Swami Vivekananda Advanced Journal for Research and Studies
Online Copy of Document Available on: www.svajrs.com

ISSN:2584-105X **Pg. 02-10**



Reforming Criminal Justice Administration through Artificial Intelligence: Opportunities and Constitutional Safeguards in India

Dr. Mithilesh Kumar Yadav
Assistant Professor, College of Law, IIMT University, Meerut, Uttar Pradesh.

Dr. Amitesh Anand
Associate Professor, College of Law, IIMT University, Meerut, Uttar Pradesh.

Accepted: 25/10/2026
Published: 04/11/2026 ***DOI: https://doi.org/10.5281/zenodo.24101669***

Abstract

India's criminal justice system is under pressure and goodwill has not been much of a balm. There are more than 5 crore cases pending in courts at all levels. According to the latest prison statistics, about three out of every four persons in Indian prisons are not convicted of any crime. There are no investigators in police stations, no examiners in forensic laboratories and legal aid, promised under article 39A of the Constitution, comes too late to too few. The State has been increasingly enthusiastic about artificial intelligence here. Facial recognition for policing, crime-mapping software, machine translation of judgments, transcription tools, and research assistants for judges are no longer laboratory curiosities. These are housed within the Crime and Criminal Tracking Network and Systems (CCTNS), the eCourts Mission Mode Project and the pilot projects of a few High Courts and have largely come in without a statute to govern them.

The paper poses a question that enthusiasm tends to overlook. What is the constitutional basis for use of such technologies in the criminal process in India? It begins with the assumption that the question is not simply technical or administrative. A criminal trial is one place where the State can take away a person's liberty by a procedure. And that procedure itself must be right, just and fair. Maneka Gandhi Any tool that affects who is investigated, who is arrested, who is released and whose evidence is believed thus crosses into constitutional territory – regardless of the intentions of its designers.

The study is doctrinal and comparative in nature. It measures the Bharatiya Nagarik Suraksha Sanhita 2023, the Bharatiya Sakshya Adhiniyam 2023 and the Digital Personal Data Protection Act 2023 against the line of Supreme Court authority from Maneka Gandhi through Selvi, Puttaswamy and Anuradha Bhasin, and it tests the Indian position against experience elsewhere: the Wisconsin decision in Loomis, the English Court of Appeal's ruling in Bridges on facial recognition, the Ethical Charter of the Council of Europe's CEPEJ, and the European Union's Artificial Intelligence Act. It also draws on empirical work on algorithmic bias and automation bias, because an argument about legal safeguards is only as strong as its account of how these systems actually fail.

The first finding is that the opportunities are real and should not be dismissed in the name of caution. Technology that can triage forensic backlogs, flag undertrial prisoners who have served the period which section 479 of the BNSS deems sufficient for release, translate judgments into regional languages, or help an overstretched legal aid lawyer organize a file, serves values that the Constitution itself protects: speedy trial, equal access to justice and reasoned adjudication. International studies of bail prediction indicate that well-constructed tools can in principle reduce unnecessary detention and avoidable offending. For an Indian prisoner waiting years for a hearing, there is little solace in theoretical objections to a machine that could have advanced his case.

The second is that the risks are of a particular kind and existing law does a bad job of handling them. Opaque proprietary models can't be interrogated. Training data based on Indian policing and prison records will reflect the overrepresentation of some castes, communities and economic groups in the system and a model trained on that data may entrench that as a prediction. Officials and judges that receive a score tend to defer to it. Facial recognition error rates are unevenly distributed, and its deployment in public spaces without a statutory footing sits uneasily with the requirements of legality, legitimate aim and proportionality laid down in Puttaswamy. Meanwhile, the Digital Personal Data Protection Act excludes processing by State agencies for the investigation of offenses in broad enough terms to put much of this activity beyond its protections.

