© IJARW | ISSN (O) - 2582-1008
February 2026 | Vol. 7 Issue. 8
www.ijarw.com

WHEN MACHINES TESTIFY: ARTIFICIAL INTELLIGENCE AND THE BHARATIYA SAKSHYA ADHINIYAM, 2023

Dr. Sunita Meena

Assistant professor Government Law College, Bundi, Rajasthan

ABSTRACT

The rapid evolution of Artificial Intelligence (AI) has transformed digital ecosystems with unprecedented precision and autonomy, simultaneously amplifying the sophistication of cybercrime and complicating the evidentiary landscape of criminal adjudication. As AI-generated outputs—particularly deepfakes and synthetic media—become increasingly indistinguishable from authentic material, courts are confronted with complex questions concerning admissibility, authenticity, reliability, and probative value. Within this emerging paradigm, the Bharatiya Sakshya Adhiniyam, 2023 marks a significant legislative shift by recognizing electronic and digital records as primary evidence. However, the statute does not comprehensively address the unique challenges posed by algorithmically generated and AI-processed material. This paper critically examines the evolving interface between AI technologies and the Indian law of evidence, assessing whether the existing statutory framework is normatively and procedurally equipped to respond to AI-driven evidentiary risks. Drawing upon doctrinal analysis, emerging judicial trends, and international jurisprudential developments concerning AI misuse, the study identifies structural gaps within the current evidentiary regime. It further undertakes a comparative exploration of regulatory approaches in the European Union and the United States to distil normative benchmarks for strengthening India's evidentiary framework. The paper ultimately argues for a principled recalibration of evidence law to preserve the integrity, fairness, and reliability of the criminal justice system in the age of intelligent automation.

Keyword: Artificial Intelligence, Deepfakes, Digital Evidence, Evidence, Reliability, Synthetic Media.

1. INTRODUCTION

Artificial Intelligence, referred to as AI, has permeated every facet of human existence in the 21st century. It is being created as a significant innovation for the improvement of humanity, with ramifications that extend far beyond its current infant stage of artificial intelligence. The application of artificial intelligence across diverse domains is consistently rising due to its ever evolving, dynamic, and adaptive learning capabilities, which often yield superior data synthesis and results compared to human intelligence. In the legal domain, where ongoing study is essential for proficiency and implementation, AI enhances outcomes through capabilities such as predictive coding, generative AI, explainable AI (XAI), continuous active learning models, and integration with blockchain technology. The Indian judicial system, constrained by inadequate infrastructure and a lack of technological expertise, faces considerable challenges regarding the implementation of AI, including the opaque nature of AI, inconsistencies in the chain of custody, biases in responses, and, crucially, judicial literacy.[1]

The notion of trial in courtrooms relies on the integrity of evidence evaluation, its admissibility, and evidentiary significance. The Bharatiya Sakshya Adhiniyam 2023 abrogated the Indian Evidence Act of 1872 and established electronic recordings as primary evidence, hence introducing issues in assessing the credibility of evidence in the evolving AI era.[2] Despite the stringent framework of Indian legislation concerning technology and evidence, including the Information Technology Act 2000 and the Digital

[1] Trishita Chatterjee, *Admissibility of Ai-Reviewed Digital Evidence in Legal Investigations,* V, IJIRL, 2056, 2060-2062, (2025).

[2] Tanmay Pradeep, *Comparison Analysis between the Indian Evidence Act, 1872 and the Bharatiya Sakshya Adhiniyam, 2023,* Volume 16, Issue 1, IJSAT, 1, 1-3 (2025).

