ADMISSIBILITY AND RELIABILITY OF AI-GENERATED EVIDENCE IN INDIAN CRIMINAL TRIALS: CHALLENGES TO PROOF, AUTHENTICITY AND LEGAL LIABILITY

Bharti Kataria[1]
Prof. (Dr.) Ashwani Kumar Dwivedi[2]

ABSTRACT

The growing integration of Artificial Intelligence (AI) into criminal investigations has introduced a new category of evidence generated through algorithmic processes, including facial recognition systems, predictive analytics, and automated forensic tools. While such AI-generated evidence enhances efficiency and investigative capabilities, it raises significant concerns regarding its admissibility and reliability in criminal trials. In India, the existing evidentiary framework under the Bharatiya Sakshya Adhiniyam, 2023, primarily addresses electronic and digital records but does not specifically account for the complexities associated with AI-generated outputs.

This paper critically examines the challenges posed by the use of AI-generated evidence in Indian criminal proceedings, particularly with respect to proof, authenticity, and legal liability. It highlights issues such as algorithmic opacity, bias in training data, lack of transparency, and the risk of manipulation, including the use of deepfakes. The study further explores the implications of relying on such evidence for the right to a fair trial and the presumption of innocence.

By analysing statutory provisions, judicial approaches, and comparative international practices, the paper argues that the current legal framework is insufficient to address the unique challenges posed by AI. It emphasises the need for clear guidelines on admissibility, enhanced standards for reliability, and a well-defined framework for fixing liability. The paper concludes by recommending legal and policy reforms to ensure that the use of AI in criminal trials aligns with principles of fairness, accountability, and justice.

Keywords: Artificial Intelligence (AI), AI-Generated Evidence, Criminal Trials, Admissibility of Evidence, Reliability, Bharatiya Sakshya Adhiniyam, 2023, Digital Evidence, Algorithmic Bias, Deepfakes, Legal Liability, Fair Trial.

[1] LL.M. Scholar, Faculty of Law, SGT University, Gurugram, Haryana, India.
[2] Professor, Faculty of Law, SGT University, Gurugram, Haryana, India.

INTRODUCTION

The increasing integration of Artificial Intelligence (AI) into the criminal justice system has significantly transformed the processes of investigation, evidence collection, and adjudication. Technologies such as facial recognition systems, predictive policing tools, automated forensic analysis, and voice recognition software are now being used by law enforcement agencies to enhance efficiency and accuracy in criminal investigations.[3] While these developments mark a progressive shift towards technologically advanced policing, they simultaneously introduce complex legal challenges, particularly in the context of criminal trials where the standard of proof is stringent and the stakes involve individual liberty and justice.

Traditionally, the law of evidence has been premised on human testimony, documentary proof, and material objects, all of which are subject to judicial scrutiny and cross-examination. However, AI-generated evidence differs fundamentally in that it is produced through algorithmic processes that may not be transparent or easily comprehensible to judges, lawyers, or the accused.[4] This "black-box" nature of AI systems raises serious concerns regarding the ability of courts to assess the reliability and authenticity of such evidence. The opacity of algorithms, coupled with the potential for bias in training data, further complicates the evidentiary evaluation process and challenges the foundational principles of fairness and due process.[5]

In the Indian context, the enactment of the Bharatiya Sakshya Adhiniyam, 2023 represents an attempt to modernise evidentiary law, particularly with respect to electronic and digital records. However, the statute does not explicitly address AI-generated evidence or the unique issues arising from its use in criminal trials.[6] This legislative gap creates uncertainty regarding the admissibility of such evidence and the standards required to establish its probative value. Moreover, the absence of clear guidelines raises concerns about the potential misuse of AI technologies, including the risk of wrongful convictions based on flawed or manipulated algorithmic outputs.

[3] Expert Systems in Law Enforcement, 21 Harv. J.L. & Tech. 45 (2018).
[4] Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147 (2017).
[5] Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist, 7 Int'l Data Privacy L. 76 (2017).
[6] The Bharatiya Sakshya Adhiniyam, 2023, §§ 61–63 (India).