Third, the paper asserts that the constitutional framework is sufficient in principle but lacking in practice. Article 14 prohibits manifest arbitrariness and therefore provides a basis for challenging an unexplained outcome of an algorithm. Article 21 demands a fair procedure, which on any reasonable reading, demands disclosure of the material on which liberty is curtailed and a real opportunity to contest it. Article 20(3) and the reasoning in Selvi circumscribe the technologies which elicit mental or physiological evidence from an accused. There is no lack of principle, but there is a lack of legislation, procurement discipline and judicial guidance. Currently deployment is largely through tenders, executive notifications and administrative circulars, none of which meets the need of a restriction on a fundamental right being authorized by law.

The paper ends by proposing a graduated framework, rather than a blanket ban or open licence. Uses that support administration, such as transcription, translation and scheduling, would require only transparency and periodic audit. Uses that inform decisions about liberty, such as risk scoring in bail or flagging suspects, would require statutory authority, disclosure to the accused, independent validation on Indian data and a duty on the decision-maker to give reasons beyond simply repeating the output. Prohibited are uses that supplant a machine's adjudicative judgment of guilt, sentence, or credibility, and uses that infer mental states from physical signals. Judicial adoption policies such as the one issued by the Kerala High Court in 2025 are a useful starting point, but a policy addressed to judges cannot discipline the police or private vendors.

The gap has to be closed by Parliament or failing that by guidelines of the Supreme Court in the manner of Vishaka and D K Basu. At the center is a modest but firm claim. Artificial intelligence can assist India's criminal justice administration without violating its constitutional commitments, but only if the assistance is answerable to human beings who can be questioned, reversed and, when necessary, blamed. Efficiency is a proper constitutional issue. It was never meant to be bought fairly.

Keywords: *Artificial Intelligence; Criminal Justice Administration; Constitutional Safeguards; Proportionality; Algorithmic Accountability*

1. Introduction

Every few months some High Court in India notes with barely concealed exasperation that an accused has been in custody for years, waiting for a trial that has not yet started. This is an age-old complaint. In Hussainara Khatoon, the Supreme Court found that the jails of Bihar were housing under trial prisoners who had already served more than the maximum sentence for the alleged offenses charged.[1] The facts are different forty years on, the structure is not. National Judicial Data Grid shows pendency is over five crore cases[2] and three out of every four prisoners in India are undertrials.[3]

And into this congestion has artificial intelligence (AI) now been invited. And very little ceremony attended the invitation. Facial recognition is being tried by police forces; analytical software maps crime by locality; translation tools convert Supreme Court judgments into regional languages; courts are experimenting with transcription and research assistants. All this happened when Parliament was busy with something else - the replacement of the colonial criminal codes by the Bharatiya Nyaya Sanhita 2023 (BNS), the Bharatiya Nagarik Suraksha Sanhita 2023 (BNSS) and the Bharatiya Sakshya Adhiniyam 2023 (BSA) that came into force on July 1, 2024.[4] The new codes talk freely of electronic records, audio-video recording and forensic investigation. They barely talk about algorithms.

That silence is important. A criminal trial is a procedure by which the State can deprive a person of his liberty by a procedure of its own making. Since the Maneka Gandhi case, that procedure has to be right, just and fair, not arbitrary, imaginary or oppressive.[5] Part of that process, whether its vendors see it that way or not, is a tool that decides who is investigated, who is arrested, who is let out on bail, and whose evidence is believed. The real question is not whether technology should come into the courtroom and the police station. It's already in there. The question is on what conditions?

My argument is in three steps. First, the case for AI in criminal justice is more robust than its critics typically admit, since delay and under-resourcing are constitutional injuries in their own right. Second, the objections are more serious than its proponents typically admit, because the typical failures of these systems - opacity, inherited bias, unwarranted deference - are visited precisely on those least able to contest them. Third, most of what is needed to regulate the technology is already in India's constitutional doctrine. But nobody has converted that doctrine into rules that a police station or a vendor has to actually follow. The method is doctrinal and comparative. I refer to Indian case law, the three new codes, the Digital Personal Data Protection Act 2023 and experience in the United States, United Kingdom and European Union. The ambition is modest: to draw a framework in which efficiency and fairness are not rivals.

