005, the eCourts mission aims to leverage technology to improve judicial productivity both qualitatively and quantitatively, and make the justice delivery system affordable, accessible, cost-effective, transparent, and accountable. It has digitised judicial administration across roughly 8,300 court establishments in India via a system known as the Court Information System (CIS). To ensure uniformity and integration, infrastructure design and specifications were centralised, including a central cloud architecture and common services. At the same time, implementation was decentralised under the core and periphery model[37], allowing courts to respond to local needs and exercise autonomy in areas such as vendor selection.

The following institutions are involved in the project:

- **Supreme Court eCommittee:** responsible for the policy planning, strategic direction, high-level design, specifications and certifications,
- **National Informatics Centre (NIC):** responsible for the development of technology and provision of support to the High Courts,
- **High Court Computer Committees:** responsible for implementing the project at the state level, and
- **The Department of Justice (DoJ):** responsible for financing[38].

In addition, the Supreme Court and some High Courts have also set up dedicated AI Committees to identify opportunities for the use of AI, with an emphasis on process automation, case management, and holistic justice dispensation[39]. As part of Phase 3 of the eCourts project, Rs. 53.57 crore has been earmarked explicitly for Future Technological Advancement, including AI and blockchain, across High Courts in India[40].

The High Courts and Supreme Court run on NIC software (except in Kerala, Delhi, and Madhya Pradesh). The NIC also maintains the National Judicial Data Grid (NJDG). This online dashboard aggregates and displays real-time data from all district courts, High Courts and the Supreme Court. From a data standpoint, one of the most significant outcomes of the eCourts project is the availability of "unified and standardised" data on cases across courts in India. Various research organisations, such as DAKSH, Vidhi Centre for Legal Policy, XKDR, and Enfold Proactive Health Trust[41], have utilised eCourts data to analyse pendency, delays, and related issues in courts. This data is also valuable to legal-tech firms developing AI tools for courts, and for the legal sector more broadly.

2.2 How AI tools are currently used

So far, AI integration in the administration of justice in courts has been ad hoc and difficult to track. However, the clearest and best-documented use case of AI is in document translation, particularly that of reportable judgments across Indian languages[42]. This is overseen by AI committees and supported by robust human review, with High Courts engaging retired judges and legal experts to vet the AI-generated translations[43]. The Supreme Court developed SUVAS (Supreme Court Vidhik Anuvaad Software) with technical support from MeitY)[44], while the NIC has created AI Saransh[45], a tool for summarisation. Private company ManCorp Innovation Labs has developed SUPACE, an AI-enabled tool designed to help Supreme Court judges examine large volumes of legal material more efficiently[46]. However, despite the proliferation of these tools, there is limited information on whether they are being used and if user feedback informs further improvement.

In parallel, private vendors have begun developing tools specifically for judges and courtroom use. One such example is Adalat AI, an AI-based transcription and translation service, currently deployed in courts across nine states (as of 8 December 2025)[47]. This use is based on MOUs signed by Adalat AI with various High Courts[48]. Some AI tools are also being piloted in courts on a non-commercial basis.

What is a pilot?

Interviewees used the term "pilot" to refer to a variety of engagements including:

- proofs of concept, where court users were given test logins or court APIs were connected to validate tools against judicial expectations,
- collaborative development with courts, in which legal-tech developers were given access to judges and/or staff over long periods to understand court processes and develop solutions customised to their needs or pain points, and
- limited deployment in assigned courtrooms or for specified case types to seek feedback and iterate software, design, or model performance.

In practice, any involvement short of full-scale deployment was referred to as a "pilot". In this section, we use the term broadly to cover this full spectrum. However, we recommend that a structured pilot must include clear documentation, indicators they seek to measure, causal attribution, feedback loops, and ideally a fixed timeframe. Without these elements, pilots fail to generate meaningful insights into implications or corrections and can become a drain on resources. We observed instances of pilot phases continuing for months without any agreed-upon end date, effectively becoming "perpetual pilots".

In addition to Kerala, Adalat AI transcription and translation tools are available in district courts in Delhi, Karnataka, and Odisha. The Supreme Court uses Technology Enabled RESolution (TERES) transcription and validation services for Constitution bench cases[49], while Karnataka uses TERES services without the human validation component. These examples show that AI tools are being integrated into the judicial system rapidly, but in ways that are piecemeal and often opaque.

While uses such as live transcription, real-time translation, and automation of court documents are visible to litigants and the public, others, like AI-based summarisation or legal research, are less apparent to those outside the system. Public disclosure on where or how AI is being used in courts, or what litigant or case data may be shared with developers, remains minimal. Although case status information, orders and judgments available on eCourts are publicly accessible, courts also hold personal, identifiable, and sensitive data, including pleadings and other documents that are not available to the public[50].

There is also limited transparency around how courts select vendors, the criteria used to approve experimental technology for use in live court contexts, and how an AI system may be deemed suitable for a specific task or whether it poses technical and ethical risks.

The Computer and/or AI Committees in the Supreme Court and High Courts, tasked with making decisions on AI adoption, often comprise judges who may not have the technical expertise to identify risks or stay abreast of measures to evaluate AI tools[51]. In the absence of prescribed technical assessment frameworks, courts risk deploying AI systems that contain biases, produce inaccurate outputs, or compromise privacy and security.

Moreover, different courts use different tools with uneven levels of integration into court management systems. In the absence of comprehensive impact assessments, courts cannot effectively measure whether AI tools actually improve judicial efficiency, accuracy, or access to justice. The opacity surrounding AI deployment also raises concerns about due process and the right to a fair trial. Litigants and their legal representatives may be unaware when AI tools are being used in their cases, limiting their ability to challenge AI-generated outputs.

Even as AI adoption accelerates, the transition to fully digital processes remains incomplete. Despite nearly 20 years of the eCourts project, much of the day-to-day judicial work still relies on physical files, paper registers, and manual processes. The COVID-19 pandemic forced courts to allow e-filing and hold virtual hearings, but courts are yet to become natively digital institutions. Case files in many courts are still maintained in physical form, hard copies still need to be filed even where e-filing is permitted, and extensive physical registers are used to track the movement and storage of documents. While paperless courts gain momentum, the absence of end-to-end digital workflows limits the ability of digital case files to serve as a reliable single source of truth. Although the Phase 3 eCourts Vision Document envisages open, interoperable digital infrastructure allowing for unified, evolutionary systems, this is far from being realised[52].

For judges and court staff accustomed to physical processes, the introduction of AI tools represents not merely a technological upgrade but a profound shift in understanding, workflow, and capacity. Effective adoption of AI in courts requires a foundational understanding of AI models' functions and limitations. Without this knowledge about the capabilities, ethical boundaries, limitations and potential sources of bias in AI tools, courts risk misusing tools, making uninformed decisions on where and how to deploy AI, and failing to detect errors or rectify unfair outcomes.

2.3 Who builds AI tools for Indian courts?

To understand the promises and limits of judicial AI[53], it is essential to examine how problems are identified and who develops the solutions. In recent years, a variety of use cases have emerged for use by judges and court staff, with transcription and translation tools seeing the most widespread adoption. The availability of infrastructure, including web cameras, microphones and desktops for virtual court proceedings, has further enabled the uptake of transcription technologies[54].

Across different levels of courts, a range of collaborations support AI development, with courts partnering with third-party organisations to compensate for limited institutional technical capacity. Some models for AI development include:

- **In-house development:** The Kerala High Court built in-house tech capacity by creating an IT Directorate to develop and customise tools for the High Court and the district courts.

- **Co-development with the National Informatics Centre (NIC):** The NIC, under the MeitY, is the technology partner of the Government of India. The Supreme Court developed SUVAS in association with the NIC[55]. In Odisha, a scrutiny tool is being developed in collaboration with NIC[56].

- **Partnerships with academic institutions:** The Supreme Court has partnered with IIT Madras to develop an AI-enabled scrutiny[57] tool[58]. In Manipur, the High Court is partnering with NIT Imphal to develop a translation tool in the local Meiti language[59].

- **Private vendors:** Given the burgeoning legal-tech sector in India, courts are exploring partnerships with startups to acquire and, in some cases, co-develop relevant tools. These organisations exist across categories with varying revenue models. For example, Adalat AI is a non-profit supported by grants, with an objective of "building India's end-to-end Justice Tech Stack"[60]. TERES is a for-profit startup with global clients from Singapore to Dubai, and offers a range of services in addition to transcription—e-discovery, document management, electronic evidence presentation and legal analytics[61]. Jhana, a for-profit startup that raised USD 1.6 million in seed funding in 2024[62], offers products to parse through documents and provide stenographer support, and is working with the Karnataka High Court to develop summarisation tools[63]. Nyaay AI is another startup developing "AI-powered case management tools" incubated by PanScience Innovations in 2022[64].

Our interviews revealed a growing interest for AI tools created by legal-tech startups, seen as offering greater reliability and accuracy. Some stakeholders noted, for instance, that SUVAS has not yet reached the accuracy levels offered by third-party vendors' tools[65].

2.4 How courts acquire AI today

When it comes to technology adoption, the judiciary has several options. Formal procurement processes may rely on model Request for Proposals (RFPs) and Expressions of Interests (EOIs) drafted by the Executive and adapt them for the provision of hardware or software licenses *(goods)*[66], managed service contracts or transformation through *consultancy* basis[67], or co-development under a variety of *PPP* models[68].

At present, AI adoption in courts is often ad hoc and relationship-based, with pilot phases lasting long periods, or even involving the personal procurement of technology. As courts engage vendors for transcription and translation tools, the pilot phase may extend for many months; Adalat AI, for instance, has been piloting its tool across courts for at least a year. Courts generally prefer vendors approved or originated by the Executive; for instance, they would opt for Bhuvan over Google Maps[69]. However, as private companies' tools increasingly offer better performance for customised use cases, courts are gravitating towards them, particularly when their teams and governing boards include lawyers with strong litigation experience and an awareness of court-specific challenges.

Our interviews also revealed a relational approach to AI integration, with vendors approaching judges, registrars and other court staff to present their tools and offer demonstrations. While this informal approach to procurement may seem barrier-free, the reliance on personal preferences for AI adoption increases the risk of overlooking a variety of concerns, which can be prevented with selection processes involving objective criteria based on performance and the protection of rights. Such informal processes also create an uneven playing field, restricting access to courts for some vendors.

Where courts have engaged in a formal process, they have relied on established public sector procurement processes. For instance, in May 2023, the Supreme Court published a RFP for the design and development of AI-based transcription tools[70]. This process involved multiple stages, including bid meetings with numerous companies, shortlisting, and demonstrations. The bid document restricts the data used for the model or transcriptions from being shared with any third party and requires human oversight supported by adequate human resource staffing. In addition, the document mandates annual training of staff, technical support and maintenance, including bug fixes and upgrades.

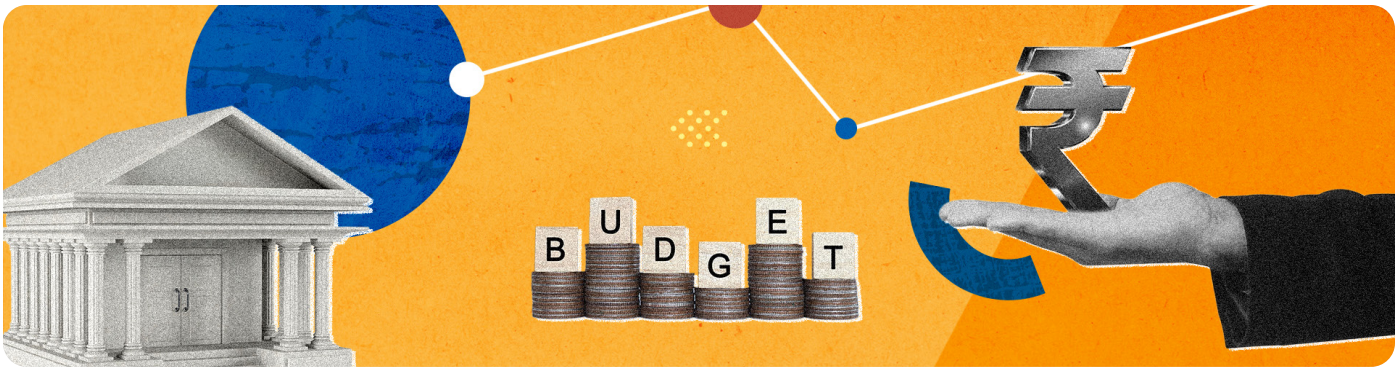
In a second example, the Karnataka High Court published an Expression of Interest (EOI) in November 2024 for tools to translate judgments from English into Kannada. This one-page document invites concept notes for the proposal and mandates accuracy checks by comparing translated output with the source text, to be undertaken by translators or linguistic experts[71].

Re-imagining procurement for AI tools

In addition to formal procurement methods such as traditional RFPs and EOIs, courts may consider innovative models such as challenge-based procurement or co-development arrangements. These approaches are especially useful when commercial, off-the-shelf solutions may be inadequate for certain use cases.

