Journal of Contemporary Law and Society, *Volume 2, Issue 4, December 2025, pp. 236-246*

TECHNICAL ARTICLE
Online ISSN: 3139-1079
Doi: 10.18311/JCLS/2025/v2i4/53258

INFORMATICS JOURNALS

Online ISSN : XXXX-XXXX

Journal of Contemporary Law and Society

Contents available at: *www.informaticsjournals.co.in/index.php/jcls*

JCLS

Bringing Bytes to the Bench: The Constitutional and Ethical Implications of AI in the Criminal Justice System

Aranya Nath[1*], Sovna Priyadarshini Parida[2], Parismita Nanda[3] and Kritika Poornomy Mohanta[4]

[1]Damodaram Sanjivayya National Law University, Visakhapatnam – 531035, Andhra Pradesh, India; subhamitanath002@gmail.com

[2]SOA National Institute of Law (SNIL), Bharatpur, Bhubaneswar - 751029, Odisha, India

[3]Madhusudan Law University, Cuttack - 753003, Odisha, India

[4]KIIT School of Law, Bhubaneswar - 751024, Odisha, India

Abstract

In the era of digitalisation, Artificial Intelligence (AI) adoption is rapidly increasing within the criminal justice system, where predictive policing algorithms, risk assessment tools, facial recognition systems, and digital forensics raise urgent constitutional and ethical challenges. The authors in this article will explore whether AI-powered innovations impact fundamental rights, including constitutional concerns, legal fairness, confidentiality, equality before the law, and the presumption of innocence. The present research compares the legal systems of India and the United States of America to examine the implications of algorithmic opacity, data bias, and automated decision-making. The Puttaswamy v. Union of India ruling, along with Articles 14, 19, and 21, sheds insight into the Indian context. At the same time, decisions on penalty systems and surveillance technology are relevant in the US context. Practical case studies, such as judicial recourse to risk-scoring systems for bail and sentencing in the United States and trial implementations of surveillance-based AI in India, demonstrate how such systems can unfairly target minority groups, worsen discrimination, and erode confidence in the justice system. Ethical analysis questions the propriety of outsourcing important judgments, such as release decisions or suspect identification, to machine-led systems that lack human empathy or contextual subtlety. It contextualises these issues within moral conceptions of justice, autonomy, and human dignity. The article advocates for essential security measures such as mandatory algorithmic transparency, consistent, accurate, and fair audits, legislative oversight, judicial review procedures, and "human-in-the-loop" systems that ensure genuine human agency in decision-making. It emphasises the significance of constitutional morality frameworks that balance AI deployment with constitutional principles and rights protections. Finally, the researcher emphasises that, while AI has the potential to improve efficiency in the criminal justice system, its abuse may jeopardise fundamental constitutional foundations. The article recommends multidisciplinary and policy-oriented frameworks to enable ethical and rights-respecting AI integration.

Keywords: *Artificial Intelligence, Constitutional Rights, Criminal Justice, Ethical Governance, Privacy*

1.0 Introduction

Integrating Artificial Intelligence (AI) into the Indian criminal justice system has the potential to transform law enforcement operations, increase efficiency, and improve outcomes. This study explores the opportunities, challenges, and ethical concerns related to using AI in the Indian criminal justice system. Like others worldwide, the

*Author for correspondence

Indian criminal justice system faces issues such as case delays, resource limitations, and the need for timely and equitable judgments. AI technologies can offer innovative solutions to address these problems. Applications in criminal justice include predictive policing, crime analysis, case management, and judicial decision-making. AI is fundamentally transforming the system and can significantly impact the Indian Penal Code (IPC) and the legal landscape. Technologies such as machine learning and data analytics are being incorporated to enhance efficiency, accuracy, and fairness. This introduction considers how AI is applied within the Indian Penal Code, including predictive policing, crime analysis, evidence management, and legal research. Since AI systems, especially those used by police and courts, depend on large datasets and machine learning algorithms, they may inadvertently reflect and reinforce existing biases. Their lack of transparency, often described as a "black-box" effect, makes it difficult for individuals to understand or challenge decisions effectively. In a domain where liberty and life are at risk, the implications of vague or opaque decisions are profound. This raises a critical question: how can AI be aligned with constitutional rights and ethical standards to ensure justice is not only done but also perceived to be done? This article examines the constitutional and ethical implications of deploying AI in criminal justice, with a focus on India but drawing comparative insights from countries like the United States. In the Indian context, the use of AI must be balanced against fundamental rights enshrined in Part III of the Constitution, particularly Articles 14, 19, and 21. These provisions protect equality before the law, freedom of expression, and the right to life and personal liberty, including, as established in "Justice K.S. Puttaswamy v. Union of India (2017), the right to privacy."(Bhatia, 2017)

Whether AI technologies impact bail judgments, execution, or monitoring procedures, there are challenges about procedural fairness, the presumed absence of guilt, and protection against arbitrary state action.

