Page 414 - 423 **DOI:** *https://doij.org/10.10000/IJLSI.112545*

INTERNATIONAL JOURNAL OF LEGAL SCIENCE AND INNOVATION

[ISSN 2581-9453]

Volume 7 | Issue 2

2025

The Use of Artificial Intelligence in Criminal Trials

SHRUTHIKA S.[1]

ABSTRACT

The integration of Artificial Intelligence (AI) into criminal justice systems marks a transformative development with profound implications for the conduct of criminal trials. AI tools are increasingly employed for a variety of purposes, including evidence analysis, predictive policing, risk assessment in bail and sentencing decisions, and even aiding judicial deliberations. These technologies promise greater efficiency, objectivity, and consistency in criminal proceedings. However, the use of AI in criminal trials also raises serious concerns about fairness, transparency, accountability, and the preservation of fundamental rights. The opaque nature of many AI systems, often referred to as the "black box" problem, challenges the principle of open justice and hampers defendants' ability to understand, challenge, or appeal AI-driven decisions. Moreover, biases embedded in training datasets risk perpetuating or even amplifying existing societal prejudices, thereby undermining the fairness of trials.

This paper critically examines the potential and pitfalls of using AI in criminal trials. It explores the legal and ethical issues involved, particularly with respect to the right to a fair trial under constitutional and international human rights norms. Comparative experiences from jurisdictions such as the United States, the United Kingdom, and India are analyzed to understand emerging regulatory trends and judicial responses. The paper argues that while AI can be a valuable tool in enhancing the criminal justice process, its use must be strictly governed by principles of due process, transparency, explainability, and accountability. Safeguards such as mandatory disclosures about AI use, opportunities for human review, and mechanisms for contesting AI-generated findings are essential to preserve the integrity of criminal trials. Ultimately, the responsible integration of AI into criminal justice demands a careful balancing of technological innovation with unwavering adherence to the rule of law and the protection of individual rights.

Keywords*: artificial intelligence, evidence, criminal law, accountability.*

I. INTRODUCTION

The growing influence of artificial intelligence (AI) across various sectors has not spared the criminal justice system. AI technologies are increasingly being used to assist law enforcement

[1] Author is an Assistant Professor in India.

 [ISSN 2581-9453]

agencies, prosecutors, and even courts in investigating, adjudicating, and predicting criminal behavior. From facial recognition software and predictive policing tools to AI-driven forensic analysis, the advent of AI-generated evidence marks a significant shift in how justice is administered today. While these technologies offer unprecedented efficiencies and capabilities, they also pose profound legal and ethical dilemmas that challenge the foundational principles of criminal law, including the right to a fair trial, the presumption of innocence, and procedural fairness.

AI evidence typically refers to data or analytical outputs produced by algorithms, machine learning models, or automated systems, which are presented in court to establish facts, assess guilt, or assist in sentencing. In jurisdictions like the United States and the European Union, AI evidence has already been admitted in several criminal trials, raising concerns about its reliability, transparency, and bias[2]. In India too, there is a growing appetite for integrating technology into criminal investigations, with government initiatives promoting AI-based policing and surveillance systems[3]. However, the Indian legal system has yet to formulate comprehensive standards governing the admissibility and scrutiny of AI-generated evidence.

One major legal concern relates to the reliability and verifiability of AI outputs. Traditional rules of evidence demand that any material relied upon in a criminal trial must be subject to rigorous scrutiny regarding its authenticity, relevance, and probative value. AI systems, particularly those employing deep learning or neural networks, often function as "black boxes," meaning that their internal decision-making processes are opaque even to their own developers[4]. This opacity creates serious challenges for defendants who seek to contest AI-based evidence, potentially undermining their constitutional right to confront the evidence against them.

Moreover, AI systems are not immune to biases. Studies have shown that algorithms trained on historical policing data may replicate or even exacerbate existing racial, socio-economic, or gender-based biases[5]. This raises ethical concerns about whether the use of such evidence can truly lead to just outcomes, or whether it perpetuates systemic discrimination under a technological guise. There is also the broader issue of accountability: when an AI system errs, it is unclear whether responsibility lies with the developers, the deploying agency, or the courts

[2] Andrea Roth, Machine Testimony, 126 Yale L.J. 1972, 1976–77 (2017).
[3] Ministry of Home Affairs, Government of India, National Strategy for Artificial Intelligence (2023), available at https://www.mha.gov.in.
[4] Jenna Burrell, How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms, 3 Big Data & Society 1 (2016).
[5] Rashida Richardson, Jason M. Schultz & Kate Crawford, Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice, 94 N.Y.U. L. Rev. Online 15 (2019).

that rely upon such outputs.

