JSW CENTRE FOR FUTURE OF LAW

National Law School of India University, Bengaluru

Comments And Suggestions on The Draft Regulations for Use of Artificial Intelligence in Courts, 2026

Draft dated 3rd June 2026

Submitted to:

Member Secretary, AI Committee, Supreme Court of India

office.regcc@sci.nic.in

Date: 20-06-2026

Authors: Members of JSW Centre for The Future of Law, National Law School of India University, Bengaluru.

Table of Contents

1. Executive Summary

The JSW Centre for the Future of Law ("CFL" or "the Centre") at the National Law School of India University, Bengaluru submits these comments in response to call for comments and suggestions on the 'Draft Regulations for Use of Artificial Intelligence in Courts, 2026' (hereinafter referred as the Draft Regulations) by the Supreme Court of India.

The Draft Regulations represent the most structurally ambitious judicial AI governance instrument issued by any court system in the world. Unlike the non-binding guidance documents issued by the judiciaries of the United Kingdom,[1] Singapore,[2] and Canada,[3] or the principles-based charters adopted by the European Commission for the Efficiency of Justice (CEPEJ) and the Council of Europe,[4] the draft establishes permanent institutional bodies, mandatory impact assessments, AI registers, audit cycles, and binding procurement conditions. This approach is well-suited to India's institutional context and demands of the judicial system, but has some limitations.

In general, the Centre points out two major issues. First, the draft concerns itself with at least five distinct categories of AI use: (i) private use by lawyers, (ii) submission of AI-assisted material to courts, (iii) private use by judges, (iv) institutional use in administrative processes, and (v) institutional use in adjudicative functions, all within a single regulatory instrument, without specifying which obligations apply to which category. We would encourage the disaggregation of the regulations by use-case, including the specific obligations and liabilities in each of these cases. Second, it assumes an adoption posture, expressed through the Regulation 16 presumption and Regulation 17 innovation-over-restraint principle, that is markedly more permissive than the caution exhibited by comparable judiciaries internationally and by institutional AI adopters more broadly. We would encourage more caution, with AI systems being adopted only after careful study of their possibilities, dangers and limitations through experimental pilots.

---

[1] Tribunals and Tribunals Judiciary, *Artificial Intelligence (AI) – Judicial Guidance (October 2025)*, available at https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance-october-2025/.

[2] *Guide on the Use of Generative Artificial Intelligence Tools by Court Users*, available at 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

[3] Canadian Judicial Council, *Guidelines for the Use of AI in Canadian Courts* (2024).

[4] *European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment*, European Commission for the Efficiency of Justice (CEPEJ), 2018 available at https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment. ("CEPEJ Charter")

The submission also identifies five areas of concern regarding the draft regulation and proposes specific recommendations or suggestions in response to each.

i. Several standards are technically difficult to achieve or are overbroad including the absolute fairness prohibition in Regulation 6(2) and the AI Content Verification mechanism in Regulations 3(y) and 44.
ii. The draft contains significant drafting and institutional design issues, including undefined or underdefined terms, unassigned obligations, an undefined risk architecture, a judiciary-dominated Apex Body lacking independent expertise, and a procurement regime without public tender or audit requirements.
iii. Reliance on the DPDP Act as the primary source of litigant data protection is undermined by its Section 17(1)(b) judicial exemption, and the draft does not address digital sovereignty or training-data risks.
iv. The draft creates extensive institutional safeguards but very few rights exercisable by litigants.
v. The draft has no enforcement architecture: no vendor liability, no compensation mechanism, and no retrospective review pathway.

A summary of specific proposals and suggestions is provided in each sub-section below. Key structural suggestions include the following:

1. Disaggregating the regulatory instrument by category of AI use.
2. Deleting the innovation presumption in Regulations 16 and 17 and replacing it with an evidence-based, pilot-tested adoption pathway.
3. Inserting a "Three-tier risk-classification framework" and categorising the AI tool or its use based on risk rather than task or any other parameters.
4. Amending the Apex Body's composition to include independent technologists and academicians;
5. Introducing a mandatory public tender and independent audit requirement for AI procurement.
6. Inserting self-standing litigant data rights and a broadened grievance mechanism; and creating a new enforcement and liability chapter.

2. Background and Institutional Perspective

2.1. About the Centre for the Future of Law, NLSIU

The JSW Centre for the Future of Law is an academic centre housed within NLSIU, Bengaluru, with a mandate to advance research, teaching, and public engagement on the relationship between law, technology, and institutional reform. The Centre's work focuses on questions of technology regulation, artificial intelligence governance, digital rights, platform accountability, and the legal architecture required for emerging technologies. Its engagement with the present Draft Regulations is therefore directly connected to its broader research agenda of examining how regulatory frameworks for emerging technologies can be designed in a manner that is constitutionally sound, institutionally accountable, grounded and attentive to both innovation and the rights of those the institution serves. Given that the proposed Regulations concern the institutional architecture of the judiciary, the standards governing AI assisted adjudication and court administration, and the rights of litigants in proceedings where AI tools are used, they raise issues that lie squarely within the Centre's domain of expertise.

2.2. Scope of this submission

This submission addresses the Draft Regulations for Use of Artificial Intelligence in Courts, 2026, published by the AI Committee of the Supreme Court of India on 3 June 2026 and open for public comment until 20 June 2026. Comments are organised thematically, addressing first the draft's general principles and institutional architecture before proceeding to regulation-wise analysis of specific provisions. The submission by JSW CFL is made in the centre's research and academic capacity and is intended to assist the Supreme Court Artificial Intelligence committee in evaluating the constitutional, legal, technical, and institutional implications of the Draft Regulations. The observations are offered with a view to strengthening the regulatory framework, enhancing legal certainty, and ensuring that the deployment of AI within the judicial system remains consistent with constitutional principles, procedural fairness, and technological realities.

2.3 Disclaimer

This submission has been prepared in the academic and research capacity of the JSW Centre for Future of Law, NLSIU, Bengaluru. The views, recommendations, and analysis expressed herein represent those of the authors alone and do not constitute legal advice. They do not necessarily reflect the views of the National Law School of India University or JSW Foundation.. This document is submitted in good faith and in the public interest of evidence-based, constitutionally grounded digital governance.

3. General Observations

3.1 The Draft's Structural Ambition

The draft regulations represent a structural departure from internationally comparable judicial AI governance instruments currently in place. The UK Courts and Tribunals Judiciary's AI Guidance for Judicial Office Holders,[5] for instance, first issued in December 2023 and subsequently revised in April and October 2025, is addressed to judicial office holders and their clerks, assistants, and support staff, and takes the form of non-binding guidance organised around six thematic headings. It creates no institutional bodies, registers, audit cycles, or litigant rights. Singapore's Supreme Court Registrar's Circular No. 1 of 2024[6] is structured around three operative themes: responsibility, disclosure, and citation verification. However, disclosure is reactive rather than proactive, arising only when court requests it and the Circular also supplements existing professional conduct obligations for lawyers with its own enforcement provisions. The CEPEJ Ethical Charter (December 2018)[7] articulates five principles: respect for fundamental rights, non-discrimination, quality and security, transparency, fairness and impartiality, and user control, but remains a soft-law instrument with no binding force on member states.

Unlike these instruments, the Draft regulations establishes a permanent Apex Authority, mandates Technical and Ethical Impact Assessments (TEIAs) prior to deployment, creates AI Registers and recurring audit obligations, and prescribes detailed procurement requirements, including indemnity clauses and data localisation obligations. If enacted, these would be one of the first binding, comprehensive judicial AI governance instruments in the world. This ambition appears to be justified given India's institutional context. As of 2026 more than 55 million cases[8] were pending across the Indian judicial system. Considering the scale and structural complexity of the judicial system in the country, any AI deployment that occurs across its multiple tiers is likely to be diffuse in nature and consequently difficult to monitor in a consistent and comprehensive manner. Hence, this Draft regulations tries to solve two issues: one, to responsibly use AI to reduce the work load, and two, to have a centralised binding regulatory structure that can monitor the adoption of AI across all levels of the judiciary.

---

[5] Tribunals and Tribunals Judiciary, *Artificial Intelligence (AI) – Judicial Guidance (October 2025)*, available at https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance-october-2025/.