The ethical consequences are no less significant. Moral theorists raise concerns about accountability in machine decisions, the alienation of justice, and the ethical constraints of obtaining information and surveillance. Could robots trained on skewed historical information make objective decisions? Is it ethical to use prediction techniques that may disproportionately affect vulnerable areas? How do we make sure that AI does not replace human empathy with cold, statistical judgments?

1.1 Research Objectives

- To examine the constitutional issues raised by AI's incorporation into the Indian criminal justice system, especially those related to fairness, privacy, and equality.
- To examine the ethical challenges raised by algorithmic decision-making in legal scenarios utilising standards such as deontology, consequences, and justice concepts.
- To provide policies and protections that will assist in reconciling AI usage with democratic norms and constitutional morals.

2.0 AI Tools in the Criminal Justice Ecosystem

Crime affects a country's reputation, economic prosperity, and standard of living. Innovative crime analytics tools are necessary to protect communities and society from criminal behaviour. Al has proven to be a valuable instrument for legal proceedings. AI's applications in the criminal justice system have grown. AI is being used to address imbalances in the criminal justice system worldwide. The Criminal Justice System has restrictions such as pending cases, repetitive tasks, and a lack of technical skills during investigations. AI is being utilised for legal research and case management. Legal practitioners can utilise these technologies to simplify administrative procedures, identify precedents, and evaluate enormous amounts of legal data. *"The Supreme Court's AI-powered legal research and document analysis initiative)."* Law Artificial intelligence in criminal justice: Balancing efficiency with fairness and accountability. (n.d.). enforcement and courts are increasingly relying on AI-powered solutions to boost efficiency, enhance choices, and eliminate human mistakes. However, the use of such devices without sufficient legal supervision raises serious concerns about responsibility, openness, and equality. Predictive policing systems, facial recognition technologies, and algorithmic risk assessments are some of the most prevalent and disputable AI applications.

2.1.1 Predictive Policing

Predictive policing is the use of information analysis and machine learning algorithms to anticipate crime locations or detect possible criminals. These systems usually use past crime statistics, social behaviour, and demographic

information to create "crime heat maps" or risk profiles. While supporters assert that predictive policing enhances resource allocation and crime prevention, critics argue that it perpetuates current prejudices in historical policing data, disproportionately targeting minority groups. ('The Rise of Big Data Policing', n.d.-a) Prediction models frequently replicate the racial and socioeconomic biases of previous enforcement trends. For example, if specific regions were once overpoliced, the information will show greater criminal activity in those areas, leading to the process of monitoring and incarceration (Lum, K., & Isaac, W. (2016). The confidentiality of such algorithms makes it impossible to challenge or audit outcomes, weakening justice and due procedure.

2.1.2 Facial Recognition Technology (FRT)

Facial Recognition Technology (FRT) is a biometric identification method similar to fingerprint and iris scanning. Facial recognition technology's local feature analysis algorithms scan video and image content to measure features such as the shape of the chin, interocular distance, and other precise facial features, and then generate a mathematical description called a face template. Facial recognition technology uses facial features to identify individuals in various inputs (e.g., photos, video, or live feed) by using visible light and infrared wavelengths. The face template is a unique identifier for a person, similar to a fingerprint, being a biometric sequence.

Generally, FRT can do two things. One is to recognise the face of an unknown person by capturing the face of the person, doing a face template transformation, and comparing it with a database to check if one gets a match is found. Facial recognition technology is used by the police for this purpose. The other application is to authenticate the identity of a known person, where the image is compared with a known template, e.g., the facial recognition feature, to open a phone.

The state-of-the-art FRT uses machine learning techniques, such as deep learning. The components, when working together during training, "learn" to recognise features like edges and shapes that correspond to specific labels. Input data (scans of faces) are fed in as a pixel matrix and processed through many statistical components called "artificial neurons." These components calculate numerical weights based on characteristics like the colour of pixels and areal density. Multiple layers of these linked components are presented with many different labelled inputs, developing a heuristic that gives feedback to the weight adjustment. When the new input is presented, the units will produce high weights when identifying features seen in the training set. These weights are normalised and used to produce labels with a degree of confidence. Additionally, these systems will frequently include a liveness verification mechanism, which can include actions like blinking or a series of head movements, to ensure that the person being identified is a living organism.

FRT is defined by the trade-off between misidentifying someone's face and failing to recognise a face. Because of the risk of misidentification, often called the false positive rate (the likelihood of a false match), the majority of systems attempt to maintain this rate at the lowest possible level. But by reducing false positives, the false negative rate (the likelihood of a failure to recognise a face) is maximised. These are built into the accuracy of each facial recognition system. There are additional variables, such as lighting, background, viewpoint, pose, and facial expressions, that contribute to and can negatively impact the accuracy of FRT.

2.1.3 Supervision of FRT in India and Potential Solutions

Facial recognition technologies offer remote and covert bulk authentication. It implies that they operate without giving notice of their presence or usage, need no direct engagement with the subject, and are often not targeted to specific suspects, but rather meant to monitor everyone. As a result, any use of such technology for justice purposes raises severe privacy and free expression difficulties. Nonetheless, regulatory control of FRT in India remains limited.