Another significant issue pertains to the question of legal liability in cases where AI-generated evidence leads to erroneous outcomes. Unlike traditional evidence, where responsibility can be attributed to identifiable human actors, AI systems operate through complex networks involving developers, data providers, and end-users.[7] Determining accountability in such scenarios poses a challenge for the legal system and necessitates a re-examination of existing principles of criminal liability and state responsibility.

This paper seeks to examine the admissibility and reliability of AI-generated evidence in Indian criminal trials, with a particular focus on the challenges it poses to established notions of proof, authenticity, and legal liability. It adopts a doctrinal approach to analyse existing legal provisions, judicial interpretations, and emerging technological concerns. The paper further aims to identify gaps in the current legal framework and propose reforms to ensure that the use of AI in criminal proceedings aligns with the fundamental principles of fairness, transparency, and justice.

Understanding AI-Generated Evidence: Concept and Scope

The rapid advancement of Artificial Intelligence (AI) has led to the emergence of a new category of evidentiary material commonly referred to as AI-generated evidence. Broadly, AI-generated evidence may be understood as information or output produced by machine learning systems, algorithms, or automated processes that is subsequently used in legal proceedings to establish facts in issue.[8] Unlike traditional forms of evidence, which are created or recorded by human actors, AIgenerated evidence is derived from computational models trained on large datasets, often operating with minimal human intervention.

At a conceptual level, it is important to distinguish AI-generated evidence from conventional electronic or digital evidence. While digital evidence typically includes emails, CCTV footage, or electronic records directly created by human input, AI-generated evidence involves an additional layer of algorithmic interpretation or decision-making.[9] For instance, a raw video recording may constitute digital evidence, whereas the identification of a suspect through facial recognition software based on that video would qualify as AI-generated evidence. This distinction is crucial, as it raises questions about the reliability of the inferential processes

[7] Andrew Selbst & Solon Barocas, The Intuitive Appeal of Explainable Machines, 87 Fordham L. Rev. 1085 (2018).
[8] Harry Surden, Machine Learning and Law, 89 Wash. L. Rev. 87 (2014).
[9] Paul Grimm et al., Authenticating Digital Evidence, 69 Baylor L. Rev. 1 (2017).

employed by AI systems.

AI-generated evidence can take multiple forms within the criminal justice system. One prominent example is facial recognition technology, which is increasingly used by law enforcement agencies to identify suspects by comparing images against databases.[10] Similarly, predictive policing tools analyse historical crime data to forecast potential criminal activity, thereby influencing investigative decisions. Automated forensic tools, such as AI-based fingerprint or DNA analysis systems, also generate outputs that may be relied upon as evidence in court. Additionally, advancements in generative AI have given rise to synthetic media, including deepfakes, which can fabricate highly realistic audio, video, or images, thereby posing serious risks of evidentiary manipulation.[11]

The defining characteristic of AI-generated evidence lies in its dependence on algorithmic processes that are often opaque, complex, and not easily explainable. This "black-box" nature of AI systems creates significant challenges for legal evaluation, particularly in determining the accuracy and trustworthiness of such evidence.[12] Unlike human witnesses, AI systems cannot be cross-examined, and their internal decision-making processes may not be fully accessible to the parties or the court. Consequently, assessing the probative value of AI-generated evidence requires a deeper understanding of the underlying technology, including the quality of training data, the design of algorithms, and the potential for systemic bias.

Furthermore, the use of AI in evidence generation raises concerns regarding standardisation and uniformity. Different AI systems may produce varying results based on the datasets and methodologies employed, leading to inconsistencies that can undermine evidentiary reliability. In the absence of universally accepted standards or regulatory frameworks, the risk of erroneous or biased outputs becomes more pronounced, particularly in criminal trials where the burden of proof rests heavily on the prosecution.