[6] Registrar, *Guide on the Use of Generative Artificial Intelligence Tools by Court Users*, Circular No. 1 of 2024, available at https://www.judiciary.gov.sg/docs/default-source/circulars/2024/registrar's_circular_no_1_2024_supreme_court.pdf.

[7] CEPEJ Charter, *supra* note 4.

[8] *NJDG-National Judicial Data Grid*, https://njdg.ecourts.gov.in/njdg_v3//?p=home&app_token=7c737cc058bf7131b7077cfd32bda7faab164f39ee13fa8bcc053256570a8c7a (last visited June 19, 2026).

3.2 The Draft Regulations combines distinct categories of AI use

The draft regulates AI in court as a single undifferentiated subject. But in practice, AI use in courts can be classified into distinct categories with each carrying different actors, risks and appropriate regulatory tool, we have identified five such categories: (i) private use by lawyers in preparing matters; (ii) submission of AI-assisted material to a court by lawyers or litigants; (iii) private use by judges in chambers; (iv) institutional deployment of AI tools in administrative court processes; and (v) institutional deployment of AI tools in adjudicative functions.

These categories differ fundamentally in who bears responsibility, what harm looks like, and what regulatory instrument fits. Private use by lawyers raises questions of professional duty and candour to the court. These are perhaps best regulated by Professional bodies like the Bar Council of India. Submission of AI-assisted material engages evidentiary integrity and the rights of opposing parties. In practice, regulating these may require amendments to existing adjective law statutes. Private judicial use implicates judicial independence and the integrity of reasoning. It is also very difficult to regulate since a judge may use Generative AI tools frequently and not disclose such use. In fact, given the difficulty in detecting GenAI use, it is almost impossible to regulate private use.

Institutional deployment in administrative processes raises procurement, data protection, and operational accountability concerns. In particular, all procurement from private parties requires the court to follow public procurement rules. The draft moves between these categories without signalling the shift. Chapter II's general principles read as applicable to all five categories; Chapter III's permissible and prohibited use list is framed around institutional deployment; Chapter VI's procurement regime is clearly addressed only to categories (iv) and (v). Compressing all five into a single instrument produces a regime that is simultaneously over-inclusive in some directions and under-inclusive in others.

We recommend that the draft distinguish clearly between these different use-cases and specify the provisions applicable to each. This could be achieved either by organising the instrument into separate chapters governing individual judicial use, adjudicatory assistance, institutional administration, internally developed systems, and privately procured systems, or by expressly limiting particular obligations to the categories for which they are intended. At a minimum, the draft should clarify the scope of each chapter, identify the actors to whom its duties attach, and state where specific use-cases fall outside the regulatory framework.

3.3 The Adoption Posture of the Draft Regulation

The draft's general tone as expressed by Regulations 16 and 17 raises structural considerations that merit closer examination. Regulation 16 introduces a "*presumption in favour of responsible AI*

adoption" and Regulation 17 establishes an "innovation over restraint" principle, providing that "all other things being equal, an approach that prefers active and responsible adoption over restraint shall be encouraged." These provisions are something unique to the Draft Regulations for which you have no precedent in any comparable judicial AI framework, including UNESCO's, CEPEJ's or judicial AI guidance documents issued by the United Kingdom, Singapore, or Canada .

More importantly, they are structurally contradictory in a regulatory instrument. In our understanding these regulations set boundaries, they should not establish policy preferences for the activities they regulate. Policy preferences regarding AI adoption should be part of a policy or institutional strategy documents, not binding court regulations. The presumption risks creating institutional pressure to deploy AI even where need is unclear, non-AI reforms remain available, or empirical evidence of benefit is absent.

The empirical literature on AI deployment in institutional contexts provides a further basis for concern.[9] A 2026 study of 1,488 U.S. workers found that the effects of AI on workers depended on how AI was deployed.[10] When AI was used to replace routine or repetitive tasks, workers reported lower burnout. By contrast, when AI use required intensive oversight and monitoring, workers experienced higher cognitive load, mental fatigue, information overload, and reported more errors.[11] Courts adopting AI under an institutional presumption of adoption are at greater risk of hasty deployment that adds oversight burden rather than alleviating workload. Literature suggests that for tasks requiring professional accuracy which encompasses many uses contemplated by the draft the net value of AI deployment may be negligible after verification costs are accounted for.[12] This body of literature suggests that the benefits of AI adoption in institutional settings are context-dependent, and that the manner and conditions of deployment may be as consequential as the decision to adopt.

We would therefore recommend that Regulations 16 and 17 be removed and replaced with a principle of cautious, evidence-based, and proportionate adoption. AI is a rapidly developing and transformative technology. Yet, precisely for that reason its institutional deployment requires foresight rather than a regulatory presumption in its favour. Courts should adopt AI systems only after studying the relevant empirical literature, examining comparable deployments in other jurisdictions, identifying a clearly

[9] Ashley M. Hopkins et al., *Brain Fry: Navigating the Cognitive Burden of Artificial Intelligence Implementation in Healthcare*, Am. J. Med. S0002934326003839 (2026).; *In Artificial Intelligence (AI) We (Dis)Trust? Navigating Institutional Pressures for Automation and Augmentation in the Implementation of AI in Organizations*, 36 Inf. Organ. 100609 (2026).

[10] Bedard, Julie et. al., *When Using AI Leads to "Brain Fry"*, Harvard Business Review (2026).

[11] *Id*.

[12] Yuvaraj, Joshua, *The Verification-Value Paradox: A Normative Critique of Gen AI Use in Legal Practice*, Vol. 52, Monash University Law Review (forthcoming), The University of Auckland Faculty of Law Research Paper Series 2026 (2026).

defined institutional need, and assessing whether AI offers demonstrable advantages over available non-AI alternatives. New systems should ordinarily be introduced through limited and independently evaluated pilots, followed by phased deployment only where measurable benefits have been established and risks can be adequately controlled. Such deployment should also be accompanied by periodic public audits, transparent impact assessments, and clear documentation of the system's design. The governing principle should not be innovation over restraint, but responsible adoption.

4. Regulation Wise analysis

Apart from the general observations above, the Centre would like to make specific observations and recommendations on the specific regulations. We group these comments under 5 thematic clusters of concern: (i) technically unachievable or overboard standards, (ii) drafting and institutional design deficiencies, (iii) data protection and privacy gaps (iv) procedural fairness and litigant rights and the (v) cluster on enforcement liability.

4.1 Technically Unachievable or Overbroad Standards

Our first set of comments concerns the fact that many obligations laid down in the regulations are technically infeasible or institutionally difficult to achieve. These include requiring complete and non-technical explanations of AI reasoning under Regulations 3(w) and 7, ensuring that no system perpetuates or introduces bias under Regulation 6(2), guaranteeing irreversible anonymisation under Regulation 3(k), ensuring that no AI system violates any constitutional or statutory provision under Regulation 23(a), requiring complete documentation of model architecture and training data under Regulation 46(4)(g), conducting all audits entirely in-house under Regulation 38(2), verifying every AI-assisted transcript, translation or submission under Regulations 8(3), 19(1)(b)–(c), and 43–44 and continuously monitoring and immediately reporting every error or bias with potential legal consequences under Regulations 9 and 39(3). The repeated use of absolute terms such as "every," "no," "complete," "continuous," "irreversibly," "all," and "ensure" converts the regulatory objectives into difficult to meet standards. In many cases, these would prevent the use of any AI system at all. These duties should instead be framed in terms of reasonable, proportionate, risk-based and technically feasible measures. We provide two examples below:

4.1.1 AI Content Verification mechanism not achievable as drafted

Regulation 3(y) provides that no GenAI "*shall be filed or produced before any Court, without mandatory disclosure of its origin and verification through a designated centralised mechanism*," and Regulation 44 establishes an AI Content Verification Authority to operate "*verification standards, tools and protocols*" for such content.

The provision does not make clear whether "verification" refers to the detection of content as being AI-generated or authenticating content already disclosed as AI-assisted. and the feasibility depends on which is meant. This critical distinction determines whether the Authority would operate a provenance-authentication framework or an AI-content detection service, and whether its function would be technically achievable.