3.0 Right to Privacy

In 2017, in the Puttaswamy judgment, the Supreme Court of India declared the right to privacy to be a fundamental right under the Indian Constitution. While the right is not absolute, reasonable restrictions on the right must satisfy a threefold requirement, i.e.,

(i) there being a law,
(ii) a legitimate state interest, and
(iii) proportionality. On the basis of this requirement, the use of FRT by the state, whether for law enforcement

or otherwise, amounts to a violation of the right to privacy.

But no legislative sanction for the State or other private agency to employ FRT exists today. The National Crime Records Bureau (NCRB) has asserted that FRT enjoys cabinet approval, and no legislative or executive order sanctioning the formation of the Automated Facial Recognition System (AFRS) is necessary. Cabinet approval is not a statutory order, and it does not grant any legislative sanction for the employment of facial recognition technology.

In contrast, within the European Union (EU), there are specific legislative provisions for authorising and restricting the use of FRT. For instance, Article 9 of the General Data Protection Regulation (GDPR) prohibits processing personal data revealing biometric data to uniquely identify a natural person. The prohibition does not apply, however, to data made available to the public by the data subject and where *"processing is necessary for reasons of substantial public interest."* Most FRT uses for non-law-enforcement purposes fall under these exceptions. One such serious concern policy paper that overviews FRT is a report published by Niti Aayog, a public policy think tank that works with the Indian government. The paper, *"Responsible AI for All: Adopting the Framework – A Use Case Approach on Facial Recognition Technology,"* is one of their series of Responsible AI papers. The report assumes that most of the privacy concerns over FRT will be addressed by the data protection law of India, which was then under consultation. Furthermore, the report does not furnish complete recommendations on how the law can enable the ethical use of FRT.

It does make a limited recommendation that the Central Government not exempt law enforcement agencies from the operation of the data protection law, as it might do under *"Section 18(1)(a). Even if the government had acted on the recommendation, however, it would not go very far in safeguarding against the ills of FRT."* One way in which data protection laws can limit the use of FRT is through the data minimisation principles. Proper application of the purpose limitation necessitates that the images taken by way of FRT cannot be supplemented with any other database to form a 360-degree profile of a citizen. But the tenders for AFRS currently on the table and discussed above contemplate the opposite and specifically invite the merging of databases. (Sinha, 2024)

3.1 Case Examples: COMPAS (USA) and Facial Recognition Technology (India)

3.1.1 The United States: Controversy over COMPAS

The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) program, developed by Northpointe Inc. (now Equivant), is the most prevalent risk assessment program in the United States. COMPAS has one sole function: to assist judges with bail, sentencing, and parole decisions by estimating the likelihood that a defendant will commit future crimes. COMPAS provides risk scores based on over 100 different variables, including criminal past, employment, educational background, and personal circumstances, many of which are not disclosed because the program is proprietary.

COMPAS court sentencing usage came under sharp criticism following the historic case of *"State v. Loomis in 2016. In this case, Eric Loomis challenged the use of COMPAS-computed risk scores during his sentencing on the argument that he could not understand or challenge the algorithmic rationale due to the algorithm's secrecy. The Wisconsin Supreme Court upheld the use of COMPAS but warned that such tools should never be the sole basis of sentencing and that the courts must notify defendants of the tool's*

Further concerns were raised when ProPublica conducted an extensive examination of the COMPAS system and indicated high racial bias in the predictive component of the tool. Under the examination, Black defendants were nearly twice as likely to be misclassified as high risk as their white counterparts, with white defendants being likely to be misclassified as low risk. (Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Contrary to the claim of objectivity, the outcomes generated by the tool were subject to the underlying data and the choice rules, showing the risk of using opaque algorithms in high-stakes legal proceedings.

3.1.2 India Implementation of Facial Recognition

In the Indian context, FRT has become a common feature of techno-based policing strategies in most of the bigger urban centres. The Delhi Police, as an organisation, have been at the forefront in the use of FRT in policing in the capital city. The technology was employed extensively during the protests and civil agitations in the capital, particularly in the Citizenship Amendment Act (CAA) protests and the

2020 Delhi riots. The officials justified the use of FRT as necessary to identify the rioters and for the preservation of public order; however, civil rights activists criticised its use as intrusive and discriminatory (Singh, 2020).

A key example was the application of facial recognition technology by the Delhi Police in the 2020 riots to recognise individuals in protest videos and compare them with existing databases like Aadhaar and voter ID cards. There were reports of more than 1,100 suspects being identified through the application of such a system. The absence of clarity about the accuracy, error rates, and auditing processes associated with such technologies was a matter of concern for human rights bodies (*Delhi Police's Claims That FRT Is Accurate at 80% Are 100% Scary*, 2022). There was no external assessment conducted to determine if the facial recognition technology was flagging individuals incorrectly, and there was no legal framework to regulate its application.