Thus, AI-generated evidence represents a paradigm shift in the law of evidence, necessitating a reevaluation of traditional legal principles governing proof and admissibility. While such evidence offers significant potential to enhance the efficiency and accuracy of criminal

[10] Clare Garvie et al., The Perpetual Line-Up: Unregulated Police Face Recognition in America, Geo. L. Ctr. on Privacy & Tech. (2016).
[11] Robert Chesney & Danielle Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753 (2019).
[12] Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (2015).

investigations, its inherent complexities demand cautious and critical judicial scrutiny to ensure that its use does not compromise the fundamental rights of the accused.

Evidentiary Framework in India and its Application to AI-Generated Evidence

The legal framework governing the admissibility and evaluation of evidence in Indian criminal trials is primarily codified under the Bharatiya Sakshya Adhiniyam, 2023, which has replaced the earlier Indian Evidence Act, 1872. This transition reflects a legislative effort to modernise evidentiary rules in response to technological advancements, particularly the increasing reliance on electronic and digital records. However, despite this shift, the statutory framework does not explicitly address the complexities associated with AI-generated evidence, thereby creating interpretative challenges in its application.

The Bharatiya Sakshya Adhiniyam, 2023 recognises electronic records as a form of documentary evidence and permits their admissibility subject to procedural safeguards designed to ensure authenticity and reliability.[13] These safeguards broadly include requirements relating to proper certification, integrity of data, and the reliability of the device or system used in the generation and storage of such records. While these provisions represent a significant step towards accommodating digital evidence, they are primarily structured to deal with conventional electronic records and do not sufficiently account for the additional layer of algorithmic processing involved in AI-generated outputs.

Judicial interpretation has played a crucial role in shaping the admissibility of electronic evidence in India. In Anvar P.V. v. P.K. Basheer, the Supreme Court held that electronic records must comply with statutory requirements relating to certification and authenticity in order to be admissible.[14] This position was reaffirmed in Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, wherein the Court emphasised the mandatory nature of such procedural safeguards in preserving the integrity of electronic evidence.[15] These decisions establish a strict evidentiary threshold for electronic records; however, they do not directly engage with the unique attributes of AI-generated evidence.

The application of these principles to AI-generated evidence raises significant concerns.

[13] The Bharatiya Sakshya Adhiniyam, 2023 (India).
[14] Anvar P.V. v. P.K. Basheer, (2014) 10 S.C.C. 473 (India).
[15] Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 S.C.C. 1 (India).

Unlike traditional electronic records, AI-generated evidence is not merely a reproduction of stored data but an output derived through complex algorithmic processing. This introduces challenges in establishing authenticity, as the reliability of such evidence depends not only on the source data but also on the functioning of the underlying algorithm. The opaque nature of many AI systems— commonly referred to as the "black-box" problem—limits the ability of courts to scrutinise the reasoning process behind such outputs, thereby complicating the assessment of evidentiary value.

Further, the procedural aspects of criminal trials, including the collection and presentation of evidence, are governed by the Bharatiya Nagarik Suraksha Sanhita, 2023. While this statute provides general mechanisms for investigation and evidence gathering, it does not contain specific provisions regulating the use of AI technologies by law enforcement agencies. The absence of clear procedural safeguards for AI-based tools increases the risk of inconsistent practices and potential misuse, which may adversely affect the fairness of criminal trials.

In light of these considerations, it becomes evident that while the Bharatiya Sakshya Adhiniyam, 2023 provides a foundational framework for the admissibility of electronic evidence, it is not fully equipped to address the challenges posed by AI-generated evidence. The existing legal regime requires further refinement to incorporate standards of algorithmic transparency, accountability, and reliability, ensuring that the integration of AI into criminal trials remains consistent with the principles of natural justice and the right to a fair trial.

Admissibility of AI-Generated Evidence in Criminal Trials

The admissibility of evidence in criminal trials is governed by fundamental principles of relevance, authenticity, and reliability. Within the Indian legal framework, these principles are primarily codified under the Bharatiya Sakshya Adhiniyam, 2023, which recognises electronic records as admissible evidence subject to compliance with procedural safeguards. However, the emergence of AI-generated evidence raises critical questions as to whether existing standards of admissibility are adequate to address the complexities associated with algorithmically produced outputs.