2. Where Has AI Already Arrived?

Any serious discussion has to begin by discussing what is actually in use, not what is announced at conferences. The picture is patchy, and much of it is less dramatic than the headlines imply. The foundation is data. The Crime and Criminal Tracking Network and Systems (CCTNS) links police stations across the country through a common database of FIRs, charge-sheets and arrest records, and the Interoperable Criminal Justice System is meant to link that database to courts, prisons, prosecution and forensic services.[6] This architecture has nothing intelligent per se. But it is the raw material on which all subsequent analytical tools depend. Its quality, as anyone who has ever seen a name entered incorrectly or a FIR duplicated will know, is variable.

On the court side, the E-courts Mission Mode Project has passed through various phases, the third phase of which was approved by the Union Cabinet in September 2023 with an outlay of more than seven thousand crore rupees.[7] The Supreme Court has developed a translation tool named SUVAS and a research-assistance portal SUPACE under this project, the latter introduced as a means to help judges manage

[1] Hussainara Khatoon v Home Secretary, State of Bihar (1980) 1 SCC 81.
[2] National Judicial Data Grid Available at https://njdg.ecourts.gov.in (last visited on 15/10/2026)
[3] National Crime Records Bureau, *Prison Statistics India 2022* (Ministry of Home Affairs 2023).
[4] Bharatiya Nyaya Sanhita 2023; Bharatiya Nagarik Suraksha Sanhita 2023 (BNSS); Bharatiya Sakshya Adhiniyam 2023 (BSA). All three took effect on 1 July 2024.
[5] Maneka Gandhi v Union of India (1978) 1 SCC 248.
[6] National Crime Records Bureau, 'Crime and Criminal Tracking Network and Systems (CCTNS)' https://ncrb.gov.in.
[7] eCourts Mission Mode Project https://ecourts.gov.in; Press Information Bureau, 'Cabinet Approves eCourts Phase III' (13 September 2023).

big records without making any decisions for them.[8] It's in policing that the constitutional thermometer heats up. Several state police forces have adopted facial recognition, and the National Crime Records Bureau has sought to develop a national automated system.[9] In March 2020, the Union Home Minister had told Lok Sabha that facial recognition was used to identify persons in connection with the Delhi violence.[10] Predictive or 'hotspot' crime mapping, such as the infamous Delhi Police's CMAPS, is based on historical incident data to guide patrols.[11] These systems are rarely based on a statute. They are empowered, if at all, by a tender, by a government order, or by a general reference to the powers of police.

But the new codes do add one thing that matters here: a vocabulary of electronic process. Information about a cognizable offense can now be given through electronic communication; searches and seizures must be audio-video recorded; forensic experts must visit the scene in cases punishable with seven years or more; and trials, inquiries, appearances can be conducted in electronic mode.[12] The Supreme Court has already accepted that evidence given by video-conference is evidence in the true sense[13] and the BNSS carries the principle into statute. The codes don't touch the layer above the data, the models that score, rank, match and recommend.

3. The Opportunities

Investigation and forensic capacity

Section 176(3) of the BNSS requires that a forensic investigation be carried out at the scene for grave offenses.[14] The regulation is commendable, but it will be unworkable unless the forensic laboratories can reduce their backlogs. The forensic report is the document awaited by those in the trial courts, and they know how often it appears. Technological support in DNA mixture analysis, fingerprint assessment, document examination, and digital forensic prioritization will not substitute the examiner but may enable the examiner to more efficiently allocate their time. An investigator who can review a large phone extraction in one afternoon, rather than a month, is less likely to detain a suspect for that amount of time. This is the mundane end point of AI, and probably where the benefits are most significant.