Personal Data Protection Act 2023, as well as significant criminal statutes, it fails to thoroughly address the unique challenges posed by AI. Consequently, courts may increasingly encounter evidence that is, firstly, AI-generated rather than human-authored; secondly, AI-processed documents rather than original records; and thirdly, AI-interpreted rather than directly observable. The absence of explicit standards and regulations to govern AI-related risks may result in issues such as heightened wrongful attributions and diminished evidentiary reliability. In light of these concerns, this article examines whether the current statutory framework regulating electronic and digital evidence—specifically under the Bharatiya Sakshya Adhiniyam, 2023—is adequately prepared to confront the doctrinal and practical challenges posed by Artificial Intelligence within the Indian criminal justice system. It also analyzes the evidentiary hazards associated with AI-generated outputs, such as deepfakes and other synthetic media, concerning admissibility, authenticity, probative value, and judicial evaluation of evidence. The study conducts a comparative analysis of developed jurisdictions to find normative and procedural benchmarks that could be effectively implemented to enhance the Indian evidential framework. This research aims to assess the sufficiency of current regulations, pinpoint regulatory deficiencies, and recommend principled revisions to ensure the integrity, reliability, and fairness of AI-augmented judicial processes.

Given that both Artificial Intelligence and The Bharatiya Sakshya Adhiniyam 2023 are new advancements, scholarly research on these topics is still nascent. This work, both timely and intellectually sound, identifies a substantial gap that warrants additional investigation. This study recognizes deepfakes and synthetic media as significant hazards to the evaluation of evidence in a court of law; hence, the pressing issue is as follows -

''Existing scholarship focuses on the transition from The Indian Evidence Act, 1872, to The Bharatiya Sakshya Adhiniyam, 2023 as a modernisation of format. However, a critical gap remains that the essence of Bharatiya Sakshya Adhiniyam preserves a 'mechanical view of technology' that is obsolete in the age of generative AI".

This paper bridges the gap by demonstrating how the BSA's certification and authentication frameworks are insufficient for synthetic media, necessitating a new judicial standard for 'Algorithmic Probity'.

2. RECALIBRATING THE BHARATIYA SAKSHYA ADHINIYAM, 2023 FOR THE AGE OF INTELLIGENT AUTOMATION

- ***When 19th-Century Legal Architecture Meets 21st-Century Artificial Intelligence***

The Bharatiya Sakshya Adhiniyam, 2023 ("BSA 2023") implements a thorough revision of India's evidentiary system, replacing the Indian Evidence Act, 1872 with measures presumably designed for digital contexts.[3] The BSA 2023, introduced during India's comprehensive criminal law reforms alongside the Bharatiya Nyaya Sanhita and Bharatiya Nagarik Suraksha Sanhita, explicitly defines "electronic records" as documents, grants them primary evidence status under certain conditions, and establishes certificate requirements for their admissibility.[4] Legislative documents highlight "technology-neutrality" and "future-proofing," addressing decades of judicial development from ***Anvar P.V. v. P.K. Basheer (2014)*** to ***Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal (2020).***[5]

This modernization coincides with the rapid proliferation of generative artificial intelligence (GenAI)—multimodal systems such as OpenAI's GPT-4o, Google's Veo, and Stability AI's Stable Video Diffusion, which synthesize text, images, audio, and video from probabilistic latent spaces, frequently indistinguishable from human-generated content.[6] GenAI's evidential disruption is existential: it creates "synthetic originals" lacking real-world referents, undermining the perceptual reliability assumed by evidentiary law.[7] Danielle Citron and Robert Chesney's

[3] BSA, 2023 § 52.
[4] *Id.* § 2(1)(d), 61-63; Statement of Objects and Reasons, The Bharatiya Sakshya (Second) Bill, 2023, Bill No. 130 of 2023, Lok Sabha (India).
[5] Anvar P.V. v. P.K. Basheer, (2014) 10 SCC 473; Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 SCC 1.
[6] Google DeepMind, Veo: A New Frontier in Video Generation (2024), https://deepmind.google/technologies/veo.
[7] Hany Farid, *Seeing Is No Longer Believing: Detecting Deepfakes*, 372 Science 1396 (2021).

foundational paradigm delineates the "liar's dividend"—wherein falsified evidence contaminates sources of truth, while authentic evidence is disregarded as a deepfake.[8] This section does a detailed doctrinal evaluation of BSA 2023's sufficiency across conceptual, procedural, and institutional aspects, assessing its stipulations against GenAI scenarios and suggesting specific improvements.