In addition, studies have shown that facial recognition technology readily available for purchase has higher error rates when identifying darker-skinned individuals, women, and youth populations that often compose a higher percentage of protestors and minority groups in India. If left unaddressed, biases can lead to false positives, unwarranted arrests, and a chilling effect on civic participation. (Buolamwini & Gebru, 2018) Compared to the United States, India lacks a clearly defined legal framework of data protection with robust surveillance technology-specific mechanisms of accountability. *"The historic judgment in Justice K.S. Puttaswamy v. Union of India (2017) established privacy as a fundamental right under Article 21; yet the unregulated spread of surveillance methods like facial recognition technology has advanced far ahead, with little public scrutiny or legislative* oversight (Bhatia, 2017)." This gap between law development and technology intrusion is a compelling case in favour of robust regulatory controls.

4.0 Constitutional Implications of AI In the Criminal Justice System

The judgment rendered by the Supreme Court in the case of "Justice K.S. Puttaswamy v. Union of India (2017) (Bhatia, 2017)" marks a landmark step in the protection of privacy as a fundamental right enshrined in Article 21 of the Indian Constitution." The judgment asserts that privacy is part of life and liberty and is entitled to protection unless withdrawn by law, based on the tests of fairness, justice, and reasonableness. The Court further elaborated on the fundamental principles of proportionality, necessity, and legitimacy concerning any government action that may infringe upon this right. The significance of the judgment is particularly critical in the context of AI-based surveillance, where the collection of big data and profile-based automation may pose a significant threat to individuals' privacy, often coupled with ineffective monitoring mechanisms.

The Puttaswamy judgment created a conceptual and legal framework to inform the government's conduct within the increasing use of digital surveillance and AI technologies. Heavily drawing on international human rights paradigms, such as the European Convention on Human Rights, the "*General Data Protection Regulation (GDPR)* (Linden *et al.*, 2020) and *US Supreme Court rulings, the judgment formulated the principle of informational autonomy, relating to personal information ownership.* ('(PDF) Recalibrating The Right To Privacy In The Digital Age', 2025)" the principle presumes particular significance in AI, where quantities of personal information are processed by algorithms that frequently lack sufficient user consent and transparency. The judiciary formulated a four-factor test-based evaluative framework for privacy rights restrictions. The four-step test is a constitutional framework utilised to determine whether AI surveillance technologies and digital forensic tools are legitimate. The court formulated a structured approach to test the legality of the state's action that invades privacy rights. It also suggested a long-term template for conceptualising privacy in the digital age, especially as law enforcement agencies increasingly use AI to monitor digital traces, intercept communications, and build profiles of individuals who are viewed as suspects. Puttaswamy contends that where AI is used to categorise an individual as a "potential threat" based on social media behaviour or patterns of behaviour, such activity must be governed by legal codes, directed at a legitimate purpose, proportionate in its application, and subject to procedural checks. (Warthon, 2024).

4.1 Article 14: Equality before law and Algorithmic Arbitrainess

Equal application under the law and similar treatment under the law. Under Article 14 of the Indian Constitution,

the phrase equality before the law and equal protection under the law is preserved, prohibiting arbitrary dealings with the state. Indian constitutional adjudication has developed the Article 14 requirement within a two-pronged test: the formal equality, equivalent to similar treatment of similarly situated persons; and substantive non-arbitrariness, as elaborated in *"E.P. Royappa v. Tamil Nadu (1974).* (*E. P. Royappa vs State Of Tamil Nadu & Anr on 23 November, 1973*, n.d.). Substantive non-arbitrariness demands that all state actions be based on rational classification and non-arbitrary standards.

AI systems often operate via opaque, data-driven processes that complicate conventional legal processes and concepts of openness. One such instance is predictive policing algorithms, which, when they are trained on past data, can disproportionately target communities that have been over-policed in the past, thereby creating a self-perpetuating cycle of criminalisation. (Lum & Isaac, 2016) This is against the principle of substantive equality since these communities get targeted with more surveillance based on past systemic discrimination rather than personal responsibility.

In addition, algorithmic arbitrariness is also problematic on doctrinal grounds under Article 14. Since the majority of AI systems function through closed proprietary black-box algorithms, those who are subject to these systems are denied access to important reasons or avenues of recourse. The case is starkly different from the legally entrenched principle that "arbitrariness is antithetical to equality," and that state action is amenable to judicial review in Maneka Gandhi v. Union of India. (*Maneka Gandhi vs Union of India on 25 January, 1978*, n.d.)" The use of black-box AI technologies in law enforcement and judicial decision-making, thus, disempowers constitutional accountability.

4.2 Articles 19 and 21Concern Autonomy, Liberty, and Inhibiting Consequences

"Articles 19 and 21 form the backbone of civil liberties protection in India. Article 19(1)(a) protects the freedom of speech and expression, and Article 19(1)(d) upholds the right to move freely within the territory of India. However, AI-enabled surveillance systems like FRT, already in use at protests, public transport hubs, and civic spaces, are gradually threatening these rights." This application has been specifically referred to as its dominant and chilling effect that causes people to self-censor and avoid lawful assembly due to the risk of retaliation, thus reducing the democratic breathing room in which to dissent and exercise civic rights.