One such method is the Ministry of Defence's Innovations for Defence Excellence (or iDEX) scheme, introduced in 2018 to "provide financial support to startups/MSMEs/individual innovators" developing prototypes and specialised technologies for hyper-specific and confidential defence needs[72]. While the Supreme Court has previously engaged with hackathons, a clear pipeline from idea identification to procurement and deployment remains underdeveloped. For iDEX winners, the Defence Acquisition Procedure 2020 provides this pathway, allowing emerging technologies to be legally procured.

Where courts consider co-development, they (or the DOJ) may need to perform "incubation functions" such as facilitating access to mentorship, financial support, and accurate, representative datasets. Courts should also proactively invite vendor demonstrations and stay informed about emerging trends and best practices from other sectors and jurisdictions.



However, neither of the above processes addresses broader ethical considerations, including the potential of the tool to violate fundamental rights like fairness, equality and protection from non-discrimination. There are no explicit requirements for safeguards to ensure that training datasets are free from bias, to assess the scope and extent of model vulnerabilities, or to specify the risk mitigation measures in place.

Financial sustainability is also a concern. We identified an instance, reported in an interview, in which a judicial officer independently procured and adopted an AI tool using funds allocated under the "computer equipment" budget head. Beyond the obvious risk posed by use of a tool that has not been vetted or approved by the court, individual subscription-based purchases may not be a financially viable option for multiple junior staff across courts.

Risks also accompany the unsanctioned use of Large Language Models (LLMs) like ChatGPT for daily work by judicial officers and court staff. For example, a city civil judge in Bengaluru cited "two non-existent Supreme Court judgments" in a 2025 judgment. When this was challenged during the appeal, the appellants' counsel pointed out that the use of AI may have influenced the reliance on "fictitious case laws"[73].

It is important to also examine how technology adoption is financed. In the case of High Courts and district courts, the state government approves the budget for technology adoption under the judicial head[74]. For the Supreme Court, a budget of more than Rs. 50 crores has been allocated under the eCourts project specifically for AI[75], to be directed by the eCommittee.

In many instances, courts have piloted tools on a non-commercial basis. Some courts explicitly prefer such financial models, noting that vendors may commercialise their tools by selling to lawyers and law firms[76]. While this approach is understandable given the need to demonstrate a tool's effectiveness before large-scale adoption, it does carry risks. Companies may use free tools to create dependencies. A common business practice is to provide baseline services for free but then charge a premium for upgrades, enterprise customisation, or additional security. Experience in other contexts suggests that the free version sometimes relies on opaque data collection practices or ad revenue, offers lower quality, or uses deceptive design practices (also known as "dark patterns") to pressure users into upgrading[77].

However, the lack of process on paper does not mean that courts are not cautious in their approach to AI adoption. We observed several good practices during our research such as:

- **Data protection:** Judges and registrars are concerned about the scope of data access, storage and use under the Digital Personal Data Protection Act, 2023 especially around huge fines in case of data breaches[78]. Vendors are expected to provide assurances that data is stored within India.

- **Translation validation:** Various High Courts[79] have published notices to recruit retired judicial officers, advocates, and court translators to validate AI-generated translation of judgments.

- **Disclosure of AI use:** In a pilot court in Delhi, litigants and lawyers are informed that a transcription is being undertaken by a software tool and requires signatures from the person deposing and the Presiding Officer, as well as a check by the stenographer before it becomes part of the case file[80].

2.4.1 Gaps in existing AI adoption processes

For courts adopting AI, data governance is a primary concern. This includes data localisation, deployment on secure or dedicated servers, user-level access controls, encryption, and the removal of sensitive information. Courts also emphasise that data shared for AI deployment should not be used for training models for unrelated purposes[81]. In general, courts desire ownership over data, although the implications of this ownership (management, systems, third-party support) remain unclear. The emphasis on data localisation is strong enough for some courts to consider establishing their own data centres[82].

While data protection is critical, this narrow focus risks obscuring other critical algorithmic harms. An insistence on checklist factors like data encryption and localisation has meant that considerations like errors and bias from datasets, and the accuracy of outputs receive less attention in existing discourse. For example, for pilots and R&D, vendors may access documents like pleadings that are not public when the case is ongoing[83]. It remains unclear if these datasets are being used to train other models, and even where vendors provide guarantees, courts lack the mechanisms to verify their claims.

Relationship-based approaches to acquiring new technologies can also disrupt court work by creating dependencies without adequate safeguards. If a vendor withdraws a free pilot after many months of usage, judicial processes may be disrupted, subtly coercing courts to continue using the tool. For a public institution like the court, such approaches also risk undermining trust and transparency. For example, the details of bid documents can be influenced by the existing offerings of successful pilot tools, constraining open competition and disadvantaging vendors without access to courts. Moreover, the current system lacks a framework for monitoring contract stipulations after adoption, requiring additional human resources and assessment frameworks to evaluate the technology's impact on overall court functioning.

A key caveat in this discourse, however, is that procurement guidelines and monitoring frameworks alone cannot guarantee responsible technology adoption. This would require stronger institutional scaffolding and a sustained commitment to protecting fundamental rights.

2.4.2 Institutional challenges to AI adoption

Beyond the absence of clear guidelines on responsible AI adoption in courts, institutional factors also impact the pace and certainty of decision-making regarding technology integration.

Technology adoption and terms of adoption are driven by individuals. Individual judges maintain relationships with vendors and articulate the institutional vision for acquiring new technology. Judge-centric digital transformation pathways thus impact the pace at which High Courts and district courts adopt technology. In the event of the transfer of a judge who champions a certain technology, procurement processes already underway can stall. Additionally, individuals in charge of procurement may also develop their own criteria for testing the efficacy of AI tools, leading to inconsistent and non-replicable standards across courts.

Variability in judicial tenures and the fact that Registrars in charge of administration are district judicial officers on deputation creates significant administrative flux. In the absence of stable institutional scaffolding, the identification and evaluation of AI tools will be disrupted each time a Registrar in charge is transferred.

Judicial officers are not best placed to make technical decisions. While the creation of technology cadres has been discussed[84], with some courts already seeking approvals for instituting such cadres, there are concerns that a) long-term cadre employees might be unable to keep pace with rapidly evolving technologies, making contractual employees appear more suitable, and b) career growth remains uncertain for long-term cadres.

Stretched institutional capacity is fostering informal AI use. The promise of AI to save time and improve productivity attracts adoption by staff, who already use freely accessible generative AI tools[85]. This is not unique to India; a survey of approximately 1400 judges in Colombia showed that 80% reported using free AI tools[86]. This risks creating an over-reliance on AI, potentially leading to automation bias. For example, in the U.K., an automated software miscalculated spouses' financial worth, leading to errors in alimony calculations[87]. This went undetected for many months, illustrating the human tendency to defer to machine-generated outputs.

2.5 Why courts need a structured AI framework

Stakeholder interviews reveal that ad hoc, personality-driven decisions guide AI adoption in Indian courts. Existing approaches tend to prioritise basic data security protocols, with limited attention to concerns around rights violations and documented harms arising from AI use in other sectors and jurisdictions. This includes insufficient consideration of risks associated with accuracy, reliability, and bias that may emerge from (i) limitations of training data and (ii) the probabilistic nature of AI.

As public institutions entrusted with the delivery of justice, with direct consequences for people's liberty and freedom, it is imperative that courts subject any AI deployment to rigorous and systematic review.

Current state: AI use in Indian courts • Endnotes

36. e-Committee, Supreme Court of India. (n.d.). E-Courts Mission Mode Project. https://ecommitteesci.gov.in/project/brief-overview-of-e-courts-project/.
37. Calcutta High Court. (n.d.). CIS 3.0 Core & Periphery Modules. https://calcuttahighcourt.gov.in/downloads/ecourt_files/cis3/core_and_periphery_module/CIS_3.0_Core_&_Periphery_Modules.pdf.
38. e-Committee, Supreme Court of India. (2022). Digital Courts: Vision and Roadmap. p. 14. https://cdnbbsr.s3waas.gov.in/s388ef51f0bf911e452e8dbb1d807a81ab/uploads/2023/04/2023042088.pdf.
39. Department of Justice, Ministry of Law and Justice, Government of India. (2022). Lok Sabha Starred Question No. *147 to be answered on Friday, the 16th December, 2022. https://eparlib.sansad.in/bitstream/123456789/1465709/1/AS147.pdf.
40. Ministry of Law and Justice. (2025, February 25), Digital Transformation of Justice: Integrating AI in India's Judiciary and Law Enforcement [Press Release]. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2106239.
41. DAKSH. (2020). DECIPHERING JUDICIAL DATA: DAKSH'S DATABASE. https://www.dakshindia.org/deciphering-judicial-data-daksha-database/; DAKSH & Vidhi Centre for Legal Policy. (2019). Litigation Landscape of Bengaluru. Series 1: Bengaluru Rural Courts. https://www.dakshindia.org/wp-content/uploads/2019/08/litigation-landscape-bengaluru-rural-full-report-july-2019.pdf ; Vidhi Centre for Legal Policy. (2022). A Decade of POCSO: Developments, Challenges and Insights from Judicial Data. https://vidhilegalpolicy.in/research/a-decade-of-pocso-developments-challenges-and-insights-from-judicial-data/ ; Manivannan, P., Raman, S., Zaveri-Shah, B. (2025, October 5) Beyond Pendency: Counting Cases Correctly. The Leap Blog. https://blog.theleapjournal.org/2025/10/beyond-pendency-counting-cases-correctly.html#gsc.tab=0 ; Enfold Proactive Health Trust (2025). The Possibilities of eCourts Data for Advancing Research on Law Implementation. https://enfoldindia.org/wp-content/uploads/2025/03/The-Possibilities-of-eCourts-Data-for-Advancing-Research-on-Law-Implementation.pdf
42. Ministry of Law and Justice. (2025, April 3). The Supreme Court is collaborating with the High Courts in translation of e-SCR Judgements in 18 vernacular languages [Press Release]. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2118241.
43. Lakshman, A. (2023, December 31). SC's translation projects raced ahead in 2023 as retd. HC judges, law clerks help AI. The Hindu. https://www.thehindu.com/news/national/scs-translation-projects-raced-ahead-in-2023-as-retd-hc-judges-law-clerks-help-ai/article67692773.ece; Mimansa. (2024, December 9). Retired Judges, Advocates Paid Rs 100 Per Page for Vetting AI Translations of Supreme Court Judgments And Why it Matters. Medianama. https://www.medianama.com/2024/12/223-rs-100-ai-translation-vetting-supreme-court/
44. The Economic Times. (2023, Aug 11). AI backed SUVAS translation tool intended to make legalese simpler, court proceedings faster: Law minister. https://government.economictimes.indiatimes.com/news/technology/ai-backed-suvas-translation-tool-intended-to-make-legalese-simpler-court-proceedings-faster-law-minister/102648151.
45. Ministry of Electronics and Information Technology, Government of India (n.d.). AI-Saransh. https://cloud.gov.in/user/services_ai_saransh.php.
46. Singh Pratap, A. (2021, April 8). Artificial Intelligence Portal SUPACE launched by Supreme Court of India. Statecraft. https://www.statecraft.co.in/article/artificial-intelligence-portal-supace-launched-by-supreme-court-of-india.
47. Adalat Ai (n.d.). Powering progress in Indian courts. https://www.adalat.ai/impact.
48. SU, Personal Communication, 8 December 2025
49. The vendor also provides a service of manual verification of documents to check for errors.
50. DAKSH. (2022). Judicial Data Regulation: Balancing Open Courts with the Right to Privacy - An Indian Perspective. https://www.dakshindia.org/wp-content/uploads/2023/08/Judicial-Data-Paper-1-1.pdf. pp 5-6.
51. High Court of Karnataka. (2025, July 28). The Computers and Technology Committee of the High Court of Karnataka has a provision to invite Technical Experts as Special Invitees. https://judiciary.karnataka.gov.in/common_folder/notification//reg-committees-28072025.pdf.
52. e-Committee, Supreme Court of India. (2022). Digital Courts: Vision and Roadmap. https://cdnbbsr.s3waas.gov.in/s388ef51f0bf911e452e8dbb1d807a81ab/uploads/2023/04/2023042088.pdf.
53. Definition of judicial AI: For the purpose of this report, judicial AI refers to any use of AI by the judiciary for functions relating to case filing, processing, management, or adjudication. We do not consider administrative functions that are purely internal, such as human resource of financial management modules.
54. Microphones capture voice data from within the courtroom, and from video in case of virtual appearance; the wires are routed through a box, and a single wire is connected to the system. The transcription happens in real time on the dashboard. The stenographer first selects the language for transcription and can keep an eye on it as it is transcribed. The steno can also edit the transcript once it is completed.
55. Press Information Bureau. (2023, August 10) Action Plan for Simple, Accessible, Affordable and Speedy Justice [Press Release]. https://www.pib.gov.in/PressReleasePage.aspx?PRID=1947490.