- ***Mapping the Electronic Evidence Architecture of the Bharatiya Sakshya Adhiniyam, 2023***

1. Definitional Expansion: "Document" and "Electronic Record": BSA Section 2(1)(d) defines "document" expansively to subsume "electronic and digital records," aligning with Information Technology Act, 2000 definitions while obviating amendment dependency.[9] Section 2(1)(e) clarifies "electronic record" as data generated, received, or stored by computer, network, or device—capturing GenAI outputs as admissible artifacts. This resolves pre-BSA ambiguities where courts strained Evidence Act language for digital nativity.[10]

2. Primary Evidence Status: Sections 61–63: Section 61 revolutionizes admissibility by deeming electronic records "produced from proper custody" as primary evidence, abrogating *Anvar*'s secondary-evidence presumption for printouts.[11] Section 62 reinforces this for standalone devices; Section 63 mandates certificates attesting device functionality, production process, and integrity (hash values, chain-of-custody).[12] Section 63A integrates digital signatures under IT Act Section 3A, enabling PKI-verified provenance for transmissions.[13] Parliamentary debates hailed this as streamlining e-evidence without compromising reliability.[14]

3. Admissibility Gatekeepers: Judicial Notice and Expert Evidence: Section 57 (facts of which judicial notice must be taken) retains admissibility for public documents and scientific facts but omits GenAI benchmarks (e.g., NIST detection error rates).[15] Section 45 (opinions of experts) remains available for forensic authentication, yet lacks GenAI-specific protocols.[16]

- ***Algorithmic Disruption and the Limits of the BSA's Evidentiary Paradigm***

1. From "Records of Events" to "Synthetic Artifacts": BSA presumes electronic records document *real-world referents*—emails chronicle communications; CCTV captures events.[17] GenAI inverts this: diffusion models sample from training distributions to generate *de novo* content sans originals.[18] A Stable Diffusion video of a politician's "confession" bears perfect metadata yet zero ontological grounding. BSA's "best evidence" rule (Section 60 analogue) falters: no "original" exists to compare against.[19]

2. Perceptual Reliability Collapse: Human vision/audition (95%+ confidence in high-fidelity deepfakes) underpins Sections 59–60 (oral/documentary evidence). The Deepfake Detection Challenge (DFDC) datasets shows that top models achieve only 65–80% accuracy on adversarial samples.[20] Courts untrained in error-rate calculus risk "illusory truth" effects.[21]

3. Device-Centric Certificates vs. Model-Centric Risks: Section 63 certificates verify *post-generation integrity* (SHA-256 hashes unchanged) but ignore *generation provenance* (prompts,

[8] Danille Keats Citron & Robert Chesney, *Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security*, 107 CALIF. L. REV. 1753, 1757-60 (2019).
[9] BSA, 2023 § 2(1)(d); IT Act, 2000, § 2(1)(t).
[10] State (NCT of Delhi) v. Navjot Sandhu, (2005) 11 SCC 600, (India).
[11] BSA, 2023 § 61; Report of the Standing Committee on Home Affairs (Bharatiya Sakshya Bill, 2023), Lok Sabha Secretariat, (Nov. 10, 2023).
[12] BSA, 2023 § 62-63.
[13] *Id.* § 63A; IT Act, 2000 § 3A.
[14] Lok Sabha Debates on Bharatiya Sakshya Bill, 2023, (Dec. 2023) (India).
[15] BSA, 2023 § 57; National Inst. Of Standards & Tech., NISTIR 8456: Face Recognition Vendor Test (FRVT) Part 9: Face Recognition Accuracy with Synthetic Images (2024).
[16] BSA, 2023 § 45.
[17] *Id.* § 59-60, 61.
[18] Patrick Esser et al., *Improving Image Generation with Better Captions*, Proc. IEEE/CVF Conf. Computer Vision & Pattern Recognition 11937 (2024).
[19] BSA, 2023 § 60.
[20] *Id.* § 59-60; Deepfake Detection Challenge Results: An open initiative to advance AI, https://ai.meta.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/ (last visited Jan. 10, 2026).
[21] Lisa Fazio, *Repetition Increases Perceived Truth Even for Known Falsehoods*, 28 Collabra Psychol. 1, 4 (2022).