Article 21, which guards life and personal liberty, has been broadly defined to include the right of privacy, procedural fairness, reputation and dignity. In *"Maneka Gandhi the Supreme Court said the procedure established by law should be just, fair and reasonable and not arbitrary. Any AI-based system that makes bail, sentencing, or parole decisions based on probabilistic models, and never gives an affected individual a chance to know how they make decisions or challenge the process, violates this procedure of due process."*

Even more importantly, the adoption of AI-based risk levels (e.g., the COMPAS system is used in the United States) to inform court decisions is questionable in India under Article 21. Arbitrary deprivation of liberty is the last thing that constitutional morality would embrace, no matter how well-intentioned the adoption of similar systems in Indian courts or in prisons may be, a lack of individualised assessment and the likelihood of automated stereotyping.

4.3 Issues of Consent, Surveillance, and Data Collection

AI technologies such as facial recognition technology and automatic number plate recognition technology are extensively used in public places without the particular or informed consent of the public. Puttaswamy stressed the autonomy principle, stating that privacy constitutes the ability to make personal decisions without uninvited interference. Lastly, indiscriminate data collection is not in line with the principle of purpose limitation, which requires data to be collected only for specific, legitimate, and essential purposes.

5.0 Adoption of Predictive Policing and Face Recognition Technologies

These systems sort through enormous amounts of social and biometric information to predict crime or identify suspects. As much as they can make law enforcement more efficient, they can potentially enhance prejudice, limit freedom of speech, and promote suspect profiling based on sensitive characteristics. India's National AFRS has been protested because it was implemented without

a functioning data protection system or an independent oversight mechanism. This scenario raises serious constitutional concerns regarding surveillance without protection and the coercive state power delegation to unaccountable algorithmic decision-making.

5.1 Proportionality Test and Due Process in Criminal Justice

The test of proportionality, re-stated in Puttaswamy and elaborated in *"Modern Dental College v. State of Madhya Pradesh"*, is the basis for examining limitations imposed on basic rights. The test requires an understanding that any encroachment upon basic rights must (*Modern Dental College And Research Centre And Others v. State Of Madhya Pradesh And Others, Supreme Court Of India, Judgment, Law, Casemine.Com*, n.d.)

(i) be sanctioned by law.
(ii) have a legitimate purpose,
(iii) be crucial in a democratic country, and
(iv) must be proportionate to the degree of restriction imposed. The four-pronged test, therefore, must be met for the use of AI surveillance technologies. Lacking a specialised surveillance regulation regime or judicial oversight in India, the legality and constitutionality of those means, nonetheless, remain doubtful.

Procedurally, the essence of due process is undermined where algorithmic decision-making is adopted in the absence of transparency, accountability, or redress. AI software used to determine the grant or denial of bail, compute risk scores, or select suspects may have a lack of explainability and shut all avenues for people to question the accuracy or fairness of the outcome. "Such a scenario undermines the due process protections necessarily ingrained in Article 21 and the right to a fair trial as enunciated in Article 14."

5.2 Tension Between State Interest and Individual Rights

There exists an underlying conflict between the duty of the state to maintain public order and security and the liberty and privacy of the individual. While the state has the right to impose reasonable restrictions on rights in the interest of national security or public order under *"Article 19(2), these must not be excessive, vague, or prone to arbitrariness. In the modern digital era, AI technologies may also facilitate mass surveillance systems that are disproportionately biased in favour of state interests."* As the *"Court emphasised in Puttaswamy, the Constitution is founded on the principle of limited government, and unrestrained technological advancement threatens the balance between citizens and the sovereign."* In conclusion, although AI presents extensive possibilities to revitalise criminal investigation, its use needs to be balanced with constitutional principles. The right to privacy, the right to informed consent, the right to due process, and the principle of proportionality should be guiding principles for the application of these tools within the criminal justice system. As India advances further to incorporate AI into governance, prioritising the creation of legal frameworks is essential to ensure the ethical, legal, and rights-based implementation of digital forensic technologies.

6.0 Regulatory and Ethical Challenges

The growing role of AI across many sectors has brought a transformative change in the criminal justice system. AI technologies are being employed to support law enforcement agencies, prosecutors, and courts in investigating, adjudicating, and anticipating criminal conduct. These technologies bring unparalleled efficiencies and capabilities, but also raise fundamental legal and ethical challenges that undermine the very building blocks of criminal law, such as the right to a fair trial, the presumption of innocence, and procedural fairness. ((*PDF*) *The Role of Artificial Intelligence in Criminal Justice*, n.d.)

Evidence based on AI generally denotes information or analytical results generated by algorithms, machine learning models, or automated systems, presented in court proceedings for establishing facts, determining culpability, or aiding in sentencing. In legal systems such as those of the United States and the European Union, evidence generated by AI has been admitted in numerous criminal proceedings, prompting worries about its reliability, transparency, and potential biases.

There is a growing demand in India for the use of technology to support criminal investigations, with the government attempting to promote AI-based policing and surveillance systems. However, the Indian law has not yet evolved clear directions that control the admissibility and analysis of evidence from AI.