If it means detection, it can not be achievable as no existing AI-content detector can reliably distinguish human-generated from AI-generated text across all contexts, writing styles, and models. Studies reveal that AI detector accuracy, while moderate to high in aggregate (AUC 0.75–1.00), remains far from perfect: false-positive rates, where genuinely human-authored text is misclassified as AI-generated reached up to 30% in some detectors, and false-negative rates varied considerably depending on the ChatGPT version tested, with older models like GPT-3.5 evading detection at rates approaching 24% on certain tools.[13] Further, Studies have demonstrated that AI-content detectors disproportionately misclassify writing produced by non-native English speakers as AI-generated with an average false-positive rate of 61.22%.[14] This raises equality of access concerns with direct constitutional resonance in a linguistically diverse jurisdiction.

Recommendation:

- Reframe the mechanism around disclosure. Rely on the Regulation 43 disclosure obligations as the primary safeguard.
- Insert a new operative provision (Chapter V): "(1) No GenAI-generated content shall be filed or produced before any Court without disclosure of its AI-generated origin in the manner prescribed under Regulation 43.

  (2) A recognised provenance credential or watermark, such as a hallucinated citations, may be relied upon as corroboration of AI origin;

  (3) The apex body shall set, maintain and update standards for recognised provenance credentials."

[13] Weber-Wulff et al., *Testing of detection tools for AI-generated text*, Vol. 19(26) Int J Educ Integr (2023); *See also* Erol et al., *Can we trust academic AI detective? Accuracy and limitations of AI-output detectors*, Vol. 167(214) Acta Neurochir (2025).

[14] Liang et. al., *GPT detectors are biased against non-native English writers*, Stanford University (2023), available at https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers.

4.1.2 Absolute Fairness and Bias Prohibition

Regulation 6(2) provides that "*no AI System shall be deployed that perpetuates, amplifies, or introduces bias*" on the grounds of race, religion, caste, sex, gender, disability, language, economic status, or any other prohibited ground.

Regulation 6(2)'s fairness mandate is constitutionally essential but, as it stands currently drafted, is difficult to satisfy in practice. Every model trained on real-world data shows some statistical disparity across groups.[15] The major fairness metrics like equalised error rates, calibration, demographic parity cannot all be satisfied simultaneously when base rates differ across groups and this is particularly relevant in Indian judicial context where the training data may reflect historical discriminations that have been perpetuated over time. A standard that is required to only "promote fairness" without specifying which metric governs does not resolve this tension.

The draft itself concedes that bias cannot be eliminated: Regulation 14(1) requires training data to be "*to the extent feasible, free from discriminatory bias.*" Regulation 6(2) then demands its absolute absence. The instrument contradicts itself; and while the draft regulations address algorithmic bias and prohibits the use of AI for certain functions like risk scoring in lieu of this,[16] the understanding was not carried into the general fairness standard. Further, keeping in mind the *Lt. Col. Nitisha* v. *UoI* judgment, wherein the Supreme Court recognised indirect discrimination,[17] the burden of ensuring that the AI system deployed does not perpetuate, amplify or introduce bias is extremely high. In this case, we saw Justice D.Y Chandrachud recognise indirect discrimination as when an action has a disproportionate impact on a certain community, even if it is seemingly 'neutral' and treats all alike.[18] Thus, while theoretical solutions exist to reduce algorithmic bias,[19] such solutions do not account for indirect discrimination. The UNESCO Guidelines' Principle 1.1(a) adopts the more workable formulation of preventing AI outcomes that "aggravate, reproduce, reinforce, or perpetuate

[15] Maarten Buyl &Tijl De Bie, *Inherent Limitations of AI Fairness*, COMMUNICATIONS OF THE ACM (2024).
[16] Regulation 20(1)(d) bans risk scoring due to the possibility of algorithmic bias.
[17] Lt. Col. Nitisha v. Union of India, 2021 SCC Online SC 261.
[18] *Id*., ¶¶49-51; Mihir R & Krutika R, *Court Recognises Indirect Discrimination & Strikes Down Army's Gender Discriminatory Promotion Practices*, SUPREME COURT OBSERVER, April 12, 2021, available at https://www.scobserver.in/journal/court-recognises-indirect-discrimination-strikes-down-armys-gender-discriminatory-promotion-practices/.
[19] Zehlike et. al., *Beyond Incompatibility. Trade-Offs between Mutually Exclusive Algorithmic Fairness Criteria in Machine Learning and Law* (2022).

discrimination", a process-and-impact standard.[20] The CEPEJ Charter similarly focuses on preventing the "development or intensification" of discrimination.[21]

Recommendation

Amend Regulation 6 to replace the absolute prohibition on bias with a risk-based obligation. The Regulation should require mandatory pre-deployment bias testing and periodic monitoring thereafter, with assessments addressing potential sources of bias arising at different stages of the AI lifecycle, including data collection, model development, deployment, and use.[22] Periodically thereafter, including for intersectional bias across combinations of protected characteristics, disclosure and documentation of measured disparities;[23] involve data augmentation to increase representativeness and reduce bias, selection and reasoned justification of a fairness criterion appropriate to the use;[24] and a prohibition on disparate adverse impact that is unjustified or exceeds a stated threshold. The burden of demonstrating compliance should rest on the deploying authority or vendor through the assessment under Regulation 35, with periodical scrutiny conducted as per the process.

4.2 Issues with Drafting and Institutional Design

4.2.1 Definition Inconsistencies and Missing Definitions

The draft carries two near synonymous terms. 'Artificial Intelligence' is the technology definition, while 'AI system' or 'AI Tool' is defined as any software, platform, application, device or process that employs AI in connection with a Court process. The operative obligations attach majorly to the 'AI system', and thus the reach of those obligations is understood by what counts as Artificial Intelligence under Regulation 3(m) through the 'employs Artificial Intelligence' wording in Regulation 3(i).

Under Regulation 3(m) the definition requires a system that 'infers, learns, and generates.' Read conjunctively, it requires the system to learn. A trained model at inference does not learn, its parameters are fixed. Based on a literal reading, such a model can not be classified as 'Artificial Intelligence' and the software can not be deemed an 'AI system.' The definition also folds 'deployed for court processes' into what should be a technology definition. Additionally, it excludes the general

[20] Gutiérrez, Juan David, *Guidelines for the use of AI systems in courts and tribunals*, United Nations Educational, Scientific and Cultural Organization (UNESCO), 2025, available at https://www.unesco.org/en/articles/guidelines-use-ai-systems-courts-and-tribunals. ("UNESCO Guidelines")

[21] CEPEJ Charter, *supra* note 4, Principle 2.

[22] Emilio Ferrara, *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, Vol. 6(1) Sci. 3 (2024).

[23] Davies et al., *Algorithmic decision making and the cost of fairness,* Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada 797-806 (2017).

[24] *Id.*

purpose software 'unless... functionally dependent upon, artificial intelligence,' making the definition circular by defining the term by a reference to itself.

The inconsistency can also be seen in the AI term definitions. 'Machine Learning' is defined as a subset of 'Artificial Intelligence' while 'Generative AI' and 'Large Language Model' are defined as a class or type of 'AI system.' With both 'Artificial Intelligence' and 'AI system' as parent terms in circulation, the family does not nest coherently.

There are similar problems with other definitions. For example, in the definition of 'AI Incident' (Regulation 3(e)) 'substantial risk of harm' can be understood from the definition of 'harm' under 3(za), and inherits the narrowness that definition carries. The qualifier 'substantial' may also exclude minor procedural or linguistic errors, such as changing the date in a translation, that can prejudice a litigant without crossing the substantial risk threshold. The definition can be addressed by broadening the definition of harm under 3(za) and clarifying the 'substantial risk' threshold.

The definition of 'harm' (Regulation 3(za)) is inclusive, and the reference to rights can be read to extend to procedural and due process harms. However, the definition is confined to 'in relation to AI Incidents' even though 'harm' is also used in the grievance and liability provisions. This leaves 'harm' undefined for where a litigant seeks a remedy. Moreover, it does not explicitly mention the harms likely to arise from AI in a judicial context, such as harm through delays in proceedings, procedural exclusion, barriers to participation, or the exposure of sensitive personal information. These forms of harm can affect due process rights and access to justice even in the absence of tangible economic or physical injury. Further, scholarship on algorithmic governance has demonstrated that automated systems may reinforce existing social inequalities and disproportionately disadvantage vulnerable groups. The definition should therefore expressly recognise procedural, linguistic, privacy related, and access to justice harms arising from the deployment of AI systems in courts.[25] The definition for 'harm' should, therefore, remove the limitation of 'in relation to AI Incidents' and be expanded to include procedural, linguistic, privacy related and access to justice harms.