56. SS, Personal Communication, 25 August 2025.
57. Scrutiny is the process of identification of defects in filings and provides the party to rectify any defects such as errors in status, provisions or information provided.
58. Press Information Bureau. (2025, March 20). Use of AI in Supreme Court Case Management [Press Release]. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2113224.
59. High Court of Manipur. (2023, August). Memorandum of Understanding between High Court of Manipur and the National Institute of Technology Manipur [Press Release]. https://ecourtsprojectmanipur.in/wp-content/uploads/2023/08/MOUNIT.pdf.
60. Adalat AI. About. https://www.adalat.ai/.
61. TERES. About. https://teres.ai/.
62. Jhana. About. https://jhana.ai/.
63. GS, Personal Communication, 30 July 2025.
64. Nyaay AI. Home. https://www.nyaayai.com/.
65. SS, Personal Communication, 25 August 2025.
66. Department of Expenditure, Ministry of Finance, Government of India. (2021, October). Model Tender Document for Procurement of Goods. https://eprocure.gov.in/cppp/sites/default/files/standard_biddingdocs/MTD%20Goods%20NIC.pdf.
67. (Department of Expenditure, Ministry of Finance, Government of India. (2023, April). Model Tender Document for Procurement of Consultancy Services (Includes Guidance Note for Procuring Entities). https://doe.gov.in/files/circulars_document/Model_Tender_Document_for_Procurement_of_Consultancy_Services_2.pdf.
68. Department of Economic Affairs, Ministry of Finance, Government of India. Model Request for Proposal [for PPP Projects]. https://www.pppinindia.gov.in/model_request_proposal.
69. RS, Personal Communication, 9 July 2025.
70. Supreme Court of India. (2023). Bid Document: Design, Development, and Implementation of AI solution tools for Transcribing Arguments and Court proceedings at Supreme Court of India. https://cdnbbsr.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/uploads/2024/01/2024012579-1.pdf
71. High Court of Karnataka. (2024). Invitation for Expression of Interest (EOI) for providing/Developing Artificial intelligence tool for Machine translation of Judgements of the Hon'ble Supreme Court of India and Hon'ble High Court of Karnataka from English to Kannada language to accomplish the project of Assisted translation of Judgments and Judicial Records in Vernacular Language for improved access to the Justice. https://judiciary.karnataka.gov.in/common_folder/notification//HCK-EOI-AI-Tool-15112024.pdf.
72. Government of India Ministry of Defence (n.d.). iDEX Details. https://www.ddpmod.gov.in/offerings/schemes-and-services/idex.
73. The News Minute. (2025, March 28). Karnataka HC calls for probe against judge for citing fake SC judgements. https://www.thenewsminute.com/karnataka/karnataka-hc-calls-for-probe-against-judge-for-citing-fake-sc-judgements
74. RJ, Personal Communication, 6 August 2025.
75. (High Court of Karnataka, 2024).
76. GD, Personal Communication, 31 July 2025.
77. Ford Foundation. (2023). A guiding framework to vetting public sector technology vendors. https://www.fordfoundation.org/learning/library/research-reports/a-guiding-framework-to-vetting-public-sector-technology-vendors/.
78. TJ, Personal Communication, 12 August, 2025.
79. Ministry of Law and Justice.(2025, April 23). The Supreme Court is collaborating with the High Courts in translation of e-SCR Judgements in vernacular languages [Press Release]. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2118241; High Court of Gujarat. (2025, March 3). Requirement for vetting/validation of translation. https://gujarathighcourt.nic.in/hccms/sites/default/files/miscnotifications/Notice_Inviting_Advocate.pdf
80. SS, Personal Communication, 13 August 2025.
81. SS, Personal Communication, 25 August 2025.
82. TJ, Personal Communication, 12 August 2025.
83. TJ, Personal Communication, 12 August 2025.
84. e-Committee, Supreme Court of India. (2022). Digital Courts: Vision and Roadmap. https://cdnbbsr.s3waas.gov.in/s388ef51f0bf911e452e8dbb1d807a81ab/uploads/2023/04/2023042088.pdf.
85. MS, Personal Communication, 19 August, 2025.
86. Gutiérrez, Juan & M., David. (2025). Use of AI tools in the Colombian Judiciary: Findings from Three Surveys. Conference on Digital Government Research. https://doi.org/10.59490/dgo.2025.1060
87. BBC. (2015, December 17). Divorce form error 'could have led to unfair settlements'. https://www.bbc.com/news/uk-35128010

3

Governing AI in courts: A rights-and-risk-based framework

Key Takeaways

- **Risk assessment must precede AI deployment in courts.** Before adopting any AI tool, courts must identify potential harms, assess their likelihood and severity, and define safeguards to protect fundamental rights.
- **Risk and harm are distinct but related.** Risk refers to the probability and scale of negative outcomes from an AI use case, while harms describe actual individual, collective, societal, or systemic consequences.
- **Judicial AI directly affects constitutional rights.** AI use in adjudication and administration can implicate rights to a fair trial, liberty, privacy, equality, free expression, and effective remedy, requiring rights-centred evaluation.
- **Binary risk classifications (high and low) are insufficient.** Treating some judicial AI uses as "low risk" primes users to be less cautious and can lead to omissions during oversight. These may compound into systemic harms over time.
- **Context matters.** AI risks vary by use case, case type, and litigant type, with heightened stakes in criminal justice, family law, and cases involving vulnerable groups.
- **India needs a rights-based, risk-sensitive framework.** Uneven digital readiness across courts makes safeguard-focused assessments essential for maintaining judicial legitimacy.


WAITING ROOM
FOR
CHILDREN / VICTIMS
UNDER P.O.C.S.O. ACT
SPECIAL P.P
POCSO COURT
FAMILY COURT
ERNAKULAM
SPECIAL COURT, CUTTACK
1st. Floor, FTC. Building
Estd. 17-05-2008.

When it comes to AI in courts, there is wide divergence between countries on where and how the use of AI is acceptable. Courts' responses can be categorised into three broad approaches:

Cautionary:
Policies that outline institutional governance structures and safeguards across higher and multiple risk levels.

Neutral:
Policies that offer broad guidance for individual court officers and judges, but do not take an explicit stance, typically adopting a "wait-and-see" approach.

Enthusiastic:
Policies that encourage the use of AI across the judiciary.

These approaches are outlined below, with examples from select jurisdictions. Policies in this space are evolving rapidly and may fluctuate between various positions on the spectrum over time.

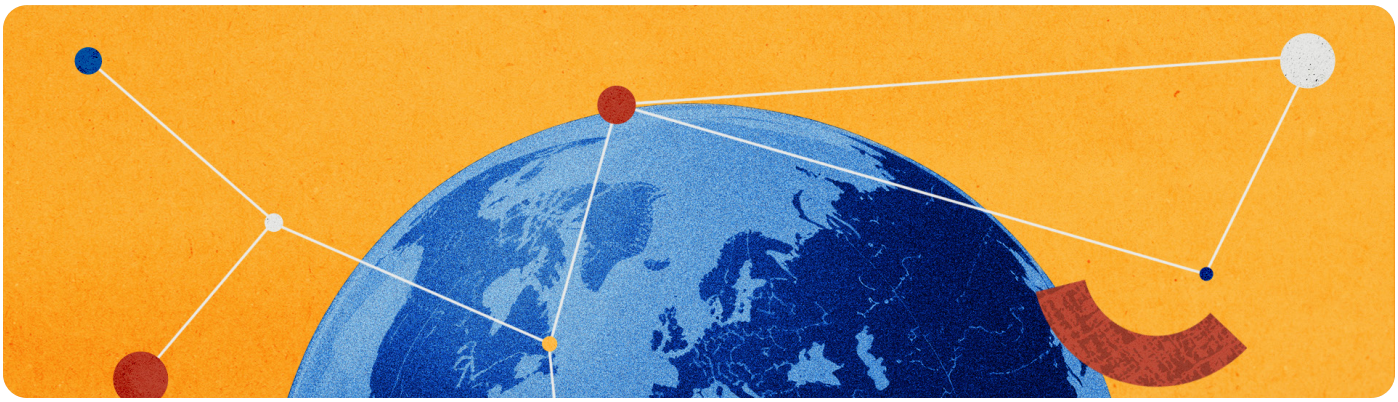
Table 1: Regulatory variations to governing AI in courts

Governance structures and safeguards for multiple risk-levels	• Colombia's policy outlines three risk levels for use cases - prohibited, high and low-risk, and a framework for AI deployment by courts[88]. • The EU's AI Act classifies the judiciary as a high-risk sector[89]. Very few activities, including administrative tasks such as communication between court personnel, anonymisation of judicial decisions, etc., are seen as non-high risk[90].
Governance structures and safeguards for high-risk AI	• In India, the Kerala High Court released a policy governing AI use in the district judiciary[91]. This policy prohibits the use of AI for judicial decision-making, requires district courts to use High Court authorised AI tools, and recognises the High Court's IT committee as an oversight body.

General guidance documents for court staff and judges; no specific governance structure mentioned	• Different provinces within Australia have issued guidance documents to judges, seeking to outline a list of dos and don'ts. For example, guidelines for New South Wales offer minimal prohibitions (using commercial and general purpose LLM tools for analysis of evidence, or reasons for judgments)[92].
Governance structure to encourage wide AI deployment with few or discretionary safeguards	• The United Kingdom's AI Action Plan for Justice[93] offers a roadmap for the widespread incorporation of AI in various judicial functions, both in administration and in criminal justice risk assessments. Judicial guidance for courts is restricted to a list of dos and don'ts for judges[94]. • Policies adopted across different states in the United States (US) vary. While policy and legal interventions have attempted to limit how tools such as predictive algorithms may be used by the judiciary[95], tools like Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), which assesses risk of recidivism in offenders have not been banned[96].

Different jurisdictions offer lessons on how to manage AI deployment in judicial settings. The National Centre for State Courts (NCSC) in the US, for example, provides guidance to judges on how AI can help assess evidence and verify authenticity[97]. As Indian courts consider regulatory approaches, they will need to weigh the benefits of speed, processing, and automated validation against possible harms to citizens.

One regulatory approach for Indian courts is to undertake structured risk assessment. Such assessments involve identifying and analysing potential harms, anticipating areas in which harms may accrue, and outlining how they will be managed[98]. While risk assessments should occur throughout the AI lifecycle[99], it is particularly crucial before deployment. Under this framework, decisions about AI adoption are not solely aimed at technical efficiency, but also at embedding safeguards to minimise and prevent violation of rights.



In this section, we use the terms:

Risk: In judicial contexts, AI use can lead to a range of outcomes, positive, negative, neutral, or a combination thereof. Here, risk refers to the composite potential of an AI system to produce negative consequences in a specific use case, the likelihood of its occurrence, and the expected magnitude of harm[100].

Harms refer to the negative consequences arising from the use of AI in a given use case. These include consequences for individuals or collectives/groups of individuals[101]. Negative consequences can extend beyond immediate impact and be societal (for example, where there is a harm caused to a specific societal interest), and systemic (occurring in patterns cumulatively over a period of time)[102].

AI use may pose risks to human rights and democratic values[103]. In the context of the judiciary (an institution entrusted with upholding the rule of law and enabling justice for rights violations)[104], the incorporation of AI has direct implications for fundamental rights. Any risk assessment undertaken before deploying AI in court must be grounded in a framework which is informed by the potential impact of automation on these rights. The use of AI in both adjudication and administration can impact the rights to a fair trial and appeal, personal liberty, privacy, equality, and even the rights to freedom of speech and expression[105].

3.1 Issues with binary approach to risk

Across sectors and countries, AI regulation and policy have largely centred on risk assessments of potential and future-facing harms[106]. These approaches typically classify AI systems as "high-risk" and "low-risk" (and sometimes, prohibited); identify certain sectors/activities as "high-risk" scenarios for AI use; and prescribe safeguards for such uses. "Low-risk" AI systems are subject to minimal regulatory oversight[107].

This binary division of AI tools by risk and the limited safeguards attached to "low risk" uses have attracted significant criticism. Risk-first approaches are criticised for being short-sighted, as they fail to account for the long-term or cumulative impact of certain AI tools[108]. A rights-based approach to AI impact assessments broadens the scope of inquiry by capturing both immediate and enduring harms to human rights.

To counter these criticisms, we propose a risk assessment framework to be undertaken by courts before AI deployment that is cognisant of the need to manage AI risk throughout its lifecycle[109]. This framework measures risk not only in terms of likelihood or scale of harm, but also by examining the presence or absence of institutional safeguards to mitigate risk. Such a framework is especially important in the context of the Indian judiciary, where digital readiness is uneven across courts.

3.2 A rights-based framework for courts

A rights-based perspective accounts for the varying degrees of harm that AI may cause to individuals' rights, including the ways in which adoption in different use cases may reinforce or perpetuate graded inequalities[110]. This is particularly important given the centrality of the protection of fundamental rights to the Indian judiciary's constitutional mandate. The subsections that follow thus propose a broader categorisation of elements to capture the impact of AI uses in court.