The Bharatiya Sakshya Adhiniyam, 2023-the new Indian law replacing the Indian Evidence Act, 1872 acknowledges the role of electronic records but does not specifically address AI-generated evidence, leaving a significant grey area[6]. As India and other nations grapple with these emerging challenges, there is an urgent need to establish clear guidelines on the admissibility, evaluation, and contestability of AI evidence in criminal trials.

Against this backdrop, this article critically examines the legal and ethical challenges posed by the use of AI evidence in criminal trials. It aims to evaluate existing frameworks, highlight comparative international approaches, and suggest reforms to ensure that the adoption of technology enhances rather than undermines the pursuit of justice.

II. UNDERSTANDING AI EVIDENCE

Artificial Intelligence (AI) evidence refers to information or outputs generated by AI systems that are presented in legal proceedings to establish or contest facts relevant to a case. In the criminal justice system, such evidence can take multiple forms, ranging from algorithmic risk assessments used during bail hearings to forensic analysis generated by machine learning models[7]. The distinguishing feature of AI evidence is that it is not the product of direct human observation or interpretation, but rather the output of complex computational processes that analyze vast datasets.

The types of AI evidence commonly encountered in criminal trials include facial recognition outputs, predictive policing data, digital forensics from automated tools, and biometric identification systems[8]. For instance, facial recognition software may identify a suspect from CCTV footage, while a predictive policing algorithm might generate data suggesting a higher probability of crime occurrence in certain areas, influencing investigative decisions[9]. Additionally, AI-driven tools are increasingly used to analyze voice samples, fingerprints, and DNA evidence, with claims of higher accuracy and efficiency compared to traditional methods.

However, reliance on AI evidence introduces novel complexities. One such complexity is the "black box" nature of many AI systems, particularly those based on deep learning. Unlike traditional forensic experts, whose methodologies can be examined and cross-examined, AI models often operate through layers of computation that are not easily explainable to humans,

[6] Bharatiya Sakshya Adhiniyam, 2023, No. 46, Acts of Parliament, 2023 (India).
[7] W. Nicholson Price II, Black-Box Medicine, 28 Harv. J.L. & Tech. 419, 423 (2015)
[8] Elizabeth E. Joh, Artificial Intelligence and Policing: First Questions, 41 Seattle U. L. Rev. 1139, 1142 (2018).
[9] Andrew Guthrie Ferguson, The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement 62–65 (2017).

even to those who designed them[10]. This lack of transparency creates hurdles for courts in assessing the reliability and credibility of the evidence.

Moreover, AI evidence often draws from large, historically accumulated datasets that may themselves be biased or incomplete. As a result, the outputs generated by these systems may inherit and perpetuate systemic biases[11]. For example, if historical policing data reflect over-policing in marginalized communities, an AI system trained on such data might unfairly flag individuals from those communities as higher risk, leading to potential violations of the principle of equality before the law.

Understanding the nature and limitations of AI evidence is crucial for legal practitioners, judges, and policymakers. Without a clear grasp of how AI-generated outputs are produced and the assumptions underlying them, there is a risk that such evidence will be accepted at face value, potentially leading to miscarriages of justice. It is, therefore, imperative that criminal justice systems develop robust frameworks to scrutinize AI evidence with the same critical rigor applied to traditional forms of proof.

III. LEGAL CHALLENGES IN ADMITTING AI EVIDENCE

The integration of AI evidence into criminal trials presents numerous legal challenges, particularly concerning its admissibility, reliability, and conformity with constitutional safeguards. Courts traditionally require that evidence must be relevant, authentic, and sufficiently reliable to be admitted. AI-generated outputs complicate this framework due to their technical opacity, susceptibility to bias, and questions surrounding their probative value.