model weights, seeds).[22] A deepfake exported from RunwayML passes hash tests impeccably, masquerading as authentic surveillance.[23]

3. STRESS-TESTING THE BHARATIYA SAKSHYA ADHINIYAM, 2023 AGAINST GENERATIVE AI–DRIVEN FACT SCENARIOS

- ***Criminal Trial: Deepfake Confession (Sections 61, 63 and 27 Illustration)***

In a scenario where the prosecution produces a custodial confession video and the defence alleges that it is a DeepFaceLab-generated fabrication, the Bharatiya Sakshya Adhiniyam, 2023 would permit its admission as primary evidence under Section 61, subject to certification under Section 63 confirming lawful extraction and an unaltered hash value. The court may assess its evidentiary weight through judicial observation under Section 60 and expert techniques such as voice biometrics under Section 57. However, the statutory framework contains a critical gap: it does not mandate disclosure of AI model architecture, prompt logs, training data, or detection confidence scores. Thus, while formal admissibility may be satisfied, meaningful scrutiny of synthetic manipulation remains structurally constrained.

- ***Civil Fraud: LLM-Generated Emails (Sections 61–63A Illustration)***

In a civil dispute, a plaintiff produces purported "emails" generated through GPT-4o bearing forged signatures and certified as originating from a business server. Under the Bharatiya Sakshya Adhiniyam, 2023, such records may qualify as primary electronic evidence if supported by a valid PKI-based digital signature under Section 63A, with hash verification establishing technical integrity. Further, the presumption relating to documents produced in the ordinary course of business under Section 61 may operate in favour of admissibility. However, the framework reveals significant vulnerabilities: there is no mandatory disclosure of AI involvement in content generation, allowing algorithmically fabricated or hallucinated material to be laundered as authentic business communication. In the absence of expert testimony under Section 45 explaining the stochastic nature of large language models, courts may inadvertently presume human authorship, thereby risking evidentiary distortion.

- ***Harassment: Voice Clone Calls (Sections 59 and 63 Illustration)***

In a criminal proceeding, the prosecution relies upon an audio recording in which the complainant alleges that the accused issued threats; the defence contends that the voice is an ElevenLabs-generated clone. Under the Bharatiya Sakshya Adhiniyam, 2023, such material may be treated analogously to oral evidence under Section 59, supported by a proper extraction certificate to establish source authenticity. However, the evidentiary framework does not mandate waveform or spectrographic forensic analysis to detect synthetic voice manipulation. Moreover, expert testimony under Section 45—crucial to explaining AI-driven voice cloning techniques—remains institutionally scarce. Consequently, courts may confront significant limitations in distinguishing genuine speech from algorithmically generated audio, thereby heightening risks of misattribution.

REFORM ROADMAP: AI-RESILIENT BSA FRAMEWORK

- ***Statutory Amendments***

1. **Section 63B (AI Disclosure)**: Mandate affidavits detailing model, prompts, seeds, detection scores.
2. **Section 63C (Provenance Standards)**: Require C2PA/ISO 42001 compliance; rebuttable presumption for biometric media.
3. **Section 57(2A) (Judicial Notice)**: Incorporate NIST/DFDC error rates.[24]

- ***Procedural Innovations***

1. **Court-Appointed Experts**: BSA Section 45A for AI forensics panels.
2. **Probabilistic Gatekeeping**: Admissibility if authenticity score >80% (Daubert-plus).[25]

[22] BSA, 2023 § 63.
[23] Introducing gen-3 alpha: A new frontier for video generation (2024) Runway Research. Available at: https://runwayml.com/research/introducing-gen-3-alpha (last visited Jan. 1, 2026).
[24] Int'l Org. for Standardization, ISO/IEC 42001:2023, AI Management Systems (2023).
[25] Daubert v. Merrell Dow Pharm., 509 U.S. 579 (1993);

- ***Institutional Ecosystem***

1. **National AI Evidence Lab**: Under NFSU/CFSL, equipped with MediFor/Sentinel.
2. **Judicial Training**: NJA modules on ELMo classifiers, watermark verification.
3. **Regulatory Harmonization**: IT Rules mandate platform provenance APIs.