AI can adversely impact rights due to multiple reasons, including gaps and biases in training data, choices around parameter selection, the open or closed nature of models, and data breaches. In the Indian judicial context, rights violations may emerge from 1) use case scenario and how AI is integrated into the process, 2) case type or proceeding type, and 3) litigant type.

a) Use case: Proposed use cases in Indian courts include translation, speech detection, case scheduling, administrative analytics and legal research assistance. Each use case has its own implications for rights[111]. In addition, not every use case requires AI intervention. This framework proposes having a clear rationale before introducing AI for a use case.

b) Case or proceeding type: Certain types of cases and proceedings are prone to severe and even irreversible harm. Errors and bias arising in cases concerning –

- life (end-of-life care cases, health emergencies, abortion petitions under the Medical Termination of Pregnancy Act, 1971, death penalty cases and others),
- liberty (criminal justice decisions such as bail, sentencing and premature release, immigration detention),
- privacy (cases pertaining to family law and divorce, mental health records, juvenile justice, testimonies of survivors of sexual violence),
- proceedings under the Official Secrets Act, 1923

and other related rights can be violated, causing irreversible harm. Case types can be categorised by the rights they can engage as shown in the table below[112].

c) Litigant type: It is impossible to prescribe a hard-and-fast rule for classifying AI risk levels based solely on specific impact on a litigant. Committees within courts responsible for procuring AI tools must recognise that risks arising from AI use vary based on the litigant type, making such risks highly context-specific and qualitative. To account for this variability, the proposed framework measures use and safeguards for AI for certain vulnerable groups of litigants.

Table 2: Mapping rights to case and litigant types

Rights group	Framework in India	Illustrative case/litigant types
Right to privacy	In 2017, a nine-judge Bench of the Supreme Court of India found that the right to privacy was a constituent of various fundamental rights under the Constitution of India[113]. The decision found that the right to privacy covered family life, personal autonomy, sexual autonomy, communications, data privacy, an individual's medical records, and more.	1. Family and marital law proceedings 2. Domestic violence cases 3. Cases under the Information Technology Act, 2000 4. Healthcare records 5. Identification of party in proceedings involving reproductive rights, including Medical Termination of Pregnancy Act, 1971, Human Immunodeficiency Virus and Acquired Immune Deficiency Syndrome (Prevention and Control) Act, 2017, Surrogacy (Regulation) Act, 2021, etc. 6. Identifying information of survivors of sexual violence 7. Identifying information of parties to a case, including identifiers, addresses, e-mail IDs, financial records and details, etc.
Free speech, expression, association and protest	Article 19 of the Constitution of India enumerates a number of liberties, including the freedom of speech and expression, the right to peaceful assembly, the right to association, freedom of movement, and the freedom of occupation. These rights come into play in various contexts, including publications on print and digital media, access to the Internet, defamation cases, and magisterial powers to prohibit gatherings.	1. Proceedings under Section 144 Code of Criminal Procedure, 1973/ Section 163 Bharatiya Nagarik Suraksha Sanhita, 2023 2. Directions to individuals or intermediaries to block/remove content on the Internet 3. Defamation suits or prosecutions for criminal defamation

Rights of protected groups and minority rights	1. Article 15 of the Constitution enumerates that no person shall be discriminated against on the ground of race, religion, caste, gender, sex, place of birth. Article 17 also prohibits the practice of untouchability. 2. Article 26 guarantees freedom of religion for religious denominations or any section thereof to establish institutions for religious and charitable purposes, acquire and administer property. 3. Article 29 offers the right to preserve distinct language, script or culture.	1. Litigation under special legislation designed for protection of vulnerable groups including Scheduled Castes and Scheduled Tribes (Prevention of Atrocities) Act, 1989, Rights of Persons with Disabilities Act, 2016, Mental Healthcare Act, 2017, Juvenile Justice (Care and Protection of Children) Act, 2015, Protection of Children from Sexual Offences Act, 2012 etc. 2. Prosecutions under Sections 196 of the Bharatiya Nyaya Sanhita, 2023
Procedural rights	Article 21, which protects the right to life and personal liberty, requires a court of law to follow "procedure established by law" in matters that may concern these rights. The Supreme Court of India has established that such procedure includes within its ambit a right to a fair trial (including requirements like a fair and impartial hearing, sufficient notice, reasoned decisions, etc.).	1. Rights of an accused in criminal procedure (framing of charge, examination and cross-examination of evidence, the right to rebut incriminating evidence, and the right to a sentencing hearing) 2. Right to be heard in writ petitions filed under Article 32 or Article 226 of the Constitution
Liberty	Article 21 of the Constitution protects the right to life and personal liberty, requiring measures that deprive such rights to be done through a "procedure established by law". Article 22 offers a framework for preventive detention.	1. Habeas corpus petitions 2. Criminal law proceedings and challenges to exercise of criminal justice powers 3. Proceedings under state and central preventive detention laws

AI use cases: Potential risks, rights violations and their causes

As AI is adopted in judicial processes, limitations of training data, model design, and the nature of human-computer interaction risk the violation of the rights listed above. These rights violations manifest differently depending on how and where AI is deployed in the court system. In the table below, we map emerging AI use cases against potential benefits, anticipate their impact on rights and identify some sources of risks.

USE CASE: Case scheduling

Current state: Registry/Judicial Officer dependent; the cause list is prepared by the administrative Registry of the court.

Potential AI use

- **Optimisation:** Support predictive analytics that prioritise cases for cause list generation based on user needs.
- **Data analytics:** Process weighted parameters such as case types, previous history of adjournments or number of hearings to estimate case timeframe, and build "optimal" schedules, and ensure cases are included in the cause list.
- Reduce discretion of the Registry in listing cases.
- Increase effective hearings by optimising the listing process.

Types of Risk

- **Output errors:** Incomplete data and subjective weights in models can lead to errors of omission; this can lead to de-prioritisation or exclusion of certain case types, litigants or older cases.
- **AI modelling errors:** Errors and hallucinations.

Rights implications

- **Right to equality before law:** Compound systemic inequalities if variables like case type, gender, caste and access are not factored into model training (A.14, Constitution of India; A.7, UDHR; A.14, ICCPR).
- **Non-discrimination:** Unintentional discriminatory scheduling (A.15, Constitution of India).

Sources of Risk

- **Data limitations:** Data entry errors may arise from gaps in data, including non-representative data.
- **Model limitations:** The model may be limited by overfitting or underfitting.
- **User-level risk** such as manipulation, or gaming by advocates.

USE CASE: Transcription and note-taking

Current state: Manual process by stenographer; lawyers' arguments are not always transcribed; witness depositions are always transcribed.

Courts are sometimes using free tools for transcription.

The manual process is prone to errors, and is time-intensive, particularly in noting, verification and approval.

Potential AI use

- **Comprehensive transcripts:** Error-free transcriptions of the entire hearing, including arguments and depositions.
- **Timesaving and resource-efficiency:** Saves time and resources for court and litigants.

Types of Risk

- **Inaccurate outputs:** Transcription tools may not accurately capture certain dialects and accents; Hallucinations fill in gaps and pauses.

Rights implications

- **Equality:** Compound systemic inequalities if variables like region, gender and caste are not factored into training for different accents and dialects (A.14, Constitution of India; A. 7, UDHR; A.14, ICCPR).
- **Right to liberty and fair trial:** Errors in transcription can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** Absence of training data representative of diverse accents and dialects in Indian languages may cause their continuous neglect. In the long term, this may lead to systemic exclusion of certain communities.
- **Model limitations:** Model may fail to account for pauses in the deposition or arguments. This may prejudice the record against litigants and witnesses who lack fluency and confidence in court.
- **Privacy and data misuse risk:** Risks of data breaches and unauthorised access to non-public data.

USE CASE: Translation between Indian languages

Current state: Empanelled translators translate documents on request. This process can be slow, expensive, with limited capacity.

Potential AI use

- **Time efficiency:** Automated translations are faster.
- **Diversity-enhancing:** Documents can be made available to litigants and the public at large in multiple Indian languages.

Types of Risk

- **Inaccurate translations:** Linguistic mistranslations, failure to translate complex legal language accurately, as well as potential for loss of nuance and context.
- **Exclusions of languages, accents and dialects:** AI-generated translations are easier for resource-rich languages, thereby excluding others.

Rights implications

- **Equality:** Compound systemic inequalities faced by communities speaking resource-scarce languages (A.14, Constitution of India; A.7, UDHR; A.14, ICCPR).
- **Discrimination:** Discriminatory outcomes from lack of representation in training data (A.15, Constitution of India; A.2, UDHR).

Sources of Risk

- **Data limitations:** Non-representative data that does not capture legal jargon, rules of interpretation, and colloquialisms.

USE CASE: Case law research

Current state: Online keyword-based search on legal databases.

Potential AI use

- **Granularity and speed:** Extract relevant information, including pointed summaries, at faster speeds, yielding more accurate and comprehensive results.

Types of Risk

- **Hallucinations:** Language models can make up case details in summaries.
- **Bias:** Self-learning algorithms can bias results based on user patterns, as an outcome of efficiency improvement.

Rights implications

- **Right to fair trial:** Inaccurate or incomplete research can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** From annotation stage and gaps from archival non-digitised reportable judgments.
- **Model limitations:** Overfitting and underfitting results.

USE CASE: Scrutiny of filings

Current state: Conducted manually by court administration based on established rules and guidelines. This is time-consuming.

Automated scrutiny is being used for certain case types in Kerala.

Potential AI use

- **Productivity improvements:** Rule-based scrutiny can be completed by AI with human oversight to verify accuracy.
- **Fast-track listing:** By improving scrutiny timelines, cases can be listed faster.

Types of Risk

- **Lack of clarity:** Documented rules and guidelines may not cover all local variations. This may confound the model.

Rights implications

- **Right to fair trial:** Delays caused by disputed defects can impact the right to a fair trial under procedure of law (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** Incorrect labelling or annotation in training data can limit defect identification accuracy.

USE CASE: Case-level research support: verification of case facts, summarisation, relevant data extraction from large case files.

Current state: Conducted manually by judges and/or law clerks.

Potential AI use

- **Time-efficiency:** Reduces time spent on sense-making of large, complex, unstructured case files.
- **Comprehensive analysis:** Avoid human errors of omission when parsing through documents and cross-verify recorded facts.

Types of Risk

- **Inaccurate outputs:** Incorrect identification of facts, and hallucinations in extracts.
- **Interpretation errors:** Limitations of training the machine to make sense of complex legal documents, loss of context and nuance.
- **Privacy:** Risk of non-public data breach.

Rights implications

- **Right to fair trial:** Incorrect, inaccurate outputs can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).
- **Privacy:** Data breaches involving non-public data will violate the right to privacy (A.21, Constitution of India; A.12, UDHR).

Sources of Risk

- **Data limitations:** Challenges in extracting data from non-machine-readable documents with handwriting or abbreviations.
- Training on limited case types which are digitised.
- **Model limitations:** Training limitations for use cases where models are unable to correctly interpret nuances in case documents.
- Models would need additional training and processing power to make human-like inferences around complex legal questions.

USE CASE: Evidence review and analysis

Current state: Conducted manually with inputs from subject-matter experts.

Potential AI use

- **Assess quality of evidence:** Verify evidence provenance and run automated checks for tampering.
- **Time-efficiency:** Reduce time spent on obtaining validation certificates from individual experts and Forensic Science Laboratories.

Types of Risk

- **Incorrect outputs:** Where reliability of forensic science methods are often questioned, there is possibility of the machine making inference errors.

Rights implications

- **Right to fair trial:** Incorrect inferences can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** From non-contextual training data.
- **Model limitations:** Errors in inferences and lack of explainability; overfitting and underfitting.

USE CASE: Drafting: AI enabled drafting of orders, decrees and judgments.

Current state: Drafting is done manually by judicial officers with support from law clerks and secretarial staff.

Some possible shadow AI use by judicial officers, law clerks and secretarial staff.

Potential AI use

- **Time-efficiency:** Faster orders and judgments with less manual work to retrieve and support.
- **Standardisation:** AI can pre-fill in templates with relevant info (for specific case types or stage/purposes).
- **Productivity gains:** Machine-enabled dictation and drafting can reduce dependence on stenographers.

Types of Risk

- AI hallucinations and inaccuracies
- **Reduced scope for application of mind:** Standardisation and drafting with generative AI risks eroding cognitive skills for legal reasoning.
- **Inadequate reasoning skills:** Judicial decision-making is also guided by human consideration of mitigating factors and reasoning that may not be entirely captured in previous decisions.

Rights implications

- **Right to fair trial:** Incorrect, inaccurate outputs can impact the right to a fair trial. (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** Gaps brought on by missing citations and lack of contextual data reflecting Indian law.
- **Model limitations:** Incorrect weighted parameters or lack of adequate parameters for training.

USE CASE: Calculation support for compensation, fines and costs

Current state: Manual
With support of spreadsheet formulae and calculators developed by individual judicial operators.

Potential AI use

- **Time-efficiency:** Automated calculations can save time, by extracting relevant factors for speedier calculations and pre-populating formula.

Types of Risk

- **Incorrect outputs:** Errors in calculation if incorrect data extracted for calculation.
- **Automation bias:** An over-reliance on AI can lead to a tendency to not monitor for errors.

Rights implications

- **Right to fair trial:** Incorrect, inaccurate calculations can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** Incomplete extraction or scoring of relevant factors.
- **Model limitations:** Poor inference training.

USE CASE: Predictive or recommendatory systems for judicial decision-making

Current state: Not in practice.