One major legal concern relates to the reliability and verifiability of AI outputs, as traditional rules of evidence

demand rigorous scrutiny regarding authenticity, relevance, and probative value. AI systems, particularly those employing deep learning or neural networks, often function as "black boxes, ('(PDF) The Black Box Problem Revisited. Real and Imaginary Challenges for Automated Legal Decision Making', 2024)" making it difficult for defendants to contest AI-based evidence.

As India and other countries are also struggling with these changing realities, it is an urgent necessity to establish parameters on the acceptability, evaluation, and challenge ability of artificial intelligence evidence in criminal trials. ('The Rise of Big Data Policing', n.d.-b)

6.1 Understanding AI Evidence

AI evidence is data or results produced by AI systems used in court proceedings to prove or disprove facts of a case. In the criminal justice system, AI evidence can be in various forms, such as algorithmic risk scores and forensic analysis. What characterises AI evidence is that it is not directly the result of human observation or interpretation but a result of advanced computational processes that scan extensive amounts of information. Typical types of AI evidence used in criminal trials are facial recognition outputs, predictive policing data, digital forensics generated by automated systems, and biometric identification systems. (Grimm *et al.*, n.d.)

AI-based technologies are increasingly being deployed for voice, fingerprint, and DNA recording analysis with the promise of greater accuracy and faster processing compared to traditional methods. However, such reliance on AI-based evidence introduces new challenges, primarily the lack of transparency of AI systems, particularly those based on deep learning algorithms. Such non-transparency is a major hurdle for judicial systems in assessing the authenticity and credibility of the produced evidence. In addition, AI-based evidence typically comes from large databases that were conventionally collected, and these might contain inherent prejudices or be incomplete, hence producing outcomes that can perpetuate systemic prejudices.

There is a need for legal practitioners, judges, and legislators to appreciate the nature and scope of AI-generated evidence. If there is a lack of appreciation of the production processes of AI output and the assumptions underlying them, then there is a danger that this evidence will be accepted at face value, which may lead to miscarriages of justice.

6.2 Algorithmic Opacity, Bias, and the Ethics of Fairness

Algorithmic opacity is an ethical issue in algorithmic justice, where the decision-making processes of AI systems are not accessible to third parties. This opacity is common in proprietary software, where private vendors maintain intellectual property rights on source codes, training data, and weighting schemes. This transparency is unhealthy as it contradicts the transparency and due process requirements of criminal adjudication. Algorithmic bias, which arises when historical data is used to train models, often reflects societal bias. Empirical studies show that facial-recognition systems are less accurate for people with darker skin tones and women. ProPublica's study found that the probability of misclassifying a black defendant as high-risk was nearly two times higher than white defendants when applying the COMPAS algorithm. (Mattu, n.d.)

Ethical concerns of distributive justice and fair treatment are highlighted, as there is no recourse to tackling erroneous algorithmic responses. Deontological ethics, based on Immanuel Kant's moral philosophy, emphasises the inherent worth of individuals and the ethical imperative to treat others as ends in themselves. Incorrectly allowing an algorithm to label an individual as a potential threat without their consent or contribution violates their dignity and autonomy.

6.3 Human Against Machine Moral Agency in Criminal Justice

AI lacks understanding, intention, and ethical considerations. The absence of moral agency is especially noticeable when such systems are tasked with enforcing fundamental human rights. In the standard paradigms of justice, accountability must be based on human judgment, ethical consideration, and the ability to correct. Judges, police officers, and forensic scientists are supposed to adhere to ethical norms and legal protections. When AI systems replace or usurp the roles of these persons, the accountability structure is disrupted, leaving victims of algorithmic failure with no obvious party to blame or against whom to seek restitution. In *"State v. Loomis (2016), the Wisconsin Supreme Court acknowledged that the COMPAS algorithm used in the defendant's sentencing evaluation lacked complete transparency and individualisation."* While the Court supported using

them, it underlined that such tools should not constitute the only basis for sentencing conclusions and underlined their drawbacks. Ethically, this represents a rising judicial recognition of the constraints of assigning moral power to non-human creatures. The Court was able to underline the need for human oversight, although in an environment marked by a lack of comprehensive regulatory definition. It is also connected to consequentialism, a moral theory that determines the rightness of an action based on its effects. Advocates of AI believe that technology may lower human weakness, accelerate justice delivery, and maximise resource efficiency. However, if the implications of utilising AI lead to wrongful arrests, self-control as an outcome of excessive monitoring, or incorrect sentencing decisions, especially against disadvantaged groups, ethical considerations take precedence over administrative efficiency. The rule of utility, which is frequently invoked in public policy debate, must be balanced by constitutional ethics in matters of rights.