The definition of 'sensitive judicial data' (Regulation 3(zh) is stated at a level of generality that cannot provide calibrated protection across the range of sensitive information judicial records may contain. Another term 'sensitive personal data' may be introduced, with a standard of protection tied to the proportionality principle in Regulation 12, so that the most sensitive data gets the highest protection.

[25] Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor (New York: St. Martin's Press, 2018); Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code (Cambridge: Polity, 2019); Solon Barocas and Andrew D. Selbst, *Big Data's Disparate Impact*, Vol. 104, California Law Review, 671-732 (2016).

Recommendations

- The two parent terms 'Artificial Intelligence' and 'AI System' can be consolidated into a single defined term, 'AI System,' aligned with the widely adopted intergovernmental standard, and reduce 'Artificial Intelligence' to a synonym. This also lets the family nest under one parent definition.

(i) AI System: "AI System" or "AI Tool" is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.[26]

(m) Artificial Intelligence: 'Artificial Intelligence (AI)' shall have the same meaning as 'AI System.'

- The definition for 'harm' should remove the limitation of 'in relation to AI Incidents' and be expanded to include procedural, linguistic, privacy related and access to justice harms.

(za) "harm" includes any kind of physical or financial damage, or damage to the reputation or rights of any individual, institution, or infrastructure; and any procedural, linguistic, privacy related or access to justice injury, including any delay in proceedings, procedural exclusion, barriers to participation and exposure of sensitive personal information.

- Introduction of another term 'sensitive personal data' with a standard of protection tied to the proportionality principle in Regulation 12, so that the most sensitive data gets the highest protection. The recommended definition is based on the definition of 'sensitive personal data' in 3(36) of the Personal Data Protection Bill, 2019.[27]

(zha) "sensitive personal data" means such personal data, which may, reveal, be related to, or constitute–

i. financial data;
ii. health data;
iii. official identifier;
iv. sexual orientation;
v. biometric data;
vi. genetic data;
vii. gender identity;
viii. caste or tribe;

[26] OECD, *Recommendation of the Council on Responsible Innovation in Neurotechnology*, OECD/LEGAL/0449 (2019).
[27] *The Personal Data Protection Bill, 2019,* Bill No. 373 of 2019 (India).

ix. religious or political belief or affiliation; or
x. comparably intimate personal data

of parties, witnesses, victims or legal representatives processed in connection with a Court process; and forms a subset of sensitive judicial data.

Amendment of Regulation 10(1) to add below:

10. Data protection and privacy.–– (1) The processing of personal data ... sensitive judicial data shall be accorded the highest standard of protection. *Sensitive personal data, being its most sensitive subset, shall in addition be:*
(a) anonymised or pseudonymised to the extent feasible before any processing through an AI System;
(b) not used to train, test or refine any AI System without the specific approval of the Appropriate Authority; and
(c) subject to access on a strict need-to-know basis in accordance with the proportionality principle in Regulation 12.

- Add an explanation to Regulation 43(1) defining 'material assistance'

Explanation: "material assistance": where it could reasonably affect the substance, outcome or procedural fairness of a matter, and not where it is limited to formatting, clerical or administrative tasks that do not bear the content of any submission, decision or step in the proceeding.

4.2.2 Explainability Requirements: Clarifying the standard of Explainability (Regulations 3(w) and 7)

The draft regulation rightly recognises explainability as a foundational principle for the use of Artificial Intelligence in courts. Regulation 3(w) defines explainability as the capacity of an AI system to generate a comprehensible account of the reasoning, factors, weightings, and process by which it arrived at a particular output, recommendation, or decision, while Regulation 7 requires all AI systems used in court processes to meet "high standards of transparency and explainability." Further, Regulation 7(3) subjects opaque or incapable-of-explanation systems to heightened scrutiny and restricts their use in high-risk applications affecting personal liberty or lawful rights. Along with the requirements for TEIA, audit mechanisms, and vendor documentation in the draft, these provisions demonstrate a commendable commitment to transparency and accountability.

The concern, however, is that the explainability safeguards lack clarity regarding the level and form of explanation required. Regulation 3(w)'s demand for *"a comprehensible account of reasoning, factors, weightings, and process"* cannot be completely satisfied by large language models, which generate

outputs through statistical relationships among millions of parameters rather than identifiable reasoning chains. Post-hoc explainability techniques such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) provide approximations of feature importance, not genuine accounts of how internal weights produced a given output.[28] This matters practically because the tools most likely to be deployed under the draft will be for translation, transcription, document summarisation, legal research assistance, and conversational chatbots are precisely those powered by LLMs.

The EU AI Act adopts a more clearer approach to explainability. It requires that high-risk AI systems used in the administration of justice be sufficiently transparent to enable users to interpret and appropriately use their outputs, while ensuring traceability through logging, technical documentation, and human oversight mechanisms.[29] This emphasis on intelligibility, auditability, and accountability provides a more technically feasible standard than demands for complete algorithmic explainability.

We have identified two issues with the Draft regulations in relation to this, First, "high standards of explainability" in Regulation 7 is not a legal standard; it cannot ground enforcement, liability, or judicial review without further specification of what it requires and to whom. Second, "opaque or incapable of explanation" in Regulation 7(3) is undefined and therefore unenforceable, since a vendor can always assert that their system provides some form of explanation without any standard specifying what that must contain.

A further design gap is the absence of any requirement that each approved AI system be accompanied by a mandatory Explanation Memorandum. The draft requires impact assessments before deployment but not a standing document, accessible to judicial officers and parties alike, that explains in non-technical language how the system reaches its outputs, what inputs are used, what limitations apply, and how a user would recognise an erroneous output. Explainability to lay users and explainability to technical reviewers are distinct obligations that must be discharged separately.

Recommendations

- Reword Regulation 3(w): Replace the current definition with a two-tier formulation distinguishing (a) inherently interpretable systems, which should provide full accounts of their reasoning, factors, and weightings; and (b) opaque machine-learning systems, including LLMs,

[28] Hooshyar, Danial & Yang, Yeongwook, *Problems With SHAP and LIME in Interpretable AI for Education: A Comparative Study of Post-Hoc Explanations and Neural-Symbolic Rule Extraction*, IEEE Access (2024).
[29] EU Artificial Intelligence Act, Article 13, available at https://artificialintelligenceact.eu/article/13'.

which should be required to provide intelligible explanations of inputs used, outputs generated, confidence levels, known limitations, and identified risks, without being required to disclose internal parameters or weights.

- Reword Regulation 7: Replace “high standards of transparency and explainability” with a defined two-track standard: (a) a lay explanation of outputs, inputs, confidence levels, and limitations described in plain language to the judicial officer, court staff, or litigant to whom the output is directed, without requiring specialist technical knowledge; and (b) a technical explanation of architecture, training data characteristics, decision logic, and audit trail described in sufficient detail to enable independent review by the AI Secretariat or a nominated technical expert.
- Define "opaque" in Regulation 7(3): An AI system should be considered opaque where it is incapable of providing either the lay or the technical explanation described above. Opaque systems should be expressly prohibited in adjudicatory contexts and in any application affecting personal liberty or the lawful rights of a party, rather than merely subjected to undefined "heightened scrutiny."
- New provision required: Every AI system approved for use in court processes should be accompanied by a standing Explanation Memorandum, lodged with the AI Register under Regulation 37 and made publicly accessible. The Memorandum should specify, in non-technical language: the function the system performs, the inputs it uses, the outputs it generates, its known limitations and error rates and how a judicial officer, court staff member, or litigant would recognise an erroneous output.

4.2.3 Unassigned Obligations

Several provisions mandate an act in the passive voice without assigning responsibility to a specific actor or specifying the criteria.

For instance, in Regulation 36(1) for Controlled Environment Testing, qualification of “in suitable cases” is given without defining the criteria for suitability. Regulation 46(2) provides that proposals for private engagement would be “subject to a comprehensive evaluation” without naming the evaluator. Regulation 48(4) provides that anonymisation would be applied to personal data “to the extent technically feasible” without identifying who determines feasibility. Regulation 48(3) provides that data minimisation would be “applied in the selection and deployment of AI Systems” but provides for no specified actor.