At the threshold level, AI-generated evidence must satisfy the test of relevance, meaning that it must have a direct or indirect bearing on the facts in issue. In many instances, outputs generated through AI systems—such as facial recognition matches, predictive analytics, or

automated forensic reports—may appear highly probative. However, the apparent accuracy of such outputs does not automatically translate into legal admissibility. Courts must ensure that such evidence is not only relevant but also legally obtained and procedurally compliant.

A significant requirement for admissibility under Indian law is the establishment of authenticity. In the context of electronic evidence, the Supreme Court in Anvar P.V. v. P.K. Basheer held that compliance with statutory conditions relating to certification is mandatory.[16] This position was reaffirmed in Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, where the Court emphasised that such procedural safeguards are essential to ensure the integrity of electronic records.[17] While these principles provide a foundational framework, their application to AI-generated evidence is not straightforward. Unlike conventional electronic records, AI-generated evidence involves an additional layer of algorithmic processing, making it difficult to establish authenticity solely through certification.

Another critical issue concerns the reliability of AI-generated evidence. Traditional evidentiary standards assume that the source of evidence can be examined and tested through cross-examination. However, AI systems often operate as "black boxes," where the internal reasoning process is not transparent or easily explainable. This lack of explainability raises concerns regarding the ability of courts to meaningfully assess the accuracy and trustworthiness of such evidence. Moreover, the inability of the defence to effectively challenge the functioning of the algorithm may undermine the principles of natural justice and the right to a fair trial.

The admissibility of AI-generated evidence also raises concerns regarding procedural fairness. Criminal trials in India are premised on the principle that the prosecution must prove its case beyond reasonable doubt. The use of AI-generated outputs, particularly where their underlying logic is not disclosed, may shift the evidentiary burden in a manner that is prejudicial to the accused. This is especially problematic in cases where courts may place undue reliance on the perceived objectivity or scientific nature of AI systems, without adequately scrutinising their limitations.

Furthermore, issues of legality in the collection of AI-generated evidence must also be considered. The deployment of AI tools by law enforcement agencies, including facial

[16] Anvar P.V. v. P.K. Basheer, (2014) 10 S.C.C. 473 (India).
[17] Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 S.C.C. 1 (India).

recognition and predictive policing systems, may raise concerns relating to privacy and surveillance. In the absence of clear statutory guidelines regulating such technologies, the admissibility of evidence obtained through potentially intrusive or unconstitutional means remains questionable.

In light of these challenges, it becomes evident that the existing legal framework governing admissibility is not fully equipped to address the unique characteristics of AI-generated evidence. While principles developed in relation to electronic evidence provide a useful starting point, they require significant adaptation to account for issues such as algorithmic opacity, bias, and lack of accountability. Therefore, courts must adopt a cautious and context-sensitive approach while determining the admissibility of AI-generated evidence, ensuring that its use does not compromise the fundamental rights of the accused or the integrity of the criminal justice system.

Challenges of Reliability and Authenticity in AI Evidence

The increasing reliance on Artificial Intelligence in criminal investigations raises serious concerns regarding the reliability and authenticity of AI-generated evidence. While such evidence is often perceived as objective and scientifically accurate, its underlying processes are far from infallible. The determination of reliability and authenticity, which lies at the core of evidentiary evaluation, becomes significantly more complex when evidence is produced through algorithmic systems rather than human observation.

One of the primary challenges associated with AI-generated evidence is the issue of algorithmic opacity, commonly referred to as the "black-box" problem. Many AI systems, particularly those based on machine learning, operate through complex computational models that do not provide clear explanations for their outputs.[18] As a result, neither the courts nor the parties involved may be able to fully understand how a particular conclusion was reached. This lack of transparency undermines the ability to assess the reliability of the evidence and restricts the defence's capacity to effectively challenge it through cross-examination.

Another significant concern is algorithmic bias, which arises from the data on which AI systems are trained. If the training data contains inherent biases—whether based on race, gender, socioeconomic status, or other factors—the AI system is likely to replicate and even

[18] Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (2015).

amplify those biases in its outputs.[19] In the context of criminal trials, this can lead to discriminatory outcomes, including wrongful identification or disproportionate targeting of certain groups. Such bias directly affects the reliability of AI-generated evidence and raises serious constitutional concerns relating to equality and fairness.