6.4 Human Dignity as a Fundamental Ethical and Constitutional Pillar

Human dignity is a high moral and constitutional value or principle, the pillar of privacy and autonomy. It states that everybody has the right to be respected, be recognised and to have control of their lives. In the context of AI systems that operate opaquely, transparency is not a requisite. The absence of consent, along with a lack of substantive interpretability, diminishes their capacity to uphold principles of respect. Ethically, any system which defines personal freedom should recognise that individual should be in liberty to express their opinions, as far as the ones which concur with the reasoning of decisions, and challenge the results. This duty is in line with the principle of Amartya Sen that suggests that the foundation of justice is something that he terms as public reasoning. The irrevocable judgments created by AI systems based on factors that cannot be perceived are not justified in any way; hence, they do not deserve to be considered from a moral and legal perspective. The Indian Supreme Court, in several judgments, has highlighted the protection of human dignity in criminal law, and it points out that the criminal justice machinery shall not act in violation of the cognitive liberty or act without conscious assent. It specifically applies to the AI-powered behavioural profiling or predictive policing, which intrudes into the cognitive and associative space of persons without appropriate respect to dignity and informed consent.

6.5 Lack of Specific Legislation on AI in Forensics

India presently lacks an adapted legal framework for the use of AI in criminal investigations and digital forensics. Existing laws under the *"Indian Evidence Act, 1872, and the Information Technology Act, 2000"*, provide minimal advice on the admissibility and trustworthiness of AI-generated evidence. These statutes were not created with autonomous or probabilistic systems in mind, and as a result, they fail to meet the distinct epistemological issues given by machine learning models, such as opacity, data reliance, and systematic bias. This legal vacuum has resulted in a patchwork approach in which security entities use AI capabilities with little supervision or judicial review.

7.0 Comparative Perspectives: India and the United States

The Indian and American due process-oriented constitutional and judicial systems show considerable divergence in their response to the due process concerns raised by AI in the criminal justice process. Both countries encounter challenges in the use of algorithmic tools in law enforcement, bail determination, and punishment. The differing constitutional systems of due process, transparency, and justice offer valuable lessons in the composition of the issues at hand as well as evolving standards within the framework.

The United States Constitution's Fourteenth Amendment and Fifth Amendment provide for the doctrine of due process of law, mandating that no individual shall be deprived of life, liberty, or property without due legal process. As AI took hold, U.S. courts began to struggle with due process implications of black box algorithmic systems that impact judicial decisions, specifically bail and sentencing. In *"State v. Loomis, the Wisconsin Supreme Court balanced the constitutionality of the use of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) algorithm in sentencing."* The defendant argued that the use of a proprietary, opaque risk assessment tool deprived him of his due process rights by not allowing him to view

or challenge its underlying methodology. The Court ultimately sanctioned the use of COMPAS, holding that it should be viewed as one of several factors in sentencing, but also warned against the use of COMPAS as the sole determinant of judicial discretion. Nevertheless, it also acknowledged limitations in the tool, citing issues with racial bias and the lack of individualised assessments.

Indian case law of the Constitution does not employ the term "due process" but preserves the substantive and procedural protections of fairness under *"Articles 14, 19, and 21. As reconfirmed in Maneka Gandhi v. Union of India (1978), the Indian Supreme Court expanded the meaning of Article 21 to encompass "fair, just and reasonable" procedure, thus giving due process a constitutional basis in Indian law. This has broad implications for the application of AI in Indian criminal justice: any use of algorithmic tools must pass the test of non-arbitrariness, reasonableness, and procedural fairness, particularly if such tools impinge upon life or liberty."*(*Maneka Gandhi vs Union of India on 25 January, 1978*, n.d.)

Indian courts, on the other hand, have not yet confronted questions of fairness directly in the context of algorithmic justice. However, questions of transparency and bias have arisen in adjacent spheres. For instance, in *"Internet Freedom Foundation v. Ministry of Home Affairs (2021)"*, a court challenge was instituted against the use of FRT for law enforcement in the absence of a legislative framework. Even though no final order has been passed, the challenge is based on parameters of arbitrariness, violation of privacy, and discriminatory impact issues directly relevant to the question of fairness. India may draw lessons from this course. Firstly, there is a need to prevent ad-hoc or extra-legal use of AI tools by police or judicial authorities, particularly in the absence of data protection law or institutional ethics oversight. As seen in surveillance cases, e.g., *"PUCL v. Union of India (1997)"*, the Supreme Court has subjected intrusive state action to procedural protection and judicial review. Such precepts must be applied to algorithmic systems with predictive or surveillance functionalities. Secondly, India needs to put in place statutory mechanisms to regulate the use of AI in law enforcement. The comparative perspective shows that legal demands of explainability, contestability, and equal treatment are required to guarantee algorithmic fairness and procedural justice.

8.0 Policy Recommendations and Conclusion

The use of AI within the criminal justice framework is becoming more prevalent, and there are concerns regarding transparency, accountability, bias, and the potential infringement of fundamental rights. The Indian legal framework must take positive measures towards regulating algorithmic systems prior to their use as entrenched and unamendable components. This manuscript provides accurate policy suggestions founded upon constitutional principles, comparative jurisprudence, and ethical theory.

The recommendations are to enact a stand-alone AI criminal justice regulation statute, create an independent AI watchdog agency, impose normative limits on the admissibility and employment of algorithmic tools, incorporate ethical theories into system design, ensure quality, representativeness, and consent of data, and implement international best practices.