Regulation 11 on purpose limitation does not adequately guard against "function creep" the gradual expansion of the use of a technology beyond its originally approved purpose.[30] This is an observed phenomenon across the use of technologies and is very much applicable to the use of AI tools.[31] On the face of it, Regulation 11 addresses this, but function creep is not a discrete, identifiable event. Instead, it is a gradual process. A tool introduced for case summarisation may be applied to case prioritisation, performance analytics, or legal research without triggering a fresh approval requirement, as they might not be flagged as functions that require fresh approval. Technologies carry institutional assumptions and can reorganise authority structures, even when described as routine tools.[32] The UNESCO Guidelines' Principle 1.12 on accountability and contestability establishes that where a tool's purpose expands without fresh approval, affected parties' ability to contest its use becomes structurally impossible as they will be unable to recognise the expansion of the use of the tool.[33] This mandates a continued monitoring mechanism that can be managed by the AI registry under Regulation 37 to check whether actual use diverges from registered purpose.

The proportionality provision in Regulation 12 does not require consideration of non-AI alternatives. The UNESCO Guidelines' Principle 1.3 on feasibility of benefits requires assessing the specific potential benefits of an AI system and the institutional capacity to realise them, which logically includes assessing whether non-AI alternatives exist and have been considered.[34] AI may not always be the most suitable response: in some settings, judicial appointments, procedural simplification, or legal aid investment address underlying bottlenecks more directly and with fewer risks.

The consultation mechanism under Regulation 33(6) allows AI Committees to invite suggestions from universities, researchers, and the private sector, but makes no mention of litigants, legal aid bodies, disability rights groups, or civil society organisations. Affected communities understand the practical risks and institutional frictions that expert-led systems overlook.[35] The UNESCO Guidelines' Principle 1.14 on human-centric and participatory design requires that governance bodies hear from those whose access to justice is most directly affected by AI deployment.[36]

---

[30] Koops, Bert-Jaap, *The concept of function creep*, Vol. 13, Law, Innovation and Technology, 1-28 (2021).

[31] Fantin, Stefano & Vogiatzoglou, Plixavra, *Purpose Limitation By Design As A Counter To Function Creep And System Insecurity In Police Artificial Intelligence*, in United Nations Interregional Crime and Justice Research Institute (UNICRI), UNICRI Special Collection on AI (2020), available at http://www.unicri.it/sites/default/files/2020-08/Artificial%20Intelligence%20Collection.pdf.

[32] Sheila Jasanoff, *Technologies of Humility: Citizen Participation in Governing Science*, 41 Minerva 223 (2003); Winner, Langdon, *Do Artifacts Have Politics?*, Vol. 109(1), Daedalus,121-36, (1980).

[33] UNESCO Guidelines, *supra* note 19, Principle 1.12.

[34] *Id*., Principle 1.3.

[35] Wynne, Brian, *Misunderstood Misunderstanding: Social Identities and Public Uptake of Science*, Vol. 1(3), Public Understanding of Science, 281–304 (1992).

[36] UNESCO Guidelines, *supra* note 19, Principle 1.14.

The AI Incident Database under Regulation 39 should be designed as a learning system rather than a catalogue of adverse events. Failures in tightly coupled socio-technical systems arise through chains of interaction.[37] These may not be visible if institutions record only major breakdowns. Thus, the database should capture near misses, user complaints, false positives in document verification, and data breaches, with a periodically published public bulletin.

Recommendation:

- Delete Regulations 16 and 17 in their current form. Replace with a provision requiring the Appropriate Authority, before approving any AI system, to document: empirical evidence of improvement in access to justice, reduction in delays, or enhancement of administrative efficiency in a comparable judicial context or in a controlled pilot within the Indian judicial system; consideration of non-AI alternatives and reasons for preferring AI deployment; and a pilot-test-deploy pathway with mandatory public reporting at each stage.
- Amend Regulation 11: add a sub-regulation providing that any material change in the function, data inputs, output use, or institutional setting of an AI system constitutes a new deployment requiring fresh approval and a new TEIA, and there should be continued monitoring to make sure that the tool is only deployed for the registered purpose.
- Amend Regulation 12: require every TEIA to include a non-AI alternatives assessment documenting alternatives considered and the reasons AI deployment is preferred.
- Amend Regulation 33(6): name litigants' representatives, legal aid bodies, disability rights organisations, and civil society groups as required consultees, not merely optional ones.
- Amend Regulation 36 (1) : define 'suitable cases' as all deployments of high-risk AI systems under the proposed risk-classification framework, making Controlled Environment Testing mandatory for such systems.
- Amend Regulation 46(2): assign the comprehensive evaluation function to the AI Secretariat, with mandatory referral to the Technical Committee for systems involving personal data, adjudicatory functions, or high-risk applications.
- Amend Regulations 48(3) and (4): explicitly assign data minimisation and anonymisation feasibility determinations to the Appropriate Authority or AI Secretariat.
- Amend Regulation 39: expand the AI Incident Database to include near misses, user complaints, false positives in document verification, and data breaches. Require a periodically published public incident bulletin.

[37] Charles Perrow, Normal Accidents – Living with High Risk Technologies (1984).

4.2.4 Risk-Based Architecture

The draft repeatedly invokes concepts of risk without ever defining them. Regulation 7(3) subjects' opaque systems to heightened scrutiny in "high-risk applications." Regulation 12 requires safeguards proportionate to the "risk profile" of a task. Regulation 46(4)(h) requires explainability documentation for "high-risk AI tools." The draft uses risk language sometimes in terms of usage, sometimes in terms of tasks, and sometimes in terms of outputs, without establishing a consistent methodology. Until risk levels are defined, the instrument's proportionality architecture would be inoperative.

The EU AI Act resolves this through explicit classification: it designates as high-risk AI systems "intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts."[38] By this standard, most tools within the scope of the draft would be classified as high-risk. The CEPEJ's 2023 operationalisation of its Ethical Charter identifies specific risk typologies in judicial AI deployment including unclear grounds for judgements, discrimination and amplified bias, and forced use of AI where opt-out systems are absent.[39] The UNESCO Guidelines' Principle 1.1 requires effective safeguards for AI systems "that might represent a high risk to human rights," presupposing that the category is defined.[40]

The Regulations would function better if they are categorised as principled risk-based architecture throughout, and not be usage- or task-based in any instance. Risk should be determined by the potential impact of the system on fundamental rights, personal liberty, procedural fairness, and the finality of the proceeding. We propose the following three-tier framework:

- Tier 0 - Absolute prohibition: uses already enumerated in Regulation 20, including risk scoring, algorithmic adjudication, undisclosed AI in liberty-affecting proceedings, and surveillance without legal authority. These remain non-derogable. Proceedings involving personal liberty, capital punishment, final constitutional appeals, and non-reviewable determinations. AI may not be used to assist in research, drafting, analysis, or case management in any matter at this tier, including for administrative functions in that matter.
- Tier 2 - High risk, mandatory human in-the-loop with documented verification: matters affecting legal rights but not personal liberty like civil final appeals, property, contract, family law. AI may assist in administrative and research functions only. All outputs must be

[38] EU Artificial Intelligence Act, Annexure III, available at https://artificialintelligenceact.eu/annex/3/.
[39] CEPEJ Charter, *supra* note 4.
[40] UNESCO Guidelines, *supra* note 19, Principle '.1.

independently verified using the substantive protocol defined under Regulation 8 by a named Designated Officer. Verification cannot be waived.

- Tier 3 - Standard risk, AI permitted with Technical and Ethical Impact Assessment and audit: administrative and registry functions, cause list management, scheduling, translation of non-adjudicatory documents. Standard TEIA and annual audit cycle applies. Expedited approval pathway under Regulation 34(4) available for functionally similar tools.

Every reference to 'high-risk' in the instrument should be replaced with a cross-reference to this framework, and the Appropriate Authority's case-by-case approval powers under Regulations 18 and 35 should operate within, not in lieu of, the tier classification.

Recommendations

- New provision required (new Regulation 18A or Schedule to Chapter III): insert a three-tier risk-classification framework as described above, drawing on the CEPEJ risk typology and EU AI Act judicial high-risk classification as comparative benchmarks, adapted to Indian constitutional requirements including Articles 14, 21, and 39A.
- Amend Regulations 7(3), 12(2), and 46(4)(h): replace unglossed references to 'high-risk' with cross-references to the classification framework, ensuring that trigger conditions for heightened safeguards are defined and consistent.