Delay, bail and the undertrial

The most persuasive argument is about freedom. Section 479 of BNSS replaces section 436A of Code of Criminal Procedure 1973 which provides that an undertrial prisoner who has undergone one-half of the maximum punishment for the offense shall be entitled to be released on bail and in the case of a first-time offender after undergoing one-third of such punishment, save in the case of a death sentence or imprisonment for life.[15] Eligibility is basically a matter of simple math: the section that applies, the date of arrest and the number of cases pending. Correctional officials and legal service agencies do the job with considerable delay and manually. A system that automatically identifies inmates that are eligible for human verification would create an additional authority, rather than establishing a legal entitlement. This is the same concern that the Supreme Court had in *Satender Kumar Antil* of arbitrary arrests and treating bail as the exception.[16]

Predicting risk is more difficult. Kleinberg and colleagues looked at bail decisions in New York and found that a machine-learning algorithm applied to the same defendants could, in theory, have reduced crimes committed while on bail by about a quarter without increasing detention rates, or reduced the number of people in jail by about 40 percent without an increase in crime.[17] The writers were careful of the constraints of that discovery, and we should be, too. What they had was American data, the legal evaluations differ, and an Indian judge deciding on bail looks at a statute, a case diary and an application not a score. But the research disputes the comforting idea that human judgment alone is sufficient protection from error. Humans make mistakes, and not at random either.

Access to justice

Article 39A directs the State to ensure the legal framework promotes justice on the basis of equal

[8] eCourts Mission Mode Project (n 8).
[9] Smriti Parsheera, 'Adoption and Regulation of Facial Recognition Technologies in India: Why and Why Not?' (Data Governance Network Working Paper 05, 2019).
[10] Lok Sabha Debates (11 March 2020).
[11] Vidushi Marda, 'Artificial Intelligence Policy in India: A Framework for Engaging the Limits of Data-Driven Decision-Making' (2018) 376 Philosophical Transactions of the Royal Society A 20180087.
[12] BNSS (n 5) ss 105, 173, 176(3), 530.
[13] State of Maharashtra v Praful B Desai (2003) 4 SCC 601.
[14] BNSS (n 5) s 176(3).
[15] BNSS (n 5) s 479 Code of Criminal Procedure 1973, s 436A (repealed).
[16] Satender Kumar Antil v Central Bureau of Investigation (2022) 10 SCC 51.
[17] Jon Kleinberg and others, 'Human Decisions and Machine Predictions' (2018) 133 Quarterly Journal of Economics 237.

opportunity. The Supreme Court has traditionally construed free legal aid as falling within Article 21.[18] In *Hoskot,* Krishna Iyer J. viewed legal aid as an element of the fair process and not an act of charity.[19] Language is the subtle barrier. The majority of defendants do not speak English fluently and the majority of High Court and Supreme Court decisions are in English. Machine translation has all its flaws, but it provides a translation in a language that a litigant or a paralegal can read. Inexact translations are permitted, if they are marked as additional and the original text is dominant.

Evidence and records

An electronic record is admissible under section 63 of the BSA, if it is accompanied by a certificate identifying the record and the device. In *Arjun Panditrao Khotkar*, the Supreme Court held that the certificate is a condition of admissibility and not a formality.[20] The use of instruments that hash, timestamp and document the chain of custody of digital evidence can improve the reliability that the certification system seeks to establish. The reasoning works in the other direction too. The certificate was never meant to authenticate content that was created or improved, such as a better CCTV image or a restored audio clip, which is a different category than a record.

4. The Constitutional Framework

There is no doctrine of algorithmic due process in Indian constitutional law. You don't need anything like that. The materials exist, they are scattered between decisions on arbitrariness, secrecy, self-incrimination and fair trial and the task is to read them together.