The problem of data integrity also plays a crucial role in determining authenticity. AI-generated evidence is dependent on large datasets, the accuracy and completeness of which are not always verifiable. Any errors, inconsistencies, or manipulations in the input data can significantly distort the output generated by the AI system.[20] Unlike traditional evidence, where the source can be directly examined, AI-generated outputs often obscure the link between input and result, making it difficult to establish a clear chain of custody.

A further challenge arises from the emergence of deepfakes and synthetic media, which have the potential to fabricate highly realistic audio, video, or images. These technologies can be used to create false evidence that is extremely difficult to detect using conventional methods.[21] The existence of such sophisticated manipulation tools undermines the presumption of authenticity traditionally attached to visual and audio evidence, thereby complicating judicial assessment.

Additionally, there is a risk of over-reliance on AI-generated evidence by courts and investigators. The perceived objectivity and technological sophistication of AI systems may lead to an uncritical acceptance of their outputs, without adequate scrutiny of their limitations.[22] This phenomenon, often described as "automation bias," can result in the undue elevation of AI-generated evidence over other forms of evidence, potentially distorting the fact-finding process.

The issue of lack of standardisation further exacerbates concerns regarding reliability. Different AI systems may employ varying methodologies, datasets, and algorithms, leading to inconsistent results. In the absence of uniform regulatory standards or certification mechanisms, it becomes difficult to ensure that AI-generated evidence meets a minimum

[19] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671 (2016).
[20] Paul W. Grimm et al., Authenticating Digital Evidence, 69 Baylor L. Rev. 1 (2017).
[21] Robert Chesney & Danielle Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753 (2019).
[22] Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147 (2017).

threshold of reliability required in criminal trials.

In light of these challenges, it is evident that traditional methods of assessing reliability and authenticity are insufficient when applied to AI-generated evidence. Courts must adopt a more nuanced approach that takes into account the technical complexities and inherent limitations of AI systems. This may include requiring greater transparency in algorithmic processes, mandating independent verification of AI outputs, and relying on expert testimony to evaluate the credibility of such evidence. Without such safeguards, the use of AI-generated evidence risks undermining the fairness and integrity of the criminal justice system.

Attribution of Legal Liability in AI-Driven Evidence

The increasing reliance on Artificial Intelligence in the generation and analysis of evidence raises a fundamental legal question: who should be held accountable when AI-generated evidence is erroneous, misleading, or results in wrongful conviction? Unlike traditional forms of evidence, where responsibility can be directly attributed to identifiable human actors, AI-generated evidence is the product of complex interactions between developers, data providers, and end-users. This diffusion of responsibility creates significant challenges in attributing legal liability within the criminal justice system.

At the outset, one possible approach is to attribute liability to law enforcement agencies that deploy AI tools during investigations. Under this framework, the State bears responsibility for ensuring that the tools used are reliable, properly tested, and deployed in accordance with legal standards. The use of flawed or biased AI systems by investigative authorities may amount to a violation of the accused's right to a fair trial, particularly under constitutional guarantees of due process and personal liberty.[23] However, placing exclusive liability on law enforcement may be problematic where the technical functioning of the AI system is beyond their expertise or control.

Another potential approach involves holding developers and technology providers accountable for defects in AI systems. From this perspective, liability may arise where an AI system produces erroneous outputs due to negligent design, inadequate training data, or failure to

[23] See Maneka Gandhi v. Union of India, (1978) 1 S.C.C. 248 (India).

address known biases.[24] However, applying traditional principles of criminal liability to software developers is complex, as it requires establishing intent or knowledge, which may not always be present. Civil liability frameworks, such as product liability, may offer a more appropriate mechanism, but their integration into criminal proceedings remains limited.

A third dimension concerns the role of judicial oversight and evidentiary evaluation. Courts play a crucial role in determining the admissibility and weight of AI-generated evidence. Failure to critically assess the reliability of such evidence may contribute to unjust outcomes. However, imposing liability on judges for reliance on AI-generated evidence is neither practical nor consistent with principles of judicial independence. Instead, the emphasis must be on developing robust standards and guidelines to assist courts in evaluating such evidence.