India must take the best from around the world, but with context, embracing a risk-tiered approach to AI use and actively engaging in global collaboration to co-create ethical and legal frameworks for AI in the justice system.

The use of AI in the criminal justice system is a watershed in the administration, oversight, and experience of justice. Although technology holds out the promise of efficiency and consistency, it also poses deep constitutional and ethical concerns. State v. Loomis (2016). Comparative experience is a warning against the use of AI absent sufficient procedural protection.

Finally, the AI application in the courts of law marks a turning point in administration, regulation, and the concept of justice itself. India must ensure that technological integration extends beyond information processing to protect rights during documentation, so that the future criminal justice system is founded on constitutional principles and ethical values.

9.0 References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias-There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*, Online Edition. - References-Scientific Research Publishing. (n.d.). Retrieved 23 July 2025, from https://www.scirp.org/reference/referencespapers?referenceid=3337609

Artificial intelligence and policing' by Elizabeth E. Joh. (n.d.). Retrieved 8 June 2025, from https://digitalcommons.law.seattleu.edu/sulr/vol41/iss4/7/

Artificial intelligence (AI) vs. machine learning. (n.d.). CU-CAI. Retrieved 22 July 2025, from https://ai.engineering.columbia.edu/ai-vs-machine-learning/

Artificial intelligence in criminal justice: Balancing efficiency with fairness and accountability. (n.d.).

Bhatia, G. (2017). *The Supreme Court's right to privacy judgment. economic and political weekly, 52*(44), 22-25.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77-91.

Delhi Police's claims that FRT is accurate at 80% are 100% scary. (2022, August 17). *Internet Freedom Foundation*

E. P. Royappa *vs.* State of Tamil Nadu & Anr on 23 November 1973. (n.d.). Retrieved 31 October 2024, from https://indiankanoon.org/doc/1327287/

Grimm, P. W., Grossman, M. R., & Cormack, G. V. (n.d.). *Artificial Intelligence as Evidence*.

Linden, T., Khandelwal, R., Harkous, H., & Fawaz, K. (2020). The privacy policy landscape after the GDPR. *Proceedings on Privacy Enhancing Technologies, 2020*(1), 47-64. https://doi.org/10.2478/popets-2020-0004

Lum, K., & Isaac, W. (2016). *To predict and serve? Significance (Oxford, England), 13*(5), 14-19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Maneka Gandhi vs Union of India on 25 January 1978. (n.d.). Retrieved 16 July 2024, from https://indiankanoon.org/doc/1766147/

Mattu, J. L., Julia Angwin, Lauren Kirchner, Surya. (n.d.). How we analyzed the COMPAS recidivism algorithm. *ProPublica*. Retrieved 31 January 2025, from https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Modern Dental College and Research Centre and Others v. State of Madhya Pradesh And Others, Supreme Court of India, Judgment, Law, casemine.com. (n.d.). Https://Www.Casemine.Com. Retrieved 8 June 2025, from https://www.casemine.com/judgement/in/5790b4f3e561097e45a4e666

(PDF) Recalibrating the right to privacy in the digital age: Legal challenges and constitutional safeguards in the era of technological intrusion. (2025). *Research Gate*. https://doi.org/10.29121/shodhkosh.v5.i6.2024.5121

(PDF) The black box problem revisited. Real and imaginary challenges for automated legal decision making. (2024). ResearchGate. https://doi.org/10.1007/s10506-023-09356-9 https://doi.org/10.1007/s10506-023-09356-9

(PDF) The Role of Artificial Intelligence in Criminal Justice. (n.d.). Retrieved 10 June 2025, from https://www.researchgate.net/publication/384910406_The_Role_of_Artificial_Intelligence_in_Criminal_Justice

Singh, V. (2020). *1,100 rioters identified using facial recognition technology: Amit Shah*. The Hindu. https://www.thehindu.com/news/cities/Delhi/1100-rioters-identified-using-facial-recognition-technology-amit-shah/article31044548.ece

Sinha, A. (2024). *The landscape of facial recognition technologies in India* | TechPolicy.Press. Tech Policy Press. https://techpolicy.press/the-landscape-of-facial-recognition-technologies-in-india

State v. Loomis, (2016) Wisconsin Supreme Court Decisions: Wisconsin Case Law: Wisconsin Law: U.S. Law: Justia. (n.d.). Retrieved 23 July 2025, from https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html

The rise of big data policing. (n.d.-a). NYU Press. Retrieved 8 June 2025, from https://nyupress.org/9781479892822/the-rise-of-big-data-policing/

The rise of big data policing. (n.d.-b). NYU Press. Retrieved 22 July 2025, from https://nyupress.org/9781479892822/the-rise-of-big-data-policing/

Warthon, M. (2024, September 30). *Restricting access to AI decision-making in the public interest: The justificatory role of proportionality and its balancing factors* [Info:eu-repo/semantics/article]. Alexander von Humboldt Institute for Internet and Society gGmbH. https://doi.org/10.14763/2024.3.1801