In Shayara Bano, the Supreme Court has recognized that laws can be struck down on the ground of manifest arbitrariness.[21] The same principle has historically applied to executive action. That's exactly what the doctrine condemns: a decision that can't be explained. You have an algorithmic result that can't be explained, produced by a model that can't be examined. The Court was of the opinion that the duty to give reasons is a fundamental part of administrative and quasi-judicial decisions and is of particular importance where personal freedom is concerned.[22] The software marked him as high-risk. That is no excuse.

Puttaswamy says any restriction on privacy has to meet the tests of legality, legitimate aim and proportionality.[23] The later Aadhaar judgment evolved the proportionality analysis into four stages: legitimate aim, rational connection between the aim and the action, necessity (no equally effective but less intrusive alternative) and fair balancing between the benefit gained and the right infringed.[24] Mostly the framework is from Modern Dental College.[25] The debate around public spaces usually concludes with the legal stipulation of facial recognition. There is no law that permits it, no rule that limits retention, no external oversight, and no way to complain. A deployment based on a tender document cannot easily be described as law in the way Puttaswamy required. The closest analogy for how the Court might treat ambiguous surveillance comes from Anuradha Bhasin where it applied a proportionality test to permanent internet blackouts and ordered publication and justification of orders.[26]

Selvi held that the non-consensual administration of narcoanalysis, polygraph and brain-mapping tests violates the protection against self-incrimination and substantive due process, as the information extracted from the accused's psyche is testimonial in nature, no matter how sophisticated the methods.[27] It was a question of compulsion, while the possibility of voluntary administration with safeguards remained open. When emotion-recognition software is applied to a video the defendant did not know was being viewed, the difference is disconcerting. This objection applies when evidence that purports to detect deception, stress, or emotion by facial expression, tone of voice, or gait is used against the defendant, regardless of how the technology is rebranded.

The Supreme Court has held that the criminal procedure is primarily aimed at ensuring a fair trial and a criminal court cannot remain a mere spectator.[28] If a defendant cannot understand how a machine-generated

[18] Constitution of India 1950, art 39A.
[19] Madhav Hayawadanrao Hoskot v State of Maharashtra (1978) 3 SCC 544.
[20] BSA (n 5) s 63; see also Arjun Panditrao Khotkar v Kailash Kushanrao Gorantyal (2020) 7 SCC 1.
[21] Shayara Bano v Union of India (2017) 9 SCC 1.
[22] Kranti Associates Pvt Ltd v Masood Ahmed Khan (2010) 9 SCC 496.
[23] Justice KS Puttaswamy (Retd) v Union of India (2017) 10 SCC 1.
[24] Justice KS Puttaswamy (Retd) v Union of India (2019) 1 SCC 1.
[25] Modern Dental College and Research Centre v State of Madhya Pradesh (2016) 7 SCC 353.
[26] Anuradha Bhasin v Union of India (2020) 3 SCC 637.
[27] Selvi v State of Karnataka (2010) 7 SCC 263.
[28] Zahira Habibulla H Sheikh v State of Gujarat (2004) 4 SCC 158.

document was produced, he cannot evaluate it, and if he cannot evaluate it, he has not had the kind of trial that Zahira Habibulla spoke of. D K Basu, through its exhaustive protections on arrest, shows that if the political entities do not do so, the Court will set out operational guidelines.[29]

The weaknesses of data security and the current legal framework. Processing in relation to prevention, detection, investigation or prosecution of offenses is exempt under section 17(1)(c) of the Digital Personal Data Protection Act 2023. Section 17(2)(a) of the Digital Personal Data Protection Act 2023 allows the Central Government to exempt certain state entities on grounds such as security and public order.[30] Despite the benefits of these exceptions, they create a gap exactly where criminal justice AI operates. It is not possible to ask a court to interpret a data protection framework. It is capable of, and Puttaswamy urges it to, assert that the exemptions are inherently liable to proportionality assessment.