The issue of liability is further complicated by the autonomous nature of AI systems. In many cases, AI operates with minimal human intervention, making it difficult to identify a single point of failure. This raises the question of whether existing legal frameworks, which are primarily designed to regulate human conduct, are adequate to address harms caused by autonomous systems.[25] The absence of clear statutory provisions dealing with AI-related liability creates a significant legal vacuum, particularly in the context of criminal justice.

Comparative legal developments provide useful insights in this regard. Jurisdictions such as the European Union have begun to explore regulatory frameworks that emphasise accountability, transparency, and risk-based classification of AI systems.[26] These approaches recognise the need for assigning responsibility across different stakeholders involved in the lifecycle of AI systems, including developers, deployers, and users. While such models are still evolving, they highlight the importance of a multi-layered approach to liability.

In the Indian context, the absence of specific legislation addressing AI liability necessitates reliance on general legal principles, including constitutional safeguards and doctrines of state responsibility. However, these principles are not sufficient to address the unique challenges posed by AI-generated evidence. There is a pressing need to develop a comprehensive legal framework that clearly delineates responsibility among various actors and establishes

[24] Andrew D. Selbst & Solon Barocas, The Intuitive Appeal of Explainable Machines, 87 Fordham L. Rev. 1085 (2018).
[25] Ryan Calo, Robotics and the Lessons of Cyberlaw, 103 Calif. L. Rev. 513 (2015)
[26] Proposal for a Regulation of the European Parliament and of the Council on Artificial Intelligence (Artificial Intelligence Act), COM/2021/206 final.

mechanisms for accountability in cases of misuse or error.

Thus, the attribution of legal liability in the context of AI-generated evidence remains a complex and unresolved issue. A balanced approach is required, one that distributes responsibility among law enforcement agencies, developers, and regulatory authorities, while ensuring that the rights of the accused are adequately protected. Without such a framework, the increasing use of AI in criminal trials risks creating gaps in accountability that may ultimately undermine the legitimacy of the justice system.

Need for Legal and Policy Reforms

The growing integration of Artificial Intelligence into criminal investigations necessitates a comprehensive re-evaluation of the existing legal framework governing evidence in India. As discussed, the current statutory regime under the Bharatiya Sakshya Adhiniyam, 2023 and the

Bharatiya Nagarik Suraksha Sanhita, 2023 provides a foundational basis for the admissibility of electronic evidence but does not adequately address the unique challenges posed by AI-generated evidence. In light of the risks relating to reliability, authenticity, and accountability, there is an urgent need for targeted legal and policy reforms.

A primary reform requirement is the introduction of statutory recognition and classification of AIgenerated evidence. The law must clearly define what constitutes AI-generated evidence and distinguish it from conventional electronic records. Such classification would enable the development of tailored admissibility standards that take into account the algorithmic nature of such evidence, including requirements relating to transparency, explainability, and verification.

Secondly, there is a need to establish standards of algorithmic transparency and explainability. Courts must be empowered to demand disclosure of the underlying logic, methodology, and data sources used by AI systems.[27] This would enable effective judicial scrutiny and ensure that the defence has a fair opportunity to challenge the evidence. In cases where full disclosure is not feasible due to proprietary concerns, mechanisms such as independent audits or court-appointed experts may be employed to evaluate the reliability of AI systems.

[27] Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147 (2017).

Another crucial reform involves the development of uniform guidelines and certification mechanisms for AI tools used in criminal investigations. Regulatory bodies should be tasked with approving and monitoring AI systems to ensure compliance with predefined standards of accuracy, fairness, and security. Such certification would reduce inconsistencies and enhance the credibility of AI-generated evidence presented before courts.