- ***Phased Implementation***

Rule-making under BSA Section 1(4): Phase 1 (2026)—disclosure mandates; Phase 2 (2027)—tool integration.

BSA 2023 vaults India into digital-evidence modernity but stumbles at GenAI's threshold. Its architecture—laudable for CCTV/emails—crumbles under synthetic loads, exposing courts to liar's dividends and miscarriages. Targeted reforms—disclosure, provenance, expertise—can forge an AI-resilient framework, preserving truth-seeking amid technological tumult. Absent these, BSA risks obsolescence by 2030, as GenAI fluency eclipses judicial safeguards.

4.DEEPFAKES AND SYNTHETIC EVIDENCE: EMERGING THREATS TO ADJUDICATORY RELIABILITY

The accelerated advancement of Artificial Intelligence has led to the emergence of synthetic media, commonly referred to as deepfakes. These hyper-realistic images, audio, and visual content are produced using sophisticated AI methodologies, including Generative Adversarial Networks and Diffusion models, rendering them indistinguishable from genuine media. This dynamic quality has resulted in a surge of malpractices and criminal activities in the digital realm, presenting challenges that may undermine judicial fact-finding, despite the implementation of countermeasures such as watermarking, metadata, and AI-based detection tools.[26]

The shift from the Indian Evidence Act of 1872 to the Bharatiya Sakshya Adhiniyam of 2023 has been substantial, as it necessitated digital certification of electronic records, hash value verification, and the acknowledgment of electronic records as primary evidence. However, these updated evidentiary standards face challenges in identifying evidentiary risks and are inadequately developed in regulating deepfakes and blockchain evidence. The courts may frequently encounter evidence generated, modified, evaluated, and processed by AI, rather than raw records created and presented by humans, due to the advanced developmental capabilities of AI. This transition results in the primary evidential dangers associated with AI, categorized into three distinct types: first, the failure of authentication and the inadequacy of conventional evidentiary standards; second, cognitive susceptibility; and third, the systematic deterioration of evidence integrity.

- ***Authentication Failure & The Insufficiency Of Traditional Evidentiary Standards***

Authentication failure occurs when the authentication process is exploited, enabling unauthorized users to impersonate legitimate users and access sensitive data and resources without permission.[27] Frequently underestimated as a trivial issue, authentication failure in judicial fact-finding undermines the procedural framework of evidentiary laws by presenting various challenges to the court and investigative bodies, often resulting in a binary dilemma for courts: either to admit seemingly authentic synthetic evidence or to exclude legitimate evidence based on unfounded claims.

The conventional evidentiary standards that have evolved from antiquated evidence laws to contemporary statutory frameworks carry specific implications, necessitating that digital evidence conforms to the probative value established in the Bharatiya Sakshya Adhiniyam, 2023, which encompasses authenticity, integrity, reliability, and admissibility.[28] Nonetheless, the deceptive characteristics of deepfakes compel the courts to ascertain whether the existing records represent authentic events, whether any manipulation transpired intentionally or inadvertently, and

BSA, 2023 § 1(4).

[26] Irene Amerini et al., *Deepfake Media Forensics: Status and Future Challenges,* J.Imaging 11, 73 (2025) [hereinafter *Status & Future Challenges*].

[27] Vinay Kulkarni et. al., *Centrify DC Authentication Failures: Patterns, Prevention, and Protocols,* Vol. 10 Issue 6 IJSRET 1, 1 (2024) .