One of the foremost challenges lies in establishing the **reliability** of AI evidence. The test of reliability typically demands that the method by which evidence is produced must be scientifically valid and widely accepted[12]. However, many AI algorithms, particularly those based on machine learning, evolve dynamically and produce results that cannot be easily explained or replicated. This "black box" phenomenon makes it difficult for courts to scrutinize the internal logic of the system, raising serious doubts about the verifiability of the evidence[13]. Another significant concern is the **authenticity** of AI evidence. For instance, AI-based facial recognition matches or predictive analytics may be manipulated, improperly calibrated, or

[10] Sandra Wachter, Brent Mittelstadt & Chris Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, 41 Computer L. & Security Rev. 105567 (2021).
[11] Kristian Lum & William Isaac, To Predict and Serve?, 13 Significance 14, 15 (2016).
[12] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 593–94 (1993).
[13] Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147, 1167 (2017).

compromised by poor-quality data inputs[14]. Establishing a clear **chain of custody**,that is, the preservation of data integrity from collection to presentation is far more complex when algorithms automatically process and interpret data at multiple stages without direct human oversight[15].

Moreover, AI evidence challenges the accused's **right to a fair trial**, particularly the right to confront and cross-examine adverse evidence. Under traditional evidentiary rules, a defendant must have the opportunity to question the source and reliability of evidence presented against them. However, when evidence is generated by proprietary algorithms protected under trade secret laws, defendants may be denied access to the underlying source code, limiting their ability to challenge its validity[16]. This tension between intellectual property protection and the demands of due process has sparked widespread debate.

In addition, there is the issue of **algorithmic bias**. Courts must consider whether an AI system's output has been influenced by discriminatory patterns embedded in the training data. Uncritical acceptance of such biased outputs could violate the constitutional guarantee of equality before the law and non-discrimination[17]. Given these challenges, there is an urgent need for developing judicial standards tailored specifically to AI evidence, including mandates for algorithmic transparency, independent validation, and the right to meaningful adversarial testing. Without such safeguards, the risk remains that AI evidence may unduly influence judicial outcomes, compromising the integrity of criminal proceedings.

IV. ETHICAL CHALLENGES IN USING AI EVIDENCE

The use of AI evidence in criminal trials not only raises legal hurdles but also introduces deep ethical concerns that cut to the heart of justice and fairness. Among the most pressing issues are the potential violations of the accused's rights, the erosion of privacy, and the question of accountability when automated systems make or influence critical decisions.

One major ethical concern is the risk to the **right to a fair trial**. Criminal trials are fundamentally adversarial, based on the principle that evidence must be openly examined and contested. When AI evidence is opaque or its methodologies are inaccessible due to corporate confidentiality or technological complexity, defendants may be unable to effectively challenge

[14] Clare Garvie, Garbage In, Garbage Out: Face Recognition on Flawed Data, Georgetown Law Center on Privacy & Technology (May 16, 2019), available at https://www.flawedfacedata.com.
[15] Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249, 1270–72 (2008).
[16] Rebecca Wexler, Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System, 70 Stan. L. Rev. 1343, 1351–52 (2018).
[17] Pauline T. Kim, Data-Driven Discrimination at Work, 58 Wm. & Mary L. Rev. 857, 863 (2017).

the evidence against them[18].This undermines the equality of arms, a cornerstone of due process recognized under both domestic constitutional frameworks and international human rights instruments[19]. Another critical issue is **privacy**. AI technologies, especially those used in surveillance and data analytics, often involve the mass collection, storage, and analysis of personal information without the informed consent of individuals[20].

Predictive policing programs, for instance, may track individuals based on location data, social media activity, or prior associations, raising serious concerns about unwarranted intrusion into private life and the chilling effect on freedoms of expression and association[21].The problem of **algorithmic discrimination** further compounds these ethical challenges. If AI systems are trained on biased or unrepresentative data, their outputs can perpetuate and even magnify existing inequalities, leading to disproportionate targeting or harsher treatment of marginalized groups[22].Such outcomes not only offend ethical norms of fairness and equality but also threaten the legitimacy of the criminal justice system itself.

Additionally, there is the issue of **accountability**. Traditional notions of criminal responsibility are premised on human agency and intent. When AI systems contribute to wrongful arrests or convictions, it becomes difficult to assign blame or seek redress, given the distributed responsibility among programmers, vendors, and law enforcement agencies[23].This "accountability gap" challenges both ethical and legal doctrines that demand clear attribution of liability for wrongful acts.