Another point concerns transparency. In *Swapnil Tripathi,* the Court accepted that access to judicial proceedings is one aspect of access to justice under Article 21 and permitted live-streaming of constitutional cases.[31] There are constitutional endorsements for advancements in technology that lead to transparency." Technological systems that slow the reasoning process are greatly reduced.

5. The Risks, and Why They Fall Unevenly

Opacity and proprietary secrecy

In *State v Loomis,* the Wisconsin Supreme Court ruled that the risk assessment instrument COMPAS could be used at sentencing, provided that a written warning be given regarding its proprietary limitations on disclosure, the non-determinative value of its scores, and the fact that its norming was based on a national sample.[32] The resolution is frequently referred to as a positive signal. It's more like amber. Rebecca Wexler has shown how American criminal proceedings have employed trade-secret protections to deny defendants access to source code and validation studies,[33] and Frank Pasquale's more general critique of undisclosed algorithms is especially salient when the decision at stake is custody.[34] Indian procurement, with its habit of confidentiality clauses, could reproduce the problem without much effort.

Inherited bias

In 2016, ProPublica found that COMPAS was more likely to label black defendants as likely to reoffend than white defendants, though the vendor said the tool was equally accurate across racial groups.[35] The dispute is instructive because each side was right on its own measure. Kleinberg, Mullainathan and Raghavan later demonstrated that, outside special circumstances, a risk score cannot satisfy several natural definitions of fairness at once.[36] The point is that equity is a choice among competing principles, and that choice should be explicit, not buried somewhere in the code of a programmer. Barocas and Selbst argue that a model trained on past decisions also incorporates the biases in those decisions.[37]

The Indian version of this problem is hardly hypothetical. The prison data shows a consistent pattern of Scheduled Castes, Scheduled Tribes and Muslims being overrepresented among undertrials, compared to their share of the population.[38] That does not by itself prove discrimination at the point of arrest, but it does mean that a model learning from arrest and custody records learns from a skewed sample, and that apparently neutral variables such as locality, prior arrests or family association may stand in for caste or religion. Ferguson's account of American predictive policing describes the resulting loop: patrols go where past records show crime, generate more records there, and the model takes the result as confirmation.[39]

Deference

No tool, no matter how transparent and calibrated, functions in a vacuum. Automation bias, the tendency to follow a computer's advice even when it is wrong, is common and difficult to train away, a thorough

[29] DK Basu v State of West Bengal (1997) 1 SCC 416.
[30] Digital Personal Data Protection Act 2023, ss 7, 17(1)(c), 17(2)(a).
[31] Swapnil Tripathi v Supreme Court of India (2018) 10 SCC 628.
[32] State v Loomis 881 NW 2d 749 (Wis 2016).
[33] Rebecca Wexler, 'Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System' (2018) 70 Stanford Law Review 1343.
[34] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015).
[35] Julia Angwin and others, 'Machine Bias' (*ProPublica*, 23 May 2016).
[36] Jon Kleinberg, Sendhil Mullainathan and Manish Raghavan, 'Inherent Trade-Offs in the Fair Determination of Risk Scores' (2017) 67 Leibniz International Proceedings in Informatics 43.
[37] Solon Barocas and Andrew Selbst, 'Big Data's Disparate Impact' (2016) 104 California Law Review 671.
[38] NCRB (n 4).
[39] Andrew Guthrie Ferguson, *The Rise of Big Data Policing* (NYU Press 2017).

review of the medical literature found.[40] And a magistrate with forty bail cases to get through before lunch is just the man who may think a score is an expedience. The reassurance that a human remains in the loop therefore means little unless the human is required to reason independently and given the time in which to do it.

Surveillance and error

In Bridges the Court of Appeal held that South Wales Police's use of automated facial recognition was not in accordance with law, because the framework left too much to the discretion of individual officers, and that the force had failed to inquire whether the software performed differently across racial and gender groups.[41] The dilemma was a matter of two faults of a found legal structure and a full investigation of prejudice. At this time neither is present in India.