[28] Tanmay Pradeep, *Comparison Analysis between the Indian Evidence Act, 1872 and the Bharatiya Sakshya Adhiniyam, 2023,* Volume 16 Issue 1 IJSAT 1, 1-3 (2025).

whether the content infringes upon an individual's rights. This complicates the differentiation between reality and alteration, thereby raising a critical issue regarding evidentiary value. The act presumes records to be reliable unless proven otherwise and stipulates the necessity of expert involvement for specialized authentication, as synthetic media may circumvent standard forensic verification. Consequently, the inconsistency in adjudicating AI-based evidence remains within the current framework, rendering the evidentiary standards inadequate.[29]

- ***Cognitive Vulnerability***

Cognitive vulnerability in psychology denotes a pattern of maladaptive thinking that heightens an individual's risk of adverse psychological outcomes, such as depression and anxiety. This condition is frequently marked by a pronounced intolerance of uncertainty, resulting in catastrophic interpretations of ambiguous information, which in turn provokes worry and anxiety.[30] The emergence of deepfakes and synthetic media has become a crucial area of research, as it may impact individuals' cognitive susceptibility and intellectual abilities. Its proliferation has resulted in the phenomenon of "impostor bias," wherein individuals frequently question the authenticity of multimedia content due to their awareness of synthetic media's capabilities, fostering a pervasive skepticism towards all digital evidence, whether genuine or not.[31] This imposter bias may activate the cognitive susceptibility of investigating officers and other personnel inside the chain of custody, thereby resulting in unfair outcomes such as wrongful convictions and acquittals. The recent case involving Ankur Warikoo, adjudicated by the High Court of Delhi, mandated the removal of deepfake synthetic content from various social media platforms. However, it failed to establish an evidentiary framework for courts to ascertain authenticity in instances where synthetic media is implicated in criminal contexts, indicating a deficiency in the systematic training of judicial officers across Indian states regarding this issue.[32]

- ***Systemic Collapse of Evidence Integrity***

Cognitive vulnerability in psychology denotes a pattern of maladaptive thinking that heightens an individual's risk of adverse psychological outcomes, such as depression and anxiety. This condition is frequently marked by a pronounced intolerance of uncertainty, resulting in catastrophic interpretations of ambiguous information, which in turn provokes worry and anxiety. The emergence of deepfakes and synthetic media has become a crucial area of research, as it may impact individuals' cognitive susceptibility and intellectual abilities. Its proliferation has resulted in the phenomenon of "impostor bias," wherein individuals frequently question the authenticity of multimedia content due to their awareness of synthetic media's capabilities, fostering a pervasive skepticism towards all digital evidence, whether genuine or not. This imposter bias may activate the cognitive susceptibility of investigating officers and other personnel inside the chain of custody, thereby resulting in unfair outcomes such as wrongful convictions and acquittals. The recent case involving Ankur Warikoo, adjudicated by the High Court of Delhi, mandated the removal of deepfake synthetic content from various social media platforms. However, it failed to establish an evidentiary framework for courts to ascertain authenticity in instances where synthetic media is implicated in criminal contexts, indicating a deficiency in the systematic training of judicial officers across Indian states regarding this issue.

1. Chain of custody: All evidence documentation adheres to a method that includes a chain of custody, which is the crucial connection required to establish the integrity of a piece of evidence and to verify its legitimacy before the court. It must record every transmission and transfer from the moment the evidence was obtained, ensuring that only authorized possession is permitted.[33] The

[29] Harmanjeet Singh & Dr. Ritu Panta, *Deepfake Evidence and the Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law, 2025,* Volume 7 Issue 6 IJFMR 1, 3-4 (2025) [hereinafter *Harmanjeet*].

[30] Taylor and Francis, Cognitive vulnerability – Knowledge and References, (last visited Dec. 22, 2025).

[31] Irene Amerini et al.; *Deepfake Media Forensics: State of the Art and Challenges Ahead, [cs. CV] arXiv:2408.00388*, 1, 1-2, (2024) [hereinafter *DMF: Challenges Ahead*].

[32] Ankur Warikoo & Anr. v. John Doe & Anr. CS(COMM) 514/2025.

[33] Badiye A, Kapoor N, Menezes RG. Chain of Custody.

chain of custody verifies if a piece of evidence has been altered. The authentication of AI-based evidence becomes more complex and demanding, since both the data and the algorithm must be verified within a specified timeframe. Nevertheless, the courts may have to depend on third parties and private enterprises to authenticate such data because of the functioning of AI within a distributed cloud infrastructure. Intermediaries possessing deepfake evidence must be proficient in recognizing such reconstructions along the chain of custody and employing case-specific verification methods, as robust protocols are essential to ensure evidence authenticity through meticulous provenance tracking.