Potential AI use

- **Reduction in bias:** Reduces scope for human bias in decision-making.
- **Time efficiency:** Potential for faster decision-making and disposal of cases.

Types of Risk

- **Lack of explainability:** For complex tasks like decision-making in courts, which are not mere decision trees, modelling reasoning behind each output is nearly impossible.
- **Discriminatory outcomes:** Potential for recommendations that are influenced by demographic factors and reinforce systemic biases.

Rights implications

- **Right to equality:** Bias and discrimination embedded in the system will violate the right to equality before the law (A.14, Constitution of India; A. 7, UDHR; A.14, ICCPR).
- **Right to fair trial:** Inaccurate reasoning in judgments can impact the right to a fair trial (A.21, Constitution of India; A.10, UDHR; A.14, 17 and 26, ICCPR).

Sources of Risk

- **Data limitations:** Historic biases in judgment data like conservative gender norms and identity-based biases can impact outputs.
- **Model limitations:** Model's inability to embody legal reasoning, ethics, or contextual judgment means that AI-generated judicial recommendations must not replace human judges.

This mapping highlights a key limitation of AI systems: because they are trained on past patterns, they lag behind evolving legal scenarios. When new guidelines are issued, priorities for case scheduling shift, or cases arise in emerging areas of law, both technical systems and oversight protocols require time to adjust. This gap reinforces the need to confine AI use in the judiciary to assistive functions, rather than extending it to substantive or determinative activities.

Governing AI in courts: A rights-and-risk-based framework • Endnotes

88. (Superior Council of the Judiciary, 2024).
89. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L 2024/1689 (2024), annex III.
90. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L 2024/1689 (2024), recital 61. https://ai-act-law.eu/recital/61/.
91. (High Court of Kerala, 2025).
92. (Supreme Court of New South Wales, 2024).
93. Ministry of Justice, GOV.UK. (2025, July 31). AI action plan for justice. https://www.gov.uk/government/publications/ai-action-plan-for-justice/ai-action-plan-for-justice
94. (Courts and Tribunals Judiciary 2025, October 31).
95. Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016).
96. Mesa, N. (2021, May 13). Can the Criminal Justice System's Artificial Intelligence Ever Be Truly Fair? Massive Science. massivesci.com/articles/machine-learning-compas-racism-policing-fairness/.
97. NCSC. (n.d.). AI-generated evidence: A guide for judges. https://www.ncsc.org/resources-courts/ai-generated-evidence-guide-judges
98. NIST. (2023, January). Artificial Intelligence Risk Management Framework (AI RMF 1.0), p. 9. https://doi.org/10.6028/nist.ai.100-1.
99. (NIST, 2023, p.6).
100. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L 2024/1689 (2024), Art 3; (NIST, 2023).
101. Smuha, N. A. (2021). Beyond the individual: governing AI's societal harm. Internet Policy Review, 10(3). https://policyreview.info/articles/analysis/beyond-individual-governing-ais-societal-harm; (De Souza, S.P, 2024).
102. (De Souza, S.P, 2024).
103. (Rangone, N., & Megale, L, 2025).
104. This position is well established by constitutional bench decisions of the Supreme Court of India. See Maneka Gandhi v. Union of India, (1978) 1 SCC 248.
105. (Rangone, N., & Megale, L, 2025, p.1-16).
106. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L 2024/1689 (2024), Art 6; (NIST, 2023).
107. Salgado-Criado, J.; Fernandez-Aller, C;. (2021). A Wide Human-Rights Approach to Artificial Intelligence Regulation in Europe. IEEE Technology and Society Magazine, 40(2) .DOI: 10.1109/MTS.2021.3056284; (Rangone, N., & Megale, L. 2025).
108. BSR. (2025). A Human Rights-Based Approach to Impact Assessment Guide 3 of the Responsible AI Practitioner Guides for Taking a Human Rights-Based Approach to Generative AI, p.6. https://www.bsr.org/files/BSR-A-Human-Rights-Based-Approach-to-Impact-Assessment.pdf; Lu, S. (2024). Regulating Algorithmic Harms Law & Economics Working Papers. 277. https://repository.law.umich.edu/law_econ_current/277.
109. European Data Protection Supervisor. (2025). Guidance for Risk Management of Artificial Intelligence systems. https://www.edps.europa.eu/system/files/2025-11/2025-11-11_ai_risks_management_guidance_en.pdf
110. Rotolo, A., Ferrigno, B., Godinez, J.M., Novelli, C., & Sartor, G. (2025). Foundations for Risk Assessment of AI in Protecting Fundamental Rights. ArXiv, abs/2507.18290.
111. (Rotolo, A., Ferrigno, B., Godinez, J.M., Novelli, C., & Sartor, G, 2025)
112. See also Janssen, H., Seng Ah Lee, M., & Singh, J. (2022). Practical fundamental rights impact assessments. International Journal of Law and Information Technology, 30(2), 200–232. https://doi.org/10.1093/ijlit/eaac018.
113. Justice K S Puttaswamy v. Union of India, (2017) 10 SCC 1.

4

Assessment Framework

Key Takeaways

- **AI adoption in courts is a governance decision, not just a technical one:** Decisions to deploy AI affect judicial processes, fundamental rights, and public trust and hence require transparent and accountable evaluation frameworks.
- **A stage-wise assessment framework is essential across the AI lifecycle:** Courts should assess (i) institutional readiness, (ii) rights-based risks at the use case level, (iii) technical robustness and vendor practices, and (iv) post-deployment performance and harms through continuous monitoring.
- **Institutional capacity is foundational:** A permanent IT and data cadre, headed by CIO/CTO-level leadership, is essential to ensure continuity, institutional memory, vendor accountability, and rights-based AI governance.
- **Readiness must precede adoption:** Courts must assess regulatory preparedness, infrastructure, data quality, staff capacity, and safeguards before deploying AI, recognising AI as a structural shift in justice delivery.
- **Problem-first, not tool-first:** AI should only be adopted where clearly justified; courts must avoid "shiny object syndrome" and assess alternatives before automation.
- **Risk assessment must be rights-centred:** Evaluations should account for use case, case type, and litigant vulnerability, recognising that harms may be irreversible in liberty, privacy, or life-affecting cases.
- **Technical scrutiny is non-negotiable:** Vendors must disclose ownership and funding, document model design and performance, enable plain-language explanations, maintain audit trails, allow human override, and support third-party audits to protect judicial independence and open justice.
- **Strong data governance and risk controls are essential before deployment:** Courts require clear data ownership rules, strict privacy and security safeguards, bias testing and harms modelling, and sandbox testing with defined thresholds to prevent inaccurate, discriminatory, or rights-violating outcomes.

- **Oversight must be continuous:** Ongoing monitoring, KPIs, human override mechanisms, and safety stops are critical to prevent automation bias and systemic harm over time.

As Indian courts explore AI tools for various functions, it is essential that they adopt a structured framework to evaluate these technologies. The following questions are designed to guide courts through their journey of AI adoption, including assessing the court's readiness, the suitability of AI tool for judicial use, and oversight after adoption. Importantly, seeing as AI deployment is not merely a technical decision, but one that impacts court processes, citizens' rights, and public trust in the judiciary, these questions are designed to support informed, cautious, and accountable decision-making.

Drawing from established practices in AI adoption lifecycles, we organised questions across three critical stages:

- pre-adoption assessment of (a) institutional readiness and (b) functional needs,
- adoption-stage evaluation of vendors and their tools, and
- post-adoption testing and monitoring.

Table 4: Assessment Frameworks		
Institutional readiness assessment	A set of questions to help the court decide whether its current human resources, infrastructure and finance considerations are adequate for the successful design, deployment, and monitoring of AI tools.	This assessment should be conducted by the court at a **generic level prior to adoption**
Risk assessment	A mechanism for courts to identify the potential risks of AI use in judicial processes. The assessment will help courts determine whether or not to proceed with deploying AI for the intended purpose and if so, with what safeguards.	This assessment should be conducted by the court at a **functional use case level prior to adoption**
Technical assessment	An assessment by the court and the vendor to help the court to understand the proposed AI tool. This detailed vendor assessment questionnaire that systematically examines vendor credentials, technical capabilities, data governance practices, transparency measures, safety protocols, and accountability mechanisms.	This assessment should be completed by a solution provider at a **specific tool level as part of adoption/ pilot planning**
Ongoing and Continuous assessment	Questions for the court and vendor to continuously monitor impact and determine success metrics of the tool once adopted.	This assessment should be periodically conducted by the court (with information provided by the solution provider) at a **project level after adoption of AI**

The next sections outline the structured assessment frameworks designed to guide courts through critical evaluation stages of AI adoption, building on the themes discussed above. The report is accompanied by **downloadable spreadsheets that will enable courts to adapt or conduct these as self-assessments** to identify both safeguards and mitigation strategies.

Each section uses a scoring matrix that translates qualitative responses into quantitative metrics, yielding classification recommendations and guiding next steps. This structured approach allows courts to move from abstract principles to concrete assessments, ensuring that AI adoption decisions are transparent, systematic, and grounded in documented evaluation of institutional readiness, project needs, and vendor capabilities.

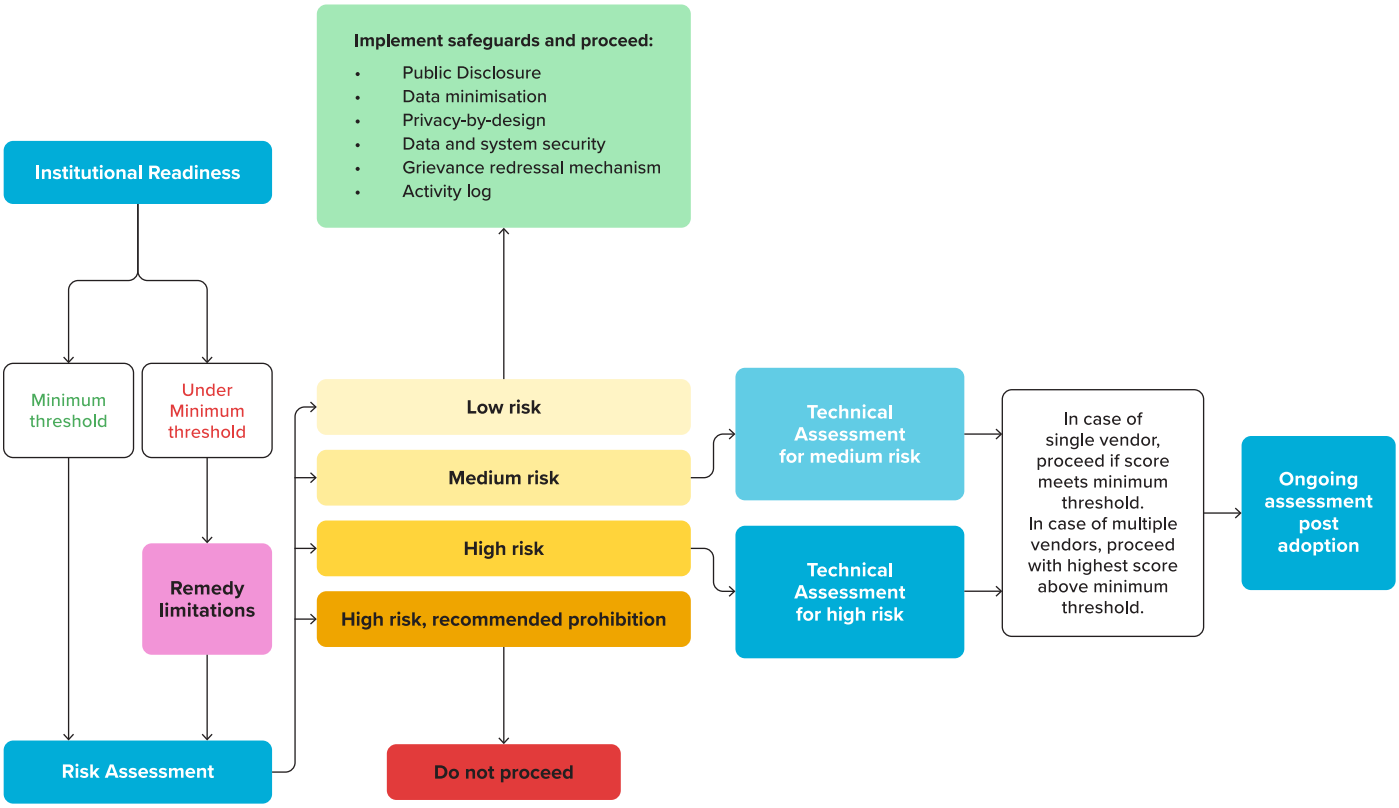
How to use the assessment tools

The assessment tools may be used either when a single vendor is under consideration for a particular use case or when the court is choosing among multiple vendors. They are intended to help courts evaluate their institutional capacity, assess potential risks to rights arising from use of the AI tool, and identify mitigation strategies to be put in place by the vendor.

We recommend courts begin with the institutional readiness assessment. Courts that score optimally, reflected in positive response to all essential questions a score of at least 60%, may proceed to the next step. If the score falls below this threshold, the court must remedy limitations before proceeding to the risk assessment.