4.2.5 Reliance on Human-in-the-loop (HITL) should be reduced

The draft regulations rely heavily on HITL as an AI fail-safe. Regulation 3(zb) defines HITL as "*a governance process in which the outputs, recommendations, or results generated by an AI System or AI Tool are subject to mandatory human review, supervision and verification, with final decision-making authority, responsibility, and accountability resting at all times with a human.*" It employs HITL as a tool to ensure that all AI processes employed in courts are subservient to human judgement (Regulation 4), to ensure proportionality in the use of AI (Regulation 12(2)), as restricting clause on the presumption in favour of responsible AI adoption (Regulation 16 *proviso*), for translation of judgments (Regulation 19(c)), to ensure functioning of AI chatbots (Regulation 19(f)), to undertake document verification in administrative processes (Regulation 19(h)). As a safeguard, the draft further provides that no adjudication or judicial outcome may be undertaken solely on the basis of the use of an AI tool, and requires mandatory HITL (Regulation 20(b), (c)). Additionally, the nominated officer supervising any AI system has full discretion to "*accept, modify, or reject any AI-generated recommendation or output within the matters falling under his charge*" (Regulation 40), and even to dispense with the requirement of verification (Regulation 8), which further demonstrates reliance on HITL.

This extensive reliance on HITL is a cause for concern for multiple reasons. *First*, there is a risk of multiple cognitive biases that arise when humans work with AI tools. Two of these biases are the automation bias, which occurs when humans become uncritical of machine output after successful use,[41] and selective adherence, which occurs when humans accept algorithmic advice only when it confirms pre-existing inclinations.[42] Studies have consistently found that automation bias cannot be eliminated, and even experts, who are resistant to it to some extent, remain susceptible.[43] Although selective adherence persists for human and algorithmic advice,[44] it remains a problem due to anchoring bias and verification drift. *Second*, the use of AI tools by humans inevitably leads to confidence in the algorithmic output, referred to as "verification drift."[45] This confidence in algorithmic output arises despite users' awareness of generative AI's limitations.[46] Additionally, anchoring bias is generated, in which users may be inclined to blindly adopt system recommendations or be disproportionately influenced by AI predictions, even when those predictions are inaccurate.[47] These concerns are adequately represented through the hundreds of court filings containing fabricated, AI-generated citations, that have crept up around the world.[48] The much-cited *Mata v. Avianca* sanctions in New York being a famous example.[49] *Third*, the aforementioned biases and verification drift are exacerbated by increased errors from AI use. Recent studies have identified a phenomenon termed "AI brain fry," referring to mental fatigue caused by excessive use of or intensive oversight of AI systems beyond an individual's cognitive capacity. Such cognitive strain has been associated with decision fatigue and a higher likelihood of workplace errors, with affected workers reporting 11% higher minor error rates and 39% higher major error rates compared to those who did not experience AI brain fry.[50]

---

[41] Romeo, G., Conti, D, *Exploring automation bias in human–AI collaboration: a review and implications for explainable AI*, Vol. 41, AI & Soc (2026); Parasuraman & Manzey, *Complacency and Bias in Human Use of Automation: An Attentional Integration*, Vol. 52(3), Human Factors and Ergonomics Society (2010).

[42] Saar Alon-Barkat, Madalina Busuioc, *Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice*, Vol. 33(1), Journal of Public Administration Research and Theory, 153-169 (2023).

[43] Romeo, G., Conti, D, *Exploring automation bias in human–AI collaboration: a review and implications for explainable AI*, Vol. 41, AI & Soc (2026).

[44] Saar Alon-Barkat, Madalina Busuioc, *Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice*, Vol. 33(1), Journal of Public Administration Research and Theory, 153-169 (2023).

[45] A. Alimardani, *Borderline Disaster: An Empirical Study on Student Usage of GenAI in a Law Assignment*, Vol 6(2), IEEE Transactions on Technology and Society, 210-219 (2025); Legg et. al., *The Promise and the Peril of the Use of Generative Artificial Intelligence in Litigation*, Vol. 48(4), UNSW Law Journal 1196 (2025).

[46] *Id*.

[47] Rosbach et. al., *Stuck on Suggestions: Automation Bias, the Anchoring Effect, and the Factors That Shape Them in Computational Pathology*, Machine Learning for Biomedical Imaging (2025).

[48] Siva, Pramod Kumar, *Citing the Unseen: AI hallucinations in tax and legal practice: A comparative analysis of professional responsibility, procedural legitimacy, and sanctions*, Vol. 53(1), The International Tax Journal, 390-401 (2026).

[49] *Id*., Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

[50] Bedard et. al., *When Using AI Leads to "Brain Fry"*, Harvard Business Review (2026).

Therefore, although the draft attempts to restrict reliance on AI tools in judicial outcomes, adjudication, and sentencing (Regulation 20(b) and (c)), they effectively permit their use as long as there is HITL. The current inadequacies in human judgement and AI detection[51] prove that HITL as a protective mechanism is inefficient and insufficient in the use of AI tools to the extent that Regulation 20 provides, as any human bias may lead to irreparable damage to parties.

Recommendation:

- Implement a substantive verification protocol for permitted uses of AI in administrative and non-adjudicative court processes. The protocol should not rely on "reasonable care" as a standard. A substantive verification protocol would involve independently checking every authority and factual proposition against primary sources to ensure no lapses that might otherwise occur due to the various types of bias.
- Amend Regulation 20(b) and (c): delete the qualifications "alone" and "solely" in Regulation 20(b) and the exception permitting AI-assisted adjudication and sentencing subject to mandatory Human-in-the-Loop in Regulation 20(c), such that no AI tool may be used in any proceeding bearing on judicial outcomes, adjudication, or sentencing, and the prohibition admits of no exception on the basis of human oversight or the advisory character of the AI output.

4.2.6 Institutional Composition and Expert Independence

Regulation 22 constitutes the Apex Body with two Supreme Court judges, two Chief Justices of High Courts, two High Court judges, one member from an institution of national importance, one MeitY Joint Secretary, one finance expert, one cybersecurity expert, one or more advocates in technology law, and "the Professor heading the field of Artificial Intelligence at the National Judicial Academy, Bhopal."

The more fundamental concern is structural: every decision-making function of the Apex Body which includes tool approval, audit oversight, grievance recommendations, standard-setting rests with the judiciary itself or with persons nominated by the Chief Justice of India. There is no independent technical verification body, no mandatory academic AI-and-law expertise, and no civil society representation.

[51] Weber-Wulff et al., *Testing of detection tools for AI-generated text*, Vol. 19(26) Int J Educ Integr (2023); *See also* Erol et al., *Can we trust academic AI detective? Accuracy and limitations of AI-output detectors*, Vol. 167(214) Acta Neurochir (2025).

The UNESCO Guidelines' Principle 1.14 on participatory design[52] and CEPEJ's emphasis on multi-stakeholder governance[53] require that bodies governing judicial AI include people who represent affected interests, not only institutional insiders. Governance bodies designed by expert insiders without structural participation from affected communities systematically overlook the practical frictions and distributional harms that those communities experience.[54]

4.2.7 Procurement Transparency and Public Audit

Regulation 46 requires prior written approval of the Appropriate Authority and "comprehensive evaluation" for private AI system engagement, but does not specify a public procurement process, competitive tender requirement, or post-deployment public audit. Regulation 46(7) provides an expedited thirty-day approval pathway for functionally similar administrative tools. The combination of an unspecified evaluation process, no tender requirement, and an expedited fast-track creates conditions for single-vendor entrenchment and opaque procurement.

More significantly, Regulation 46 addresses private entities but says nothing about publicly developed AI systems. The National Informatics Centre (NIC) has developed and deployed in Indian courts the SUVAS translation tool,[55] the SUPACE legal research assistant,[56] the LegRAA tool,[57] and a RAG-based judicial search assistant built on NVIDIA infrastructure.[58] None of these tools is subject to a mandatory independent technical audit requirement under the draft regulations, which are silent on AI systems not procured from private vendors. NIC-developed systems serve courts at scale, and their performance has not been independently evaluated against the TEIA requirements that the draft would impose on private systems.