Fabrication

One last risk is newer, and more humiliating. The generative models produce In Mata v Avianca, a New York federal court sanctioned attorneys for filing a brief citing fake rulings created by a chatbot.[42] If a judge or attorney uses generated text without checking it, they haven't used a tool. They have outsourced a job. Verification is no courtesy in a framework where a citation can mean the difference between bail and detention.

6. Comparative Experience and a Framework for India

Three external reference points will be useful, but none should be copied in their entirety.

The Artificial Intelligence Act of the European Union, Regulation (EU) 2024/1689, bans certain practices, including (subject to limited exceptions) real-time remote biometric identification in a public area for law enforcement purposes and the assessment of a person's likelihood to commit an offense on the basis of profiling alone. It considers most other law enforcement and justice administration systems to be high risk and requiring risk management, data integrity, record keeping, transparency and human supervision.[43] The CEPEJ Ethical Charter, established by the Council of Europe, is older, but more lenient (2018). It has 5 tenets: respect for fundamental rights, non-discrimination, quality, and security, transparency and fairness, and user empowerment. These have influenced subsequent regulations.[44] The Kerala High Court in 2025 came out with a policy for the district judiciary that restricts judges to sanctioned instruments and debars the use of AI for making determinations, granting remedies or passing orders.[45] It is a wise policy. That is a directive to judges, and does not reference any actions that could be taken by law enforcement or prosecutors.

The directive from the Punjab and Haryana High Court in the Jaswinder Singh case, where a judge utilized ChatGPT during the consideration of a bail application related to purported cruelty, serves as a valuable lesson for another purpose. The court had made it clear that the reference was for the purpose of giving a wider context and the ruling was not dependent on it.[46] The effort made to articulate this is, in itself, an acknowledgment that dependence on such an undisclosed tool would pose an issue. NITI Aayog's framework on responsible AI enumerates principles such as safety, equality, transparency, and accountability[47]; however, these are abstract and lack enforceability. An individual whose bail was denied based on a score cannot invoke a principle.

A graduated framework

I propose a framework focused on the impact of a tool on freedom, utilizing the proportionality assessment that Indian and comparative judiciary systems currently employ.[48] The concept is derived from the fundamental principle of administrative law that decisions affecting an individual's core interests warrant heightened procedural safeguards, which also

[40] Kate Goddard, Abdul Roudsari and Jeremy Wyatt, 'Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators' (2012) 19 Journal of the American Medical Informatics Association 121.
[41] R (Bridges) v Chief Constable of South Wales Police [2020] EWCA Civ 1058.
[42] Mata v Avianca Inc 678 F Supp 3d 443 (SDNY 2023).
[43] Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence, arts 5 and 6, Annex III.
[44] European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment* (Council of Europe 2018).
[45] High Court of Kerala, *Policy Regarding Use of Artificial Intelligence Tools in District Judiciary* (July 2025).
[46] Jaswinder Singh @ Jassi v State of Punjab CRM-M-22371-2022 (P&H HC, 27 March 2023).
[47] NITI Aayog, *Responsible AI: Approach Document for India, Part 1: Principles for Responsible AI* (February 2021).
[48] Aharon Barak, *Proportionality: Constitutional Rights and their Limitations* (CUP 2012).

underpins Danielle Citron's advocacy for technological due process.[49]

Entry level: administrative assistance. Transcription, translation, scheduling, document retrieval, and analogous tools that do not influence the essence of any decision would require merely the disclosure of their utilization, a log of identified errors, and regular audits. The meaning and tone of the following document are preserved in English. No information is added or omitted. No other text is added to the output. The human document takes precedence.