The risk assessment enables courts to classify the intensity of risk associated with the tool. The subsequent action should be calibrated accordingly. For low-risk tools, deployment may proceed with basic safeguards in place. Medium-risk tools should trigger non-negotiable assurances from vendors in the technical assessment. In the case of high-risk tools, we recommend that courts undertake the full technical assessment and proceed only in case of a reasonably high score (>=60%). In case of high risk with prohibition, courts should refrain from use AI, as the risks outweigh potential benefits.

Finally, once AI is deployed and ready for adoption, the monitoring assessment tool provides supervisory questions to keep the AI system in check and mitigate harms over time.



4.1 Who should conduct AI assessments in courts?

As courts increasingly integrate AI and other data-driven technologies, the need for specialised personnel to manage digital systems and court data becomes more acute. This need will intensify under Phase 3 of the eCourts project, which introduces more advanced digital infrastructure. Yet most courts rely on a small number of staff with minimal technical qualifications, rather than on formally recruited cadres with a stable professional pathway[114].

A **dedicated IT and data cadre** in High Courts and district courts is therefore necessary, following the examples of Kerala, Patna, Allahabad, Uttarakhand, and Manipur High Courts[115].

- This cadre should comprise professionals with varied technical and data competencies, supported by regular capacity building to keep pace with emerging technologies.
- At the High Court level, such a cadre should be headed by an official equivalent to a Chief Information Officer (CIO) or Chief Technology Officer (CTO) in a company, reporting to the Chief Justice's nominee or the Computer Committee.
- Designated implementation and compliance officers should support the CIO/CTO in overseeing the rollout of systems and ensuring compliance with ethical frameworks.
    - As and when a "Judicial Digital Data Management Policy"[116] is brought into effect, High Courts will also require an Implementation Officer and Compliance Officer.
    - The cadre could include systems architects, enterprise engineers, software engineers, data scientists, AI/ML specialists, cybersecurity officers, and AI safety and ethics officers, among others.
    - Dedicated staff will also be required at district and trial court complexes to oversee systems administration and user support.
- To account for variations in capacity across High Courts, the cadre framework should permit flexible hiring models, including outsourcing or contractual appointments where needed[117].

A stable technology cadre would have multiple benefits from:

- establishing institutional memory, lifecycle planning, and ability to build resilient, future-proof systems,
- ensuring continuity, maintenance, troubleshooting, and iterative improvement of technology development,

- granting courts autonomous capability to evaluate, negotiate, co-develop with, supervise, and hold vendors accountable
- building trusted, in-house capacity for AI governance and data stewardship, and
- providing robust support, while enabling judges to remain final decision-makers.

4.2 Institutional Readiness Assessment

Before deployment of an AI tool, the court must assess whether the problem it seeks to address is clearly defined, and whether adequate regulatory frameworks, institutional capacity, infrastructure, data, and monitoring mechanisms are in place to support its safe and responsible use. This is necessary because AI adoption in the judiciary represents a fundamental shift in how courts process information, make decisions, and deliver justice. This framework enables courts to evaluate their readiness considering regulatory scaffolding, institutional infrastructure and knowledge base required to support AI use. In addition, foresight and planning can ensure AI serves the interests of justice.

The questions for courts cover the following themes:

Regulatory framework
Before adopting AI tools, courts should establish clear policies governing their use. This includes creating frameworks that define permitted and prohibited AI uses, set rules for handling sensitive data, establish disclosure requirements to litigants about AI use, and create grievance redressal mechanisms. The framework should also provide for an oversight mechanism supported by experts in technology, data governance, information design, and procurement.

In-house capacity
Given that AI differs significantly from traditional software, courts need comprehensive training programmes for judges and staff on AI fundamentals, limitations, risks (including hallucinations and bias), and responsible use. Training should be integrated into the curricula at judicial academies and entry-level programs with periodic refreshers.

Technical Infrastructure

In terms of digital infrastructure, courts must ensure adequate hardware (computers, reliable internet, uninterrupted power, secure encrypted networks), high-quality audio/video equipment for transcription services, and appropriate data storage infrastructure. Critically, court data must also be digitised and maintained in machine-readable formats.

Problem definition

It is vital that courts articulate the problems they seek to address, and assess whether AI is the most appropriate solution as opposed to stronger data analytics or process reforms. Courts should guard against "shiny objects syndrome", where AI is assumed to be the default solution. Courts must also survey the market to determine whether commercial, off-the-shelf solutions exist or if customised development is required. For emerging technologies, Expression of Interest (EOI) documents should focus on outcome-based requirements rather than prescribed technical inputs, allowing vendors to propose innovative solutions.

Performance metrics and Impact

Courts must prepare themselves to define success with specific, measurable outcomes, such as reduction in case backlogs or decreased disposal times, and establish baseline measurements of current metrics before AI deployment to enable meaningful evaluation. Key Performance Indicators (KPIs) should be linked directly to identified objectives and supported by a robust monitoring and evaluation framework that allows for timely adjustments.

Courts should also assess who will be affected by the AI tool, the duration and reversibility of impacts, and potential risks to rights, equality, dignity, privacy, and procedural fairness. Courts also need to identify the necessary policy, rule, or process changes required to support AI adoption, including amendments to data policies, procedural rules relating to e-filing, and authorisation of digital processes.

4.3 Risk Assessment

While considering AI deployment for a specific use case, courts must assess whether the task is suitable for automation, identify potential risks to rights, and evaluate whether specific case or litigant types may be particularly vulnerable to harm.

Suitability of the AI system for the identified use case

AI regulation experts caution that public institutions like courts tend to treat AI as a one-stop solution to address institutional issues[118], especially as Indian courts are challenged by backlogs and bottlenecks in case disposal. Courts might pilot AI in use cases without examining whether AI is the most appropriate solution given the context. In turn, widespread deployment may cause courts to overlook tensions between AI systems and fundamental rights, as well as their own institutional readiness to deploy AI in a manner consistent with these rights.

Applying the proportionality principle can harmonise some of these tensions[119]. Proportionality requires courts to clearly articulate the goals to be achieved in a particular use case, examine the alternatives to AI in meeting these goals, and explain why AI serves as the best possible solution[120].

Thus, the primary question courts must ask themselves at the outset is this: is AI necessary, or the only means for a use case?[121] What problem is the court seeking to resolve, and how might AI assist in the same?

Nature of AI system required for the identified use case

The proposed risk assessment framework identifies model design as a key source of potential harms to fundamental rights. General-purpose tools that are neither domain-specific nor tailored to court requirements can pose significantly higher risks of inaccuracy, contextual misunderstanding, and bias. This is particularly so when courts have little control over the datasets used for training, confidentiality safeguards, and other compliance standards. Such limitations can have serious consequences, including the compromising of private information, an increased likelihood of errors, and reduced capacity to correct biases and discriminatory patterns in the dataset.

Case and litigant type controls

AI systems can disproportionately harm litigants from socially and economically vulnerable backgrounds. Further, harms caused by errors and bias in certain case types or proceedings, such as loss of privacy and inaccuracies may even be irreversible or challenging to undo.

Dataset quality and ownership

An AI output is only as good as its dataset. Biased and discriminatory outcomes are linked to the quality of the dataset, which is determined by i) who has created or provided the dataset and ii) whether the dataset has been treated and tested for discrimination, errors and domain context. Questions of data ownership and data security are also closely tied to the protection of judicial data and of litigants' privacy. For instance, granting vendors access to data in the absence of cybersecurity measures could have serious implications for fundamental rights. In such cases, personal data may be particularly vulnerable to compromise.

Safeguards proposed

Safeguards such as human oversight of AI outputs primarily address case-specific harms. However, mechanisms for aggregate review of system performance with a focus on patterns of errors and bias are necessary to mitigate longer-term harms.

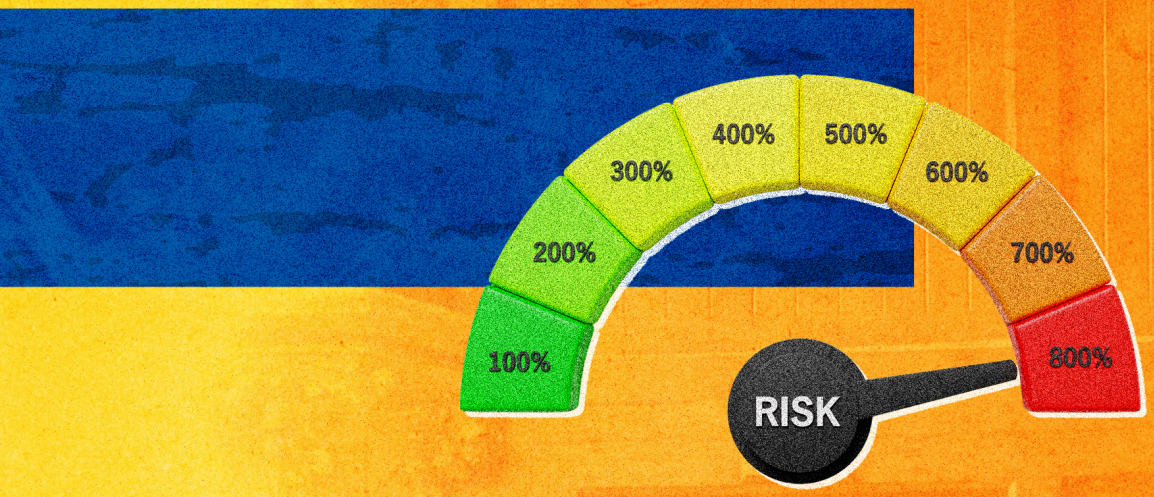
Categorisation of risks

Human rights-based approaches to AI governance reject a simplified division[122] of systems as simply "high-risk" and "low-risk." Instead, they advocate for multiple tiers to assess risks[123]. A multi-tiered approach better accounts for how harm may vary depending on the interaction between AI product, the specific use case, and both case and litigant type.

In line with recommendations from rights-based commentators[124], this report recommends a four-tiered framework to categorise AI harms in courts, recognising the variable impact of AI on rights and identifying specific use cases to prohibit the use of AI. Table 5 illustrates a scale and the rationale for the different risk levels.

Table 5: Risk categories

Risk level	Framework in India
Low risk	Presents minimal to no risk for users or intended beneficiaries of the AI tool
Medium risk	Increased likelihood of negative impacts such as rights violations which can be mitigated with good practices
High risk	Higher likelihood of negative impacts which require robust and stricter mitigation measures
High risk recommending prohibition	Assured likelihood of harm irrespective of safeguards


100%
200%
300%
400%
500%
600%
700%
800%
RISK


DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER
DO NOT PROCEED FURTHER

4.4 Technical Assessment

Assessments so far have focused primarily on the institutional context of courts, including readiness, capacity and proactive governance. As courts move towards testing and deployment, scrutiny must shift to the specific AI tools under consideration and the entities developing them.

An assessment completed by solution providers as part of due diligence can enable predictability in evaluation, guide questioning during demos or live presentations, and enable meaningful comparison across vendors, regardless of procurement mode. Such an assessment supports informed decision-making about the suitability of a particular tool. Questions in this assessment focus on model transparency and documentation, performance tracking mechanisms, and how the tool may support appropriate oversight protocols.

Transparency

Vendors must disclose their investors, funders, and organisational structure to safeguard judicial independence and prevent conflicts of interest. Courts should be informed of vendors' political interests and financial ties to industries and parties that could undermine neutrality. For nonprofits, this includes transparency around funder identity and sustainability of funding pipelines. For commercial entities, this includes disclosure of sources of funding, including private equity and business revenue.

Data

Vendors must specify what additional data they need from courts beyond what is in their possession, explaining data sources, quality, and procurement methods. Where court data is used, vendors must demonstrate how they address data limitations and ensure representativeness across case types, litigants, and geographic regions. Any use of synthetic data must be disclosed, along with its provenance and limitations.

Vendors must also address some critical ownership questions: Who owns training data, and who owns data created during the project lifecycle? They must disclose whether they will retain rights to aggregated or anonymised insights and whether court data will be used to train commercial models deployed elsewhere. Finally, vendors must detail privacy and security measures in place, including encryption, role-based access controls, and secure storage procedures, with attention to sensitive data about minors, survivors of sexual violence, medical records, and confidential financial information.

Explainability

Vendors must provide courts with detailed documentation on the algorithms and models used, including the choice of variables and AI techniques used (such as supervised, unsupervised, or reinforcement learning). They should inform courts about whether they are compliant with recognised standards like ISO 42001 or SOC 2, and such compliance must be substantive rather than symbolic. Documentation provided to courts must be sufficiently detailed to enable independent auditing, validation, and maintenance without vendor dependency.

Given the need for open justice, AI tools used in courts must be explainable. This means that vendors should be able to justify how their tools arrive at decisions in terms understandable to laypeople. While full technical explainability may not be feasible for complex deep learning models, vendors should provide to courts human-understandable explanations of the model's logic, parameters, and performance to the extent possible. Where intellectual property constraints limit transparency, courts should consider requesting custom algorithm development or treating the tool as unsuitable for judicial deployment.