---

[52] UNESCO Guidelines, *supra* note 19, Principle 1.14.
[53] CEPEJ Charter, *supra* note 4.
[54] Wynne, Brian, *Misunderstood Misunderstanding: Social Identities and Public Uptake of Science*, Vol. 1(3), Public Understanding of Science, 281–304. (1992).
[55] PIB, *From Digitisation to Intelligence: How AI is Enhancing Access to Justice in India, Technology as an Enabler of Justice, not a Substitute for Judgement*, PIB, February 11, 2026, available at https://www.pib.gov.in/PressReleasePage.aspx?PRID=2226283&reg=48&lang=2.
[56] *Id*.
[57] *Id*.
[58] NVIDIA Success Stories, *How India's National Informatics Centre Is Modernizing the Indian Judiciary With AI*(no date), NVIDIA, available at https://www.nvidia.com/en-us/case-studies/role-of-nic-modernizing-indian-judiciary-using-ai/.

4.3 Data Protection, Privacy and Digital Sovereignty

4.3.1 Reliance on the DPDP Act Despite Judicial Exemptions

Regulation 10(1) provides that the processing of personal data through AI Systems "*shall be governed by the principles of purpose limitation, data minimisation and data privacy by design in accordance with the provisions of the Digital Personal Data Protection Act, 2023*[59]," and Regulation 47 requires compliance with that Act. The draft thus locates the data protection guarantee for litigants primarily in the DPDP Act.

That reliance may not deliver what it appears to. Section 17(1)(b) of the DPDP Act[60] provides that Chapter II (except sub-sections (1) and (5) of section 8), Chapter III, and section 16 do not apply where personal data is processed by a court or tribunal, or a body entrusted by law with a judicial function, "where such processing is necessary for the performance of such function." On a broad reading, a court processing personal data in adjudication is relieved of the data fiduciary obligations of notice, purpose limitation and data minimisation, and the data principal loses every Chapter III right – access (section 11), correction and erasure (section 12), and grievance redressal (section 13)[61] – leaving only the general accountability and security safeguard duties under section 8(1) and 8(5), which confer no enforceable entitlement on the litigant. On that reading, the protection the regulations may not apply to situations where AI might most affect a litigant.

The scope of that exemption is itself unsettled. The words "*necessary for the performance of such function*" may be read narrowly, so that personal data processed through a procured AI tool, not strictly required to decide the case, falls outside the exemption and the DPDP Act continues to govern. Whichever reading prevails, a regulation should not rest litigant protection on a statutory exemption whose reach is uncertain. This is especially important since the incorporation of AI into judicial decision-making involves the collection, storage and analysis of sensitive personal data for responses and improved efficiency and accuracy.[62] The commencement position is a further difficulty: the DPDP Act and its 2025 Rules were notified on 14 November 2025 with staggered effect, and the data fiduciary obligations and data principal rights take effect only on 13 May 2027,[63] the Information Technology Act, 2000 and the 2011 Rules governing until then.

---

[59] Digital Data Protection Act, 2023.
[60] Digital Data Protection Act, 2023 § 17(1)(b).
[61] Digital Data Protection Act, 2023 §§§ 11, 12, 13.
[62] Socol de la Osa DU, Remolina N., *Artificial intelligence at the bench: Legal and ethical challenges of informing—or misinforming—judicial decision-making through generative AI*, Vol. 6, Data & Policy (2024).
[63] Ministry of Electronics and Information, Notification, G.S.R. 846(E), available at https://www.dpdpa.in/dpdpa_rules_2025/.

Although Regulations 10 and 48 set out the Court's own data protection principles, these bind the Court institutionally rather than conferring rights a litigant may invoke, and Regulation 10 draws its content from the very DPDP provisions the exemption may remove.

4.3.2 Digital Sovereignty, Data Residency, and Training Data Prohibition

Three key questions relating to digital sovereignty are inadequately addressed by the draft: where court data may be hosted, who may access it under foreign law, and whether court data may be used to train AI models.

Regulation 46(4)(j) requires "*on-premises or sovereign cloud deployment*" for AI systems processing sensitive judicial data but does not define sovereign cloud or address the implications of Indian judicial data being processed on infrastructure operated by entities subject to foreign legal compulsion. The US CLOUD Act empowers US law enforcement to compel any US-headquartered company to produce data stored anywhere in the world, regardless of physical server location.[64] A requirement that data be stored "in India" does not guarantee Indian data sovereignty if the infrastructure is operated by a US-headquartered company. Several AI systems already in use in Indian courts, including NIC's NVIDIA-based infrastructure for transcription and legal research, involve supply chain dependencies on entities subject to such compulsion.[65] The regulations should define sovereign cloud by reference to Indian legal jurisdiction and prohibit the processing of sensitive judicial data by any entity subject to foreign legal compulsion inconsistent with Indian law.

On training data, Regulation 46(4)(k) prohibits re-training, fine-tuning, or modifying AI models using court data without AI Committee approval. But this is a contractual term applicable to private vendors, not a prohibition in the regulations themselves. The prohibition should be elevated to a non-derogable regulatory requirement applying to both private vendors and public bodies including NIC, and should cover the use of court data for any training, fine-tuning, or evaluation purpose without the express written approval of the AI Committee.

Recommendation:

- Establish, within the instrument itself and as procedural, rules governing the Court's own processing of the data. And before that an enforceable set of data handling protections for adjudicatory processing such as, notice of AI processing, purpose limitation, data

[64] Clarifying Lawful Overseas Use of Data (CLOUD) Act, 2018 (U.S.); *See also*, *CLOUD Act Resources*, U.S. Department of Justice: Criminal Division, available at https://www.justice.gov/criminal/cloud-act-resources.

[65] NVIDIA Success Stories, *How India's National Informatics Centre Is Modernizing the Indian Judiciary With AI*, NVIDIA, available at https://www.nvidia.com/en-us/case-studies/role-of-nic-modernizing-indian-judiciary-using-ai/; NVIDIA Cloud Agreement, Clause 16, available at https://www.nvidia.com/en-us/agreements/cloud-services/nvidia-cloud-agreement/.

minimisation, and avenues for access, correction, erasure and grievance. Regulations 10 and 47 should state these expressly, with a transitional provision for the period before the DPDP Act's substantive provisions commence on 13 May 2027.

- Insert a new provision in chapter VII: "Where personal data is processed through an AI System in the performance of an adjudicatory function, the Court shall, as a matter of procedure – (a) inform the parties of such processing; (b) limit the processing to the purpose of the proceeding; (c) collect and retain only such personal data as is necessary for that purpose; and (d) provide a means by which a party may seek access to, correction or erasure of, and raise a grievance concerning, such personal data. Until the substantive provisions of the Digital Personal Data Protection Act, 2023 are in force, the Information Technology Act, 2000 and the rules made thereunder shall apply to such processing in addition to this Regulation."
- Amend Regulation 46(4)(j): define 'sovereign cloud' as cloud infrastructure where all processing, storage, and access control is subject exclusively to Indian legal jurisdiction, and where no entity with authority over the infrastructure is subject to foreign legal compulsion inconsistent with Indian law. Prohibit the processing of sensitive judicial data by entities subject to such compulsion.
- Amend Regulation 46(4)(k): elevate the training data prohibition from a contractual condition to a non-derogable regulatory requirement applying to both private vendors and public bodies. Prohibit the use of any court data sensitive or otherwise for training, fine-tuning, evaluation, or benchmarking any AI model without the express written approval of the AI Committee, which should record reasons for any approval granted.

4.4 Procedural Fairness and Litigant Rights

4.4.1 Absence of Party-Side Rights Within Proceedings

The draft establishes extensive governance structures and institutional safeguards but provides very few rights directly exercisable by litigants. HITL requirements (Regulation 3(zb)), audits, oversight mechanisms, and disclosure obligations primarily operate at the institutional level. (Regulation 8); and disclosure under Regulation 43 tells the party that AI was used without giving any means to respond to it.

Regulations 4 and 20(b)&(c) already mandate HITL across all adjudicatory functions: no AI system may perform adjudication or sentencing without mandatory human review, and no judicial outcome may be reached through algorithmic decision-making alone. The gap is the lack of any defined means with the litigant to learn that an AI system has "materially" assisted the court in any step affecting the matter,[66] and seek human reconsideration of that step, or to obtain an explanation addressed to them.

[66] Regulation 43(1), 2026 Regulation for the Use of Artificial Intelligence (AI) in courts.

This gap is further exacerbated by the general lack of knowledge of the understanding of AI systems, especially in India, which has ranked last in a survey of over 25 countries.[67] Where an automated process contributes to a determination affecting personal liberty or a substantive legal right, the affected party should have access to a meaningful and particularised account of that contribution.[68] A general disclosure that AI was used in the proceedings does not satisfy this requirement.