2. Perpetual Detection Lag: Perpetual Detection Lag denotes the phenomenon wherein deepfake detection methodologies encounter significant challenges in realistic contexts, including time constraints and the ongoing training of AI detection models without succumbing to catastrophic forgetting. These models must also possess the ability to adapt, interpret, and accurately identify data. However, the advanced detection techniques struggle to keep pace with emerging generative technologies due to data drift, as deepfake creation technologies consistently progress at a rate that outstrips detection methods, resulting in an incessant reactive lag that compromises the reliability of these systems.

3. Explainability and Black-Box Evidence: A significant challenge in the realm of deepfake detection within the criminal justice system is the absence of explainability in the outcomes produced by deep learning-based detectors. While striving for optimal accuracy, these datasets function as Black boxes, rendering their internal reasoning and learning processes opaque and incomprehensible to judges and expert witnesses. This lack of transparency often obstructs the identification of the specifications responsible for the results, thereby complicating the decision-making process in courts.[34] Explainable and interpretable AI utilized in forensic investigations is crucial in legal and high-stakes contexts, as these synthetic data systems must deliver transparent and comprehensible decision-making processes. This clarity is vital for identifying biases and sources of error, fostering trust and credibility, and promoting expert collaboration and ongoing enhancement in the incorporation of AI within the criminal justice system.[35]

5. EVOLVING EVIDENTIARY DOCTRINES IN ADVANCED LEGAL SYSTEMS: A COMPARATIVE INQUIRY

- ***EUROPEAN UNION***

The principal evidentiary risks associated with AI are evident not only in India but also within the European Union framework. According to UNESCO, the admissibility of AI-generated evidence in courts presents a significant challenge. Judicial operators must cultivate an understanding of algorithms, their potential risks, biases, principles, and possible misapplications to navigate the complexities of AI and render informed decisions regarding evidence admissibility.[36]

However, the challenges got duly acknowledged and to address such AI posed risks, the European Union introduced the European Union Artificial Intelligence Act[37] which entered into force on August 1, 2024 and thereafter be effective from August 2, 2026.[38]

The EU Artificial Intelligence Act establishes several benchmarks, including a four-tier risk classification that organizes AI-related risks in order of severity: unacceptable risk (prohibited), high risk (stringently regulated), limited risk (transparency obligations), and minimal risk (unregulated). Additional comparative

[Updated 2023 Feb 13]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK551677/.

[34] E. Hydara et al.: *Empirical Assessment of Deepfake Detection: Advancing Judicial Evidence Verification,* Volume 12 IEEE Access 151188, 151191 (2024).

[35] Status & Future Challenges, *supra* note 34, at 15-16.

[36] *How to determine the admissibility of AI-generated evidence in courts?,* UNESCO NEWS (July 21, 2023), https://www.unesco.org/en/articles/how-determine-admissibility-ai-generated-evidence-courts?.com.

[37] THE ACT TEXTS | EU ARTIFICIAL INTELLIGENCE ACT, The AI Act Explorer | EU Artificial Intelligence Act [hereinafter *EU AI Act*] (last visited Dec. 22, 2025).

[38] Timo Gaudszun, et al., *AI Watch: Global regulatory tracker - European Union,* WHITE & CASE (July 21, 2025),https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-european- union?.com.

benchmarks established by the EU encompass transparency requirements to adequately inform end-users regarding interactions; a prohibition on deceptive AI to prevent the use of manipulated evidence for misleading purposes; a ban on exploiting vulnerabilities to ensure that AI-altered sensitive attributes are inadmissible; restrictions on criminal profiling and predictive evidence to ensure they do not supplant human judgment or objective facts; and a high-risk classification to guarantee stringent oversight and documentation when AI is involved in evidence generation or adjudication.[39]

The European Union has significantly advanced its evidentiary framework for regulating AI and deepfake synthetics by integrating legislation such as the EU Artificial Intelligence Act, the EU General Data Protection Regulation, and the Digital Services Act. This comprehensive approach aims to uphold standards of evidential adjudication, ensuring that synthetic media adheres to transparency obligations, prohibitions against manipulative, deceptive, and discriminatory practices, and rigorous oversight within justice and democratic contexts.