Bias testing and harms modelling

Vendors must document their processes for testing datasets for biases and unexpected outcomes, maintaining data cards, model cards, and audit documentation. They should explain the mitigation strategies used to address low-quality or unreliable datasets, including data augmentation and resampling when gaps or bias risks are identified.

It is also necessary that vendors conduct comprehensive harms modelling for the proposed tool. This involves identifying, assessing, and addressing potential negative outcomes, including violations of privacy, equality, and procedural fairness across different stakeholder groups. Harms modelling should specify the categories of bias tested, the likelihood and severity of biased outputs, stakeholders most at risk, frequency of potential harms, and whether users will be able to recognise bias when it appears. The assessment should consider both individual and systemic impacts, including dehumanisation, public shaming, loss of effective remedy, interference with private life, forced association, digital exclusion, economic loss, and social detriment.

Vendors must further explain how their systems handle outliers and edge cases, particularly those involving underrepresented groups whose needs they may not have considered. They should demonstrate comprehensive risk mitigation plans, detailing concrete measures for preventing, detecting, and responding to identified risks.

Audit

Vendors must maintain audit trails recording all outputs, clearly identify key decision points, and document all changes to the model over time. AI tools should generate explanations for outputs in plain language, and audit trails must specify the exact tool version used for each decision as well as the identity of the authorised decision-maker. Vendors also need to establish robust mechanisms to collect user feedback, enable processes for contesting AI-generated outputs, and ensure human override capabilities.

Vendors must agree to periodic third-party audits (especially crucial for frequently updated generative AI models) and provide timely, up-to-date documentation and record maintenance. The tool should incorporate privacy-by-design principles from the concept stage, with clear explanations about whether the tool operates offline/isolated or relies on external cloud services and APIs, and if it incorporates appropriate safeguards against data leakage.

Testing

Vendors must agree to sandbox testing before large-scale deployment for medium and high-risk tools. They should outline proposals for testing duration, scale, protocols, and use of representative dummy datasets. Vendors must also define acceptable thresholds and confidence scores required for wider deployment and explain how success will be evaluated against baseline metrics.

In addition, vendors must provide courts with comprehensive technical documentation, instructions for use, automatic event logging, and secure, accessible logs for authorised personnel. The service of auditors is also necessary to enable independent court oversight without forced vendor dependency.

4.5 Ongoing and continuous assessment

Technical partners must assist courts in developing protocols to monitor the performance of tools over time, through traditional KPIs to track how their AI tools improve, adapt, or degrade over time. This should include evidence of performance in both controlled testing environments and real-world pilots, while acknowledging that actual courtroom conditions can be complex and differ from controlled settings.

Safety stop mechanisms are essential in AI tools. These include kill switches for emergency intervention, blinding techniques to prevent learning from sensitive variables (like caste), and boxing techniques to isolate systems from external networks. For reinforcement learning systems particularly prone to producing harmful outcomes, vendors

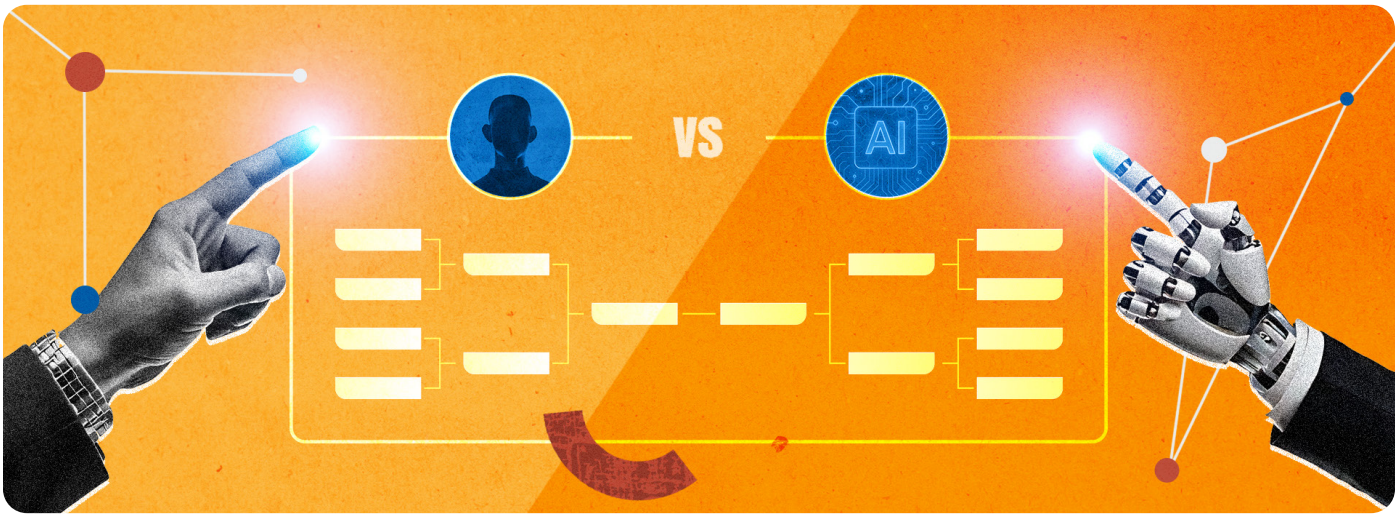
must explain techniques to ensure agents generalise well without overfitting. Vendors should also clarify whether systems have automatic feedback / retraining loops or human-in-the-loop review processes.

Evaluation through benchmarking

The process of benchmarking involves identifying tasks that enable comparison of the relative performance of AI systems. These may include simple side-by-side comparisons of easily available tools on the same function; for example, entering domain-specific prompts into ChatGPT, Claude, and DeepSeek to compare accuracy, comprehensiveness, or merit of generated responses.

Existing benchmarks have been criticised for relying on broad "anglo-centric and domain-agnostic" parameters. In response, recent efforts have focused on developing more diverse datasets, tests, and formal evaluation objectives to facilitate more "culturally and contextually aware evaluation"[125]. These include tasks that are multilingual or involve lesser-known dialects and hybrid languages, as well as questions that are both domain-specific and jurisdiction-specific. As courts begin using AI more across use cases, it is important to build a protocol for benchmarking, so that tools can be built using the best-performing models for an Indian judicial context.

Teams have developed structured technical evaluations for models that collect detailed performance data and assign weights to compute scores. Beyond quality, speed, and cost, such rankings may also factor in parameters like domain-specific skills, memory usage, or safety in order to support informed model selection. In addition, AI-generated outputs may be scored by expert human evaluators for relevance and grounded thinking.



4.6 After the assessment: Mitigating and responding to risks and harms

AI use in court settings carries the risk of rights violations. To ensure AI systems improve the quality of delivery of justice, it is imperative to invest resources in capacity building, continuous performance monitoring, and need-based correction of new systems.

Below, we highlight some good practices, including both mitigation and response strategies, that courts (judicial officers and court staff) can adopt to manage risks arising from AI use. Recognising that certain limitations, like hallucinations in large language models, cannot be fully addressed at the application level, we recommend sustained vigilance and consistent communication with vendors and developers. Actions marked [C] indicate activities to be undertaken by the court and those marked [D] indicate responsibilities of developers and vendors. Where the court itself functions as the developer and can undertake checks (e.g., assessing data representativeness), we use [D].

Data limitations: Drawbacks in the training dataset

1. Data entry errors	1.1 Data documentation standardisation (e.g., datasheets with pre-defined formats; rule-based detection of anomalies) [D]
2. Gaps in data: Lack of representation, context, non-digitised data	2.1 Legal language glossary for machine learning [C]
	2.2 Representativeness checklists [C]
	2.3 Data audit to check for representativeness [D]
	2.4 Edge case checklist for context [C] + [D]
	2.5 Systematic digitisation of legacy records based on case activity (prioritise pending cases), legal importance (prioritise constitutional law and personal liberty cases) and physical condition of records [C]

[C] Action to be undertaken by the AI-adopting court
[D] Action to be undertaken by the AI developer

3. Annotation errors	3.1 Clear annotation instructions [D]
	3.2 Peer review process for reviewing annotations [D]
	3.3 Implementation of adequate quality controls [D]
	3.4 Annotator training to address annotator bias [D]
	3.5 Maintain diversity and expertise in annotator pool [D]
4. Data does not capture legal syllogisms	4.1 Consult with domain expert [D]
	4.2 Allow descriptive data capture [C]
5. Biases in data	5.1 Data audit to check for bias [D]
	5.2 Bias checklist [D]

Model limitations: General issues at the model training and deployment stages

1. Generalisation / overfitting	1.1 Test for overfitting through cross-validation techniques [D]
	1.2 Pruning or feature selection to eliminate irrelevant parameters [D]

[C] Action to be undertaken by the AI-adopting court
[D] Action to be undertaken by the AI developer

	1.3	Increase diversity of data – e.g., types of cases in the dataset [D]
2. Hallucinations	2.1	Be vigilant – log hallucinations [C]
	2.2	Seek information from model developer – strategies to mitigate risk of hallucinations [C]
3. Developer constraints: Biases of the developer, incorrect assumptions and variables	3.1	Conduct fairness evaluation [D]
	3.2	For incorrect assumptions and variables, conduct back testing, stress testing and/or sensitivity analysis [D]
4. Model drift	4.1	Continuous monitoring of model performance [D]
	4.2	Report identification of model drift and what is causing it [D]
	4.3	Continuous retraining strategy [D]
5. Lack of explainability	5.1	Where possible, use an interpretable model, especially for rule-based applications. Interpretability is a prerequisite for explainability. [D]

[C] Action to be undertaken by the AI-adopting court
[D] Action to be undertaken by the AI developer

5.2	Maintain model cards documenting feature importance [D]
5.3	Use model-agnostic interpretation tools [D]

Other sources of risks and harms

1. Misuse or manipulation	
1.1	Filters to prevent misuse – that detect and block certain prompts [D]
1.2	Ignore harmful prompts [D]
1.3	Monitoring-based restrictions [D]
2. Privacy violations	
2.1	Anonymisation techniques and safeguards to prevent re-identification [D]
2.2	Compliance with applicable laws and standards [D]

[C] Action to be undertaken by the AI-adopting court
[D] Action to be undertaken by the AI developer

The steps outlined above depend on sustained cooperation between courts and the entity that has developed AI systems and trained their underlying models. In general, some good practices include:

- maintaining thorough documentation, including datasheets and model cards that are updated regularly,
- continuously monitoring model performance against agreed metrics to detect performance "drift" or degradation over time,
- maintaining detailed logs of errors, hallucinations, and failure rates,
- ensuring proactive and continuous communication with vendors or developers, treating them as partners in building India's judicial capacity and infrastructure and not one-off service providers, and
- checking the tool for compliance with applicable laws and standards such as ISO 27001 and SOC2 for AI system risk management and data management.

On the institutional side, to protect the rights enumerated above in addition to our guiding principles like transparency, do no harm, fairness and non-discrimination, data protection, and human oversight and accountability, we propose the following safeguards to be considered across various use cases and risk categories:

- Domain expert consultation during AI tool R&D
- Clear public disclosure requirements
- Human-in-the-loop (with legal/substantive domain knowledge) ensures that no AI output is applied without human approval
- Mandatory oversight requires judicial, registry, or counsel verification of AI-generated summaries, translations, or recommendations before use
- Data minimisation and privacy-by-design rules
- Enabling litigants (and their lawyers) to opt out or restrict the use of AI in their proceedings
- Dedicated complaint/redress mechanisms for reporting AI-related errors
- Seeking technical expertise on possible cybersecurity and data security measures that should be developed
- Bias and fairness auditing requirements
- User-based audit logs that log all AI interactions for accountability, so they may be reviewed and/or challenged when needed

This list is not exhaustive and should be understood as a minimum set of good practices. Courts are encouraged to adopt additional safeguards and strengthen their capacity to implement and monitor them over time.