Given these profound concerns, it is ethically imperative to ensure that the deployment of AI evidence in criminal trials is governed by principles of transparency, fairness, accountability, and respect for fundamental rights. Otherwise, the reliance on AI risks eroding public trust in the justice system and entrenching systemic injustices under the guise of technological objectivity.

V. COMPARATIVE PERSPECTIVES

Different jurisdictions have approached the challenges posed by AI evidence in criminal trials with varying degrees of caution and regulation. Comparative analysis reveals both innovative practices and persistent gaps in addressing the complex legal and ethical issues arising from

[18] Aziz Z. Huq, A Right to a Human Decision, 106 Va. L. Rev. 611, 624–25 (2020)
[19] International Covenant on Civil and Political Rights art. 14(1), Dec. 16, 1966, 999 U.N.T.S. 171.
[20] Shoshana Zuboff, The Age of Surveillance Capitalism 94–97 (2019).
[21] Sarah Brayne, Big Data Surveillance: The Case of Policing, 82 Am. Soc. Rev. 977, 981–83 (2017).
[22] Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor 5–7 (2018).
[23] Mark Coeckelbergh, *AI Ethics* 52–55 (The MIT Press 2020).

AI's use in the justice system.

In the **United States**, courts generally assess scientific and technical evidence under the standards set by *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, which emphasize reliability and relevance[24].However, when it comes to AI tools, particularly proprietary algorithms like COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) used in sentencing, courts have struggled with the tension between due process rights and the protection of trade secrets[25].In *State v. Loomis*, the Wisconsin Supreme Court allowed the use of COMPAS risk assessments but cautioned that judges must not base sentencing decisions solely on such algorithmic evaluations[26].This illustrates an emerging but cautious acceptance of AI, tempered by due process concerns.

In the **European Union**, the regulatory framework is more robust. The General Data Protection Regulation (GDPR) specifically grants individuals the right not to be subject to decisions based solely on automated processing[27]. Additionally, the proposed Artificial Intelligence Act seeks to classify AI systems used in law enforcement and judiciary as "high-risk," subjecting them to strict compliance obligations, including transparency, human oversight, and accountability measures[28].European legal systems, with their greater emphasis on data protection and human rights, offer valuable lessons on the necessity of regulating AI evidence proactively.

In contrast, **India** has yet to develop a specialized framework for AI in the criminal justice process. Although the Bharatiya Sakshya Adhiniyam, 2023, acknowledges the role of electronic records, it does not specifically address AI-generated outputs[29].Law enforcement agencies are rapidly adopting AI technologies, such as facial recognition systems like the Automated Facial Recognition System (AFRS), without comprehensive legislative safeguards[30].This regulatory vacuum raises significant risks to due process and equality rights under Articles 14 and 21 of the Indian Constitution.

Comparative experiences underscore the need for a cautious, rights-based approach to AI evidence. Jurisdictions like the EU provide a model for embedding ethical safeguards into legal

[24] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993).
[25] Rebecca Wexler, Code of Silence: How Private Companies Hide Flaws in the Criminal Justice System, 128 Yale L.J. 1238, 1244–45 (2019).
[26] State v. Loomis, 881 N.W.2d 749, 761–62 (Wis. 2016).
[27] Regulation (EU) 2016/679 of the European Parliament and of the Council, art. 22, 2016 O.J. (L 119) 1 (General Data Protection Regulation).
[28] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence, COM (2021) 206 final (Apr. 21, 2021).
[29] Bharatiya Sakshya Adhiniyam, 2023, No. 46, Acts of Parliament, 2023 (India).
[30] Internet Freedom Foundation, Project Panoptic: Automated Facial Recognition Systems in India, available at https://panoptic.in.

frameworks, while the U.S. experience highlights the pitfalls of balancing innovation with fundamental rights. For India and other developing legal systems, these comparative insights are crucial in designing appropriate standards that ensure AI serves justice without undermining it.

VI. THE INDIAN CONTEXT: EMERGING ISSUE AND GAPS

India's rapid digitization of its criminal justice system has led to an increasing reliance on AI technologies, but the legal and regulatory frameworks remain underdeveloped. Despite the deployment of advanced tools like Automated Facial Recognition Systems (AFRS), predictive policing algorithms, and AI-assisted forensic analysis, there is a notable absence of specific legal standards governing the admissibility, reliability, and ethical use of AI-generated evidence[31].One major gap is the **lack of statutory guidance** on the admissibility of AI evidence. While the Bharatiya Sakshya Adhiniyam, 2023, recognizes electronic records as admissible, it does not differentiate between static digital evidence (such as CCTV footage) and dynamic AI-generated outputs that involve algorithmic interpretation[32].Consequently, courts are left without clear criteria to assess whether AI outputs meet traditional evidentiary standards like relevancy, authenticity, and probative value.