The UNESCO Guidelines' Principle 1.12 (Accountability and Contestability) goes further: it requires mechanisms that allow affected parties to contest and cross-examine any output produced by an AI system that may influence the outcome of their case, and to hold judicial operators accountable for errors.[69] The CEPEJ Charter's 'under user control' principle similarly requires that users be informed actors in control of their choices.[70] This is the area in which the draft most clearly falls below the international standard.

The absence is especially conspicuous given the nature of the instrument. Most comparable frameworks are non-binding guidance documents addressed to judges. India's draft is a binding regulatory instrument. A binding regulation that creates extensive institutional obligations, but no corresponding individual entitlements, produces an accountability structure with no external enforcement mechanism.

Recommendations:

- Confer a short set of procedural party side entitlements, framed as rules of procedure within the courts' competence: to be informed, in an accessible manner, that an AI system has materially assisted a step affecting the party's matter; to request human reconsideration of that step by the responsible officer; to a meaningful explanation, in matters affecting rights or liberty; and to basic information on how the system works and on what data it was trained, linked to the grievance route in Regulation 52.
- Insert a new provision: "A party to a proceeding shall be entitled: (a) to be informed, in an accessible manner, where an AI System has materially assisted a step affecting the party's matter; (b) to request human reconsideration of such step by the responsible officer;[71] and (c) in matters affecting rights or liberty, to a meaningful explanation of the basis of the AI-assisted

[67] The Hindu Data Team, *according to a survey of 25 countries, Indians are least aware of AI*, The Hindu, October 31, 2025, available at https://www.thehindu.com/data/according-to-a-survey-of-25-countries-indians-are-least-aware-of-ai/article70220820.ece.
[68] Danielle Keats Citron, *Technological Due Process*, Vol. 85, Wash. U. L. Rev., 1249 (2008).
[69] UNESCO Guidelines, *supra* note 19, Principle 1.12.
[70] CEPEJ Charter, *supra* note 4.
[71] Christen et. al., *Who Is Controlling Whom? Reframing "Meaningful Human Control" of AI Systems in Security*, Vol. 25, Ethics & Info. Tech. 10 (2023)

step. A party may raise any such matter through the grievance mechanism under Regulation 52."

4.4.2. Grievance redressal confined to prohibited uses

Regulation 52(1) gives a dedicated grievance route only for harm caused by a "*prohibited use of AI under regulation 20*." Harm from a permitted but defective tool, such as an inaccurate translation or transcription that prejudices a party, falls outside it and into the general saving in Regulation 53. When in the legal profession precision and accountability is paramount, the homogenization of linguistic diversity and the perpetuation of biases inherent in training datasets threaten the cultural integrity of translated texts.[72]

Regulation 53 preserves the ordinary remedies, so the harm still has remedies, but those remedies are slower and less accessible than the Regulation 52 mechanism and permitted use error is in practice the more likely source of harm.

The UNESCO Guidelines recommend that grievance mechanisms extend to all AI-related harms, be simple and multilingual, and be capable of ordering correction, disclosure, suspension, review, and remedies.[73] Regulation 52(1) further restricts applications to "any party to a proceeding." Where AI-related harm affects witnesses, legal representatives, court staff, or third parties through a data breach or erroneous surveillance output, no mechanism is available.

Recommendation:

- Broaden Regulation 52(1) to cover harm caused by any use of AI in a Court process, whether prohibited or permitted, retaining the requirement of a reasonable opportunity to be heard.
- Amend Regulation 52(1): " Where any harm is caused to any party to a proceeding as a direct or indirect effect *of any use of AI in a Court process*, he or his legal representative, may file an application at the earliest opportunity to the Court in which the relevant AI System was or is being used, and such Court shall, after giving a reasonable opportunity of being heard, pass appropriate orders as it may deem fit."

4.5 Enforcement and Liability

A significant structural gap in the draft is the absence of any enforcement architecture. Regulation 23(a) assigns the Apex Body a function of ensuring compliance with the Constitution and applicable law, but no provision specifies what follows when a violation is identified. Regulation 21 provides for

[72] Roya Shahmerdanova, *Artificial Intelligence in Translation: Challenges and Opportunities*, 2 Acta Globalis Humanitatis et Linguarum, no. 1, at 15 (Veris, 2025).
[73] UNESCO Guidelines, *supra* note 19, Principle 1.12.

remedial measures after a violation of the prohibitions in Regulation 20 principally suspension of the relevant AI system but operates prospectively only. There is no provision for retrospective correction of harm caused by a violation; civil liability of the vendor for system failures, disciplinary consequences for Designated Officers who fail to comply with the verification protocol, compensation for litigants whose proceedings were affected; or notification to affected parties that a systemic failure has been identified.

A governance instrument without enforcement mechanisms shifts the cost of non-compliance from the violating institution or vendor to the affected litigant, who must seek relief through ordinary legal remedies that are slower, more expensive, and more complex than the specialised mechanism the draft promises. The UNESCO Guidelines' Principle 1.12 on accountability and contestability requires not only the ability to contest AI outputs prospectively but the ability to hold judicial operators accountable for errors retrospectively.[74] The draft's current architecture satisfies only the first requirement.

This absence of an enforcement architecture raises two liability questions that require particular attention. First, Regulation 8(1) provides that accountability for all AI-assisted decisions rests with the supervising officer, and that the AI system's opaqueness or hallucination cannot excuse a palpably incorrect decision. This correctly places human accountability for the decision on the human. However, it does not address vendor liability for a system that performed differently from what was specified in the TEIA or approval documentation. Regulation 46(4)(i) provides for indemnity clauses in contracts with private vendors, but contractual indemnity is weaker and less accessible to affected litigants than statutory vendor liability. Second, there is no mechanism for retrospective review of proceedings in which a subsequently identified AI failure may have affected the outcome.

Recommendations

- New provision required (new Chapter or expanded Chapter IX): create an enforcement architecture including: (a) a right of affected persons to compensation where harm is caused by a prohibited use or by a defect in an authorised system attributable to the vendor's failure to meet stated specifications; (b) statutory vendor liability for AI systems that produce output materially inconsistent with performance specifications, operating independently of and not dischargeable through the contractual indemnity under Regulation 46(4)(i); (c) a mechanism for retrospective review of proceedings where a subsequently identified AI failure may have affected the outcome, with standing for affected parties to apply; (d) disciplinary consequences for Designated Officers who fail to comply with the verification protocol, separate from their

[74] UNESCO Guidelines, *supra* note 19, Principle 1.12.

broader accountability for the decision; and (e) a duty on the AI Committee to notify affected litigants where a systemic AI failure has been identified that may have affected their proceedings.

- Amend Regulation 21: extend the remedial measures following a violation of Regulation 20 to include compensation for persons harmed by the prohibited use, retrospective review of affected proceedings, and public notification of the violation.
- Amend Regulation 23(a): link the Apex Body's compliance function to the enforcement architecture by requiring that where a systemic violation is identified, the Apex Body directs the relevant AI Committee to notify affected litigants and initiate retrospective review within a specified period.

5. Conclusion

The Draft Regulations for Use of Artificial Intelligence in Courts, 2026 represent a serious and structurally ambitious attempt to govern AI within the Indian judicial system. No comparable jurisdiction has yet produced a binding, comprehensive judicial AI governance framework of this scope, and the initiative reflects a genuine institutional commitment to responsible technology adoption.

The Centre's recommendations are directed at strengthening the draft's regulatory purpose. We believe these suggestions could make the draft's underlying regulatory vision practically enforceable, grounded, and technically credible. They would ensure that adoption of AI in courts occurs within a framework that reflects the actual capabilities and limitations of contemporary AI systems, the distinct risks posed by different use-cases, and the institutional realities of the Indian judiciary. Greater clarity in scope, terminology, standards, accountability, and implementation would reduce uncertainty, prevent formal obligations from becoming unworkable in practice, and provide courts, judicial officers, litigants, and service providers with a more coherent basis for compliance. We hope the committee considers these comments and the Centre would be happy to expand or provide further clarity on any of the recommendations provided above.

The JSW Centre for the Future of Law is grateful for the opportunity to submit these comments and remains available to assist the AI Committee in any further technical or academic capacity during the finalisation of the Draft Regulations.