Assessment Framework • Endnotes

114. TJ, Personal Communication, 12 August 2025.
115. National Court Management System Sub-Committee. (2024). Baseline Report on Human Resource Development Strategy in the District Judiciary. https://cdnbbsr.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/uploads/2024/11/2024111282-1.pdf , p. 13
116. Ibid.
117. (National Court Management System Sub-Committee, 2024).
118. (Hickok, 2024, p.52).
119. UNESCO. (2023). Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000386276. p. 10.
120. Ibid.
121. (Hickok, 2024, p.52).
122. European Center for Not-for-Profit Law (2021). Evaluating the Risk of AI Systems to Human Rights from a Tier-based Approach. https://ecnl.org/sites/default/files/2021-06/Evaluating%20the%20Risk%20of%20AI%20Systems%20to%20Human%20Rights_ECNL%20proposal.pdf; Canales, M.P., Barber, I., Rowe, J., (2023, September 19). What would a human rights-based approach to AI governance look like?. Global Partners Digital. https://www.gp-digital.org/what-would-a-human-rights-based-approach-to-ai-governance-look-like/.
123. (European Center for Not-for-Profit Law, 2021); (Canales, M.P., Barber, I., Rowe, J., 2023); Hidvegi, F., Leufer, D., Masse, E., (2021, February 17). The EU should regulate AI on the basis of rights, not risks. Access now. https://www.accessnow.org/eu-regulation-ai-risk-based-approach/.
124. (European Center for Not-for-Profit Law, 2021).
125. Devane, V., Nauman, M., Patel, B., Wakchoure, A.M., Sant, Y., Pawar, S., Thakur, V., Godse, A., Patra, S., Maurya, N., Racha, S., Singh, N.K., Nagpal, A., Sawarkar, P., Pundalik, K.V., Saluja, R., Ramakrishnan, G. (2025). BhashaBench V1: A Comprehensive Benchmark for the Quadrant of Indic Domains. Arxiv. https://arxiv.org/html/2510.25409v1

5

Moving Forward: Priorities for Courts

Key Takeaways

- AI can support courts and improve justice delivery, but only when adopted carefully and with strong safeguards.
- The report proposes a four-part adoption process: assessing institutional readiness, identifying rights-based risks, evaluating technical quality, and monitoring AI tools throughout their lifecycle.
- Responsible AI use must be backed by robust contracts, clear rules for data use, and practical guidance for everyday use by judges and court staff.
- Courts need supporting institutional structures, including dedicated AI Committees, trained technical cadres, and a central register to track all AI tools and pilots.
- Moving away from ad hoc experiments toward structured testing, sandboxes, and time-bound pilots will improve accountability and learning.
- With sound governance, AI can strengthen efficiency, fairness, and public trust; without it, AI risks deepening inequality and undermining judicial legitimacy.

Responsible adoption of AI in courts requires a structured framework that involves systematic assessments by both courts and technology vendors and is anchored in sound contractual safeguards. We recommend a four-part assessment framework encompassing the Institutional Readiness Assessment (evaluating capacity and infrastructure), the Risk Assessment (identifying harms at the use case level), the Technical Assessment (evaluating technical aspects of the AI tool), and Ongoing/Continuous Assessment (monitoring impacts over time). To protect judicial independence and sensitive data, engagements with IT vendors must be governed by robust contractual agreements. Minimum contractual clauses are documented in Annexure 1. Recognising that AI use can extend beyond court-directed tools, especially given the increasing ubiquity of easy-to-use generative AI tools, we provide a guidance list for responsible use of AI in day-to-day work in Annexure 2.

However, assessment tools alone are insufficient. These frameworks and contractual protections will be effective only when embedded within broader institutional structures in courts that support them. As Indian courts

move from early experimentation towards deeper institutional engagement with AI, the need for responsible AI adoption becomes increasingly clear. This requires reimagining technology governance structures within courts, establishing mechanisms that ensure transparency to advocates and litigants, and investing in ongoing capacity building for informed use. Without this scaffolding, AI use will remain fragmented, personality-driven, and vulnerable to misuse, over-reliance, and public distrust.

At the institutional level, formal oversight mechanisms are necessary. Adequately empowered **AI Committees** in High Courts and the Supreme Court should sit at the centre of the new governance architecture, and these committees must go beyond ad hoc advisory roles and take responsibility for the full AI lifecycle. Their roles should include defining acceptable and prohibited use cases, establishing safeguards and standards, managing partnerships with AI vendors, coordinating capacity building efforts for judges and staff, and commissioning audits.

For these committees to function credibly, they should be balanced and multidisciplinary, combining expertise of judges with technologists, data protection experts, AI ethicists, and other experts. Gender and age diversity are also essential to reduce cognitive bias and ensure sensitivity to the varied needs of court users. Institutionally, these AI Committees must have a clearly defined relationship with existing Computer Committees and the eCourts Committee. While Computer Committees focus on infrastructure, digitisation, and IT systems, AI Committees must function separately, ensuring that AI governance is treated as a pathway towards innovation rather than routine technical upgrades.

Additionally, these Committees require the support of skilled technology cadres within the judiciary. Given the push to make courts natively digital, this cadre is vital for ensuring that technology is introduced thoughtfully and with necessary safeguards. Establishing a defined career progression pathway, together with opportunities for regular upskilling, will help retain skilled professionals within the judiciary and build the institutional memory needed for effective long-term governance.

Internal capacity must be complemented by system-wide coordination and visibility. To prevent fragmented experimentation and uneven adoption, the Supreme Court eCommittee should consider creating a **central Judicial AI Register** that documents all AI initiatives across courts in India, including ongoing pilots, proposed projects, discontinued efforts, or full-scale adoption. This register should record the tools in use across India, their specific use cases, details of vendors, collaborators, and funding models (including Public-Private Partnerships), details of contractual arrangements with vendors, sandbox and pilot outcomes, and documentation of risks and remedial actions. A public-facing version of this register (excluding sensitive data) could enhance transparency and public trust. An internal version could support cross-learning among courts, enable benchmarking of maturity levels, and inform evidence-based decisions on scaling AI use.

Given the significant technological shift represented by AI, **structured capacity building** for judges and court staff is imperative. Such training must be institutionalised and ongoing rather than episodic. It should cover a technical understanding of AI; ethical risks and constitutional rights implications; relevant safeguards, and mitigation strategies; responsible-use norms; and practical guidance on how AI should and should not be used.

Courts must also transition from informal pilots to a **structured pipeline for innovation.** Drawing on established procurement processes in public services, from presentations of early prototypes by developers, courts should consider engaging in a collaborative process where vendors can refine the tool with feedback from the court. Courts should also experiment with sandboxes, which are controlled testing environments in which tools are first evaluated on synthetic data and subsequently on real-world data, to measure performance, ensure ethical compliance, and test explainability. Eventual pilots should be time-bound and jurisdiction-specific with clearly defined success metrics, reporting obligations, and feedback loops. The AI Committee should develop testing and evaluation protocols that define acceptable thresholds and confidence scores for deployment at scale.

Crucially, these processes must be designed to endure beyond individual tenures. The deployment of AI in courts must be backed by **long-term planning** and adoption frameworks that outlast transfers of judges. Without multi-year strategies, AI adoption will remain vulnerable to discontinuity, perpetual pilots, and dependence on vendors. Structured adoption processes, strong institutional memory, and continuity in contractual arrangements are essential to transform isolated experiments into reliable digital infrastructure and scalable reforms.

AI must be adopted in courts with care, but caution should not translate into technological paralysis. If governed well, AI can improve processes and make courts more effective. If governed poorly, it risks entrenching inequality and weakening the credibility of the judiciary. The measures recommended in this report emphasise cautious adoption while enabling stronger and more relevant implementation of AI tools that target key pain points in the judicial system. Courts must remain open to informed experimentation, draw on global experience, and adapt to the evolving technological landscape, while remaining firmly anchored in constitutional values and their duty to deliver justice.

Annexures

Annexure 1: Contractual terms

Data protection

1. The agreement should eliminate any ambiguity that could allow the vendor to use court data, whether directly or indirectly, for AI training, analytics, or commercial development, or for integration into external systems. No other use, whether analytical, developmental, or derivative, should be permitted. The emphasis should be on safeguarding judicial data from repurposing or external use.

2. The vendor should disclose its data retention policy, including period of storage, storage platform and access controls.

3. Where necessary (e.g., upon completion of the project period), the vendor should be obligated to securely delete all court data (including backups and derivative copies) within a defined number of hours and provide a formal certificate confirming permanent and secure deletion.

4. The contract should mandate full compliance with the Digital Personal Data Protection Act, 2023 (DPDP Act) and its accompanying Digital Personal Data Protection Rules, 2025 (DPDP Rules) and any applicable data localisation laws and regulations in force at the time of execution and throughout the term of the agreement.

Safety and security

1. The agreement should require the vendor to ensure robust privacy and security practices. The court may request documentation or evidence of periodic independent audits or assessments of the vendor's privacy/security controls, to ensure ongoing compliance and accountability.

2. The vendor should be obligated to maintain end-to-end encryption and security standards consistent with applicable Indian data protection laws, ISO/IEC 27001, and any government-prescribed or subsequently notified technical standards.

3. The vendor should provide documentation for model design, training data provenance, bias/fairness mitigation measures, impact assessments, human oversight for decisions where appropriate, and, where feasible, explainability or output rationale, especially for decisions affecting individuals (e.g., classification, inference, recommendation).

Audits and incident reporting

1. A standardised incident-reporting protocol should be defined, specifying timelines (e.g., within six hours of a serious incident, consistent with CERT-In Directions), escalation levels, reporting content, and supporting documentation. The vendor should also cooperate with regulatory bodies and support post-incident forensic analysis.

2. The vendor must maintain detailed logs (traffic, access, modification, deletion) for a specified retention period (e.g., at least one year after completion of processing), including immutable, tamper-proof logs, with secure storage, and accessible for audit.

3. The agreement should provide for both periodic and for-cause audits, including in situations of breach or suspected non-compliance. Audit scope should expressly cover data localisation, encryption controls, integrity of logs, and security practices, not merely IT/operational controls. The vendor should be required to implement corrective measures within a defined timeframe and at the vendor's own cost, as agreed upon by parties.

Ownership and intellectual property

1. The contract should clearly define ownership of any tools, source code, datasets, algorithms, or models developed during the engagement, and should expressly state the scope and extent of how such intellectual property (IP) relates to the court or resulting from the engagement vests exclusively in the court. The vendor should not assert residual rights, nor reuse any court-related IP, models, or configurations in commercial products or deployments.

2. The vendor should be required to provide periodic disclosures of ownership and control throughout the duration of the contract, in order to pre-empt risks arising from change of control, mergers, acquisitions, or restructuring activities.

3. Any subcontracting (including non-core services with identifiable data exit and entry points such as UX development) should require prior written approval from the court. The vendor should remain fully liable for all subcontractors and should ensure that all confidentiality, localisation, data processing, and audit obligations flow down contractually to such parties. A current subprocessor register should be maintained and made available for court review.

4. Confidentiality obligations should survive termination of the contract and continue indefinitely, with no end term, including after the completion of the project.

Annexure 2: Court User Guidance for AI

In some courts, use of only approved AI tools is allowed; however, in many others, they are not sanctioned but not banned and many individuals are using them to streamline their work. These are some good practices for the latter category.

The use of generative AI is set to increase, with wider access and improving performance in day-to-day tasks. Given considerations of privacy, judicial impartiality and non-discrimination, it is essential that court users relying on LLM-based applications approach use responsibly. Drawn from guidelines regulating the use of AI in other jurisdictions, this document offers a list of dos and don'ts for court personnel (judges, court staff, clerks and researchers) while using an AI system that has not been officially sanctioned, for daily work.

Dos

1. Check and verify every output that is generated by the AI application for accuracy.
   Conduct a three-step check for -
   - relevance (is the output appropriate to the case?),
   - accuracy (are the facts, spellings, citations correct?) and
   - bias (has any issue, litigant or case type been unfairly framed?)
2. Avoid free trials or applications of AI. Trial-based or free continued use of AI may have limited protections for maintaining data privacy and may permit wider sharing of entered data with third parties.
3. Provide detailed, specific and narrow prompts. Vague or broad prompts reduce reliability. A legal research prompt that lacks specificity may produce incomplete or irrelevant precedents. For example, use cases like legal research assistance might not be able to provide a comprehensive list of precedents pertaining to a proposition you seek if provided with a vague prompt.
4. Before using any AI system for an actual work task, run test prompts with dummy data to understand how it behaves.
5. Where feasible, note when and for what purpose AI was used and what prompts were used (without storing sensitive data) to enable audits.
6. AI assistance must be disclosed for each use case. For example, if AI has been used for transcribing a hearing, this must be clearly disclosed before proceedings and in the final output documentation.

Don'ts

1. Do not use unsanctioned AI for judicial decision-making, analysing a pending case, tasks involving sensitive case types or litigants, or tasks that can undermine judicial independence.
2. Do not enter personal or sensitive information of persons who are parties or in any way involved in cases (e.g., full name, identifying documents such as Aadhaar/PAN card numbers, health records, bank details) to an AI system.
3. Do not provide permission for use of data for further training. Many applications allow users to opt out of model-training; ensure this setting is enabled.
4. Do not assume that the model will correctly understand the context of a case. Even if the tool approved for court use is domain specific (law), the various number of case and litigant types in a court mean that the tool cannot be fully relied on to understand the context for the task. Ensure that human judgment remains the primary filter.
5. Do not use AI for official work in your personal devices (phones, laptops, tablets).

Being wary of hallucinations

Due to the probabilistic nature of Large Language Models (LLMs), outputs can include hallucinations or fabricated responses that are inaccurate.

Hallucinations may be identified and corrected at the:

(i) prompt level, where models are instructed to provide chain-of-thought reasoning or admit to uncertainty/lack of knowledge rather than providing guesses,

(ii) model level, where a structured ontology is embedded within the model and used to guide or validate the explainability of responses,

(iii) agent level, where the LLM agent cross-verifies the response against a provided database/ through an internet search or runs a query across multiple agents to seek inconsistencies, or

(iv) through human checking.

Given widespread concerns about hallucinations, developers are increasingly implementing technical guardrails within their models and AI systems.

United Nations Development Programme

55, Lodhi Estate, New Delhi - 110003, India

+91-11-46532333 | www.undp.org/india


@UNDPInIndia


@UNDP_India


@undpinindia


@undpinindia


@UNDPIndia