- ***UNITED STATES OF AMERICA***

The incorporation of artificial intelligence with the United States judicial framework has not led to bespoke statutory measures. When we look at how artificial intelligence fits into the American court system, no new laws have specifically emerged to handle it. Courts end up tweaking the Federal Rules of Evidence instead, those old standbys known as the FRE[40]. Mostly there exists a clash between how much courts trust outputs from algorithms and the sheer murkiness of those black-box systems that hide their inner workings.

Getting AI evidence into a trial relies heavily on judges acting as gatekeepers. Think back to ***Daubert v. Merrell Dow Pharmaceuticals***[41] that case really hammered down the above- mentioned idea, and it has been stirred and baked into FRE 702[42] now. Now we have to check if a method holds up scientifically, look at its error margins, see what peers say about it. But applying all that to machine learning setups? Serious hurdles pop up. Proprietary algorithms just don't offer the openness needed for proper checks. At this point let's take ***State v. Loomis***[43] in consideration, where Wisconsin's top court examined the COMPAS tool for predicting repeat offenses in sentencing. The usage of the tool was allowed, however, it was pointed out that there exists due process issues because the defendant cannot by any means peek at the secret code, setting a kind of standard in its operation. Evidence from these opaque systems can get misleading unless a system of checks and balances is being introduced.

Then there's authentication under FRE 901(b)(9)[44], which deals with verifying processes or setups, and it brings its own tough barriers. Traditional digital checks follow a clear chain of logic, step by step. Modern neural networks? They run on complex, twisting paths that aren't easy to explain. Add in generative AI or deepfakes, and suddenly courts demand strict verification steps to block faked stuff. The aim: keep out anything tampered with. Experts in law point out how trade secret laws often clash with the core right to challenge evidence through questioning, a conflict that lingers without clear fixes.

In this phase of AI's role in U.S. evidence rules, things seem headed toward demanding more clarity overall. Business motives frequently cloud how reliable this evidence really is, pitting openness against claims of ownership in ways that spark ongoing debates. Overall, U.S. evidence law in the Al era is moving toward a stronger expectation that Al systems be explainable. Judicial bodies have now started to confront the concept of AI in procedure implementation with constitutional protections, amid a context where evidentiary trustworthiness is often veiled by commercial interests.

6. CONCLUSION

In conclusion, the realm of AI presents both significant growth, progress, and innovation inside Indian Courts, as well as a new array of obstacles. The BSA's provisions represent a significant advancement over its predecessor, devoid of colonial remnants; yet, its current stipulations fail to meet the essential

[39] EU ARTIFICIAL INTELLIGENCE ACT, High-level summary of the AI Act | EU Artificial Intelligence Act (last visited Dec. 22, 2025).
[40] Fed. R. Evid. 28 U.S.C. app (2012).
[41] 509 U.S. 579 (1993).
[42] Fed. R. Evid. 702, (2012).
[43] 881 N.W.2d. 749 (Wis 2016).
[44] Fed. R. Evid. 901 (b) (9) (2012).

requirements stemming from the increasing utilization and abuse of AI. Particular electronic evidence, including CCTV footage, core internal software data, facial recognition, and the application of AI in hacking computer resources, undermines the traditional system of evidence evaluation, which relies on expert opinions that may not keep pace with the rapid advancements in AI technology. It is imperative to implement specific education and awareness initiatives for individuals engaged in cybersecurity and for agencies tasked with the gathering and examination of electronic evidence at the operational level. AI not only engenders scepticism but also heightens the risk of introducing false and fabricated evidence, which may appear compelling at first glance, potentially altering the trajectory of police investigations or, conversely, undermining the entire prosecution narrative. Although current penal laws encompass nearly all aspects