A second pressing issue is **transparency and accountability**. Many AI systems deployed by law enforcement agencies operate without any publicly available information regarding their design, data inputs, error rates, or operational biases. This lack of transparency impedes the ability of defendants to challenge AI evidence and undermines the right to a fair trial under Article 21 of the Constitution[33]. Furthermore, the absence of independent audits or certification requirements for AI tools exacerbates concerns about wrongful convictions based on flawed or biased outputs.

Privacy violations represent another emerging challenge. Indian authorities have rolled out facial recognition and biometric systems in public spaces without comprehensive data protection legislation. The pending Digital Personal Data Protection Act, 2023, while addressing some privacy concerns, does not specifically regulate law enforcement access to or use of AI systems[34].This regulatory gap risks infringing individuals' fundamental rights under

[31] Vidushi Marda, Artificial Intelligence Policy in India: A Framework for Engaging the Limits of Data-Driven Decision-Making, 13 Indian J.L. & Tech. 1, 5–6 (2017).
[32] Rishabh Dara, Information Technology Act, 2000 and Challenges of Admissibility of Digital Evidence, 10 NUJS L. Rev. 43, 47 (2017).
[33] Chinmayi Arun, "AI and the Rule of Law in India," 17 *International Journal of Constitutional Law* 870, 872–74 (2019).
[34] The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).

Article 14 (equality) and Article 21 (privacy and personal liberty), as recognized in *Justice K.S. Puttaswamy v. Union of India*[35]. Finally, the Indian judiciary has yet to evolve clear **judicial standards** for evaluating AI evidence. Unlike the Daubert standard in the U.S. or the emerging risk-based frameworks in the EU, Indian courts typically rely on general principles of evidence law, which may not be equipped to deal with the technical intricacies of AI[36].As a result, there is an urgent need for judicial training, expert panels, and specialized procedural rules to responsibly manage AI-based evidence.

Unless these gaps are addressed, the introduction of AI into India's criminal justice process risks entrenching systemic biases, violating due process guarantees, and undermining public trust in judicial outcomes.

VII. CONCLUSION AND RECOMMENDATIONS

The integration of AI into criminal trials presents both unprecedented opportunities and profound challenges. AI evidence has the potential to enhance efficiency, uncover hidden patterns, and support judicial decision-making. However, its use also raises significant legal, ethical, and procedural concerns — particularly regarding transparency, accountability, fairness, and the protection of fundamental rights. Without appropriate safeguards, reliance on AI could exacerbate existing biases, compromise the right to a fair trial, and undermine public trust in the justice system.

Comparative experiences from jurisdictions such as the United States and the European Union reveal the necessity of a cautious and rights-based approach. India, while embracing technological innovation, must address its current regulatory and institutional gaps to ensure that the deployment of AI in criminal justice serves the ends of justice and not expediency alone.

Key Recommendations include:

- **Establish Clear Admissibility Standards**: AI-generated evidence must meet scientifically validated standards before being admitted in court.
- **Mandate Transparency and Explainability**: Developers and users of AI in criminal justice must be required to disclose system methodologies, data inputs, error rates, and biases.

[35] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (India)
[36] R. Venkata Rao & E. Ranganathan, Artificial Intelligence and the Indian Judiciary: Challenges and Opportunities, 62 JILI 178, 183 (2020).

- **Strengthen Privacy Protections**: AI use must align with stringent data protection laws, ensuring that individuals' rights to privacy and dignity are upheld.
- **Promote Judicial Capacity Building**: Judges, prosecutors, and defense attorneys must receive training to understand and critically assess AI evidence.
- **Ensure Human Oversight and Accountability**: Automated decision-making must never replace human judgment, and mechanisms must be in place to hold human actors accountable for AI-influenced outcomes.

By proactively crafting a clear legal and ethical framework, India can harness the benefits of AI in criminal trials while safeguarding the constitutional promises of due process, equality, and justice.

*****