Secondary level: aid that infringes on freedom. Instruments that guide choices regarding arrest, bail, detention, investigative prioritization, or evidence handling would necessitate, at the very least, five essential components. A legal foundation is required that specifies the objective, the allowable data, and its duration of retention, as Puttaswamy necessitates legislation, and a tender document does not constitute law. The accused and the court require disclosure regarding the usage and sufficient details about the model to enable a challenge, with trade-secret objections subordinated to a protective inspection protocol. Independent verification of Indian data is required, with error rates categorized by caste, religion, gender, and region where applicable, alongside periodic audits conducted at regular intervals. A decision-maker is required to document, in their own language, the rationale behind the decision, ensuring that the output serves as a basis for judgment rather than a replacement for it. They require a solution, signifying an available platform where an impacted individual can request an evaluation, along with the removal of content produced in violation of these stipulations.

Third level: forbidden applications. A machine ought not to ascertain culpability, impose a penalty, or evaluate trustworthiness. The identification of emotions and the deduction of mental states from physical cues, which cannot be aligned with Selvi, ought to be banned in investigations and trials. The implementation of real-time facial recognition in public areas for general monitoring ought to be banned, with exceptions strictly confined to grave crimes, contingent upon prior judicial approval, and delineated in legislation.

Evidence produced, augmented, or reconstructed by a model should be accepted solely on the basis of a foundation detailing the model's identity and version, its validation history and known error rates, as well as its application, alongside the certification mandated by section 63 of the BSA.[50]

Dissenting views

Two objections warrant a response. The initial point is that these protections will hinder adoption and increase expenses. They shall, and that is the crux. A system that identifies its mistakes solely after an individual has been detained for months has not been time-efficient. The second criticism is that human adjudicators are also obscure and prejudiced, so why impose a stricter criterion on machines? There exist two responses. A human judge is subject to inquiry, appeal, and, in extreme cases, removal, with their rationale documented; a model's parameters are not. And the scale of a system matters: one biased judge affects his docket, while a biased model deployed across a State affects every docket at once.

The form of regulation is a separate matter. The cleanest route is a dedicated chapter in the BNSS or a standalone statute. Parliament is slow, however, and the Supreme Court has before now filled a legislative vacuum with binding guidelines, as in Vishaka case, and with operational safeguards, as in D K Basu.[51] A petition seeking comparable guidelines on facial recognition and risk scoring in criminal proceedings would not be an act of judicial adventurism. It would be a continuation of established practice.

7. Conclusion

Nobody who has watched a bail application adjourned for the fifth time is likely to be sentimental about the status quo. Delay is a form of punishment imposed before conviction, and a technology that reduces it has a claim on the Constitution’s sympathy. That claim should be taken seriously, which is why this paper has not argued for prohibition.

But sympathy is not license. The same technologies, introduced without a statute, without disclosure and without any means of challenge, would convert a justice system that is slow into one that is both slow and unaccountable. Indian constitutional law has the concepts it needs: arbitrariness, proportionality, fair

[49] Danielle Keats Citron, 'Technological Due Process' (2008) 85 Washington University Law Review 1249.

[50] BSA (n 5) s 63; see also Arjun Panditrao Khotkar (n 21).

[51] Vishaka v State of Rajasthan (1997) 6 SCC 241; DK Basu (n 30).

procedure, the privilege against self-incrimination, the duty to give reasons. What it lacks is a body of rules that turn those concepts into obligations for the people who buy, build and use these systems. Parliament should write them. If it does not, the Supreme Court has shown that it will.

The test I would apply to any proposed deployment is simple. Can the person affected find out that a machine was used, understand roughly how it reached its output, challenge it before someone willing to disagree, and obtain a remedy if it was wrong? Where the answer is no, the technology is not ready for the criminal process, however efficient it may be.

Disclaimer/Publisher's Note: The views, findings, conclusions, and opinions expressed in articles published in this journal are exclusively those of the individual author(s) and contributor(s). The publisher and/or editorial team neither endorse nor necessarily share these viewpoints. The publisher and/or editors assume no responsibility or liability for any damage, harm, loss, or injury, whether personal or otherwise, that might occur from the use, interpretation, or reliance upon the information, methods, instructions, or products discussed in the journal's content