Further, the law must incorporate procedural safeguards to prevent misuse of AI technologies. This includes regulating the deployment of tools such as facial recognition and predictive policing systems, particularly in light of concerns relating to privacy and mass surveillance. Judicial oversight mechanisms should be strengthened to ensure that the use of such technologies does not violate fundamental rights, especially the right to life and personal liberty under constitutional principles articulated in Maneka Gandhi v. Union of India.[28]

The issue of legal liability also requires legislative intervention. A clear framework must be established to allocate responsibility among different stakeholders, including law enforcement agencies, developers, and data providers. This may involve adopting a multi-layered liability model, where accountability is distributed based on the role played by each actor in the lifecycle of the AI system.[29] Such a framework would ensure that victims of wrongful convictions or procedural injustice have access to effective remedies.

In addition, there is a need for capacity building and judicial training. Judges, prosecutors, and defence lawyers must be equipped with the necessary technical understanding to evaluate AIgenerated evidence. Specialised training programmes and expert assistance can play a crucial role in bridging the knowledge gap between law and technology.

Comparative legal developments offer valuable guidance in this regard. Regulatory initiatives such as the European Union's proposed Artificial Intelligence framework emphasise risk-based regulation, accountability, and human oversight.[30] While the Indian context requires a tailored approach, such models highlight the importance of proactive regulation in addressing the challenges posed by emerging technologies.

Finally, a broader policy-oriented approach is required to balance innovation with the

[28] *Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248 (India).
[29] Ryan Calo, Robotics and the Lessons of Cyberlaw, 103 Calif. L. Rev. 513 (2015).
[30] Proposal for a Regulation of the European Parliament and of the Council on Artificial Intelligence (Artificial Intelligence Act), COM/2021/206 final.

protection of fundamental rights. The integration of AI into the criminal justice system should not come at the cost of fairness, transparency, or due process. Therefore, reforms must aim to create a legal environment where technological advancements enhance, rather than undermine, the administration of justice.

Conclusion

The integration of Artificial Intelligence into the criminal justice system marks a significant transformation in the nature and functioning of evidence in criminal trials. While AI-generated evidence offers substantial benefits in terms of efficiency, speed, and analytical capability, it simultaneously introduces complex legal challenges that test the foundational principles of criminal jurisprudence. The existing evidentiary framework under the Bharatiya Sakshya Adhiniyam, 2023, though progressive in recognising electronic records, does not adequately address the distinctive features of AI-generated evidence, particularly its reliance on opaque algorithmic processes.

As this paper has demonstrated, the admissibility of AI-generated evidence cannot be assessed solely through traditional standards developed for electronic records. Issues relating to authenticity, reliability, and procedural fairness become significantly more complex when evidence is derived from systems that are not fully transparent or explainable. Judicial precedents such as Anvar P.V. v. P.K. Basheer and Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal provide a foundational framework for electronic evidence, but their applicability to AI-generated outputs remains limited. The absence of clear legal standards governing such evidence creates uncertainty and increases the risk of inconsistent judicial approaches.

The challenges of reliability and authenticity are further compounded by issues such as algorithmic bias, lack of transparency, data integrity concerns, and the emergence of deepfake technologies. These factors not only undermine the probative value of AI-generated evidence but also pose serious threats to the right to a fair trial. Additionally, the question of legal liability remains unresolved, as existing legal doctrines are ill-equipped to address harms arising from autonomous or semi-autonomous systems.

In light of these concerns, it is imperative to adopt a cautious and principled approach towards the use of AI-generated evidence in criminal trials. The law must evolve to incorporate clear

standards of admissibility, robust safeguards for ensuring reliability, and well-defined mechanisms for

---

accountability. As suggested, this requires a combination of legislative intervention, judicial innovation, and policy reform, including the development of guidelines on algorithmic transparency, certification of AI tools, and allocation of liability among stakeholders.

Ultimately, the legitimacy of the criminal justice system depends on its ability to balance technological advancement with the protection of fundamental rights. While AI has the potential to enhance the administration of justice, its unregulated or uncritical use may lead to serious miscarriages of justice. Therefore, it is essential that the integration of AI into evidentiary processes is guided by the principles of fairness, transparency, and accountability, ensuring that technological progress serves as an instrument of justice rather than a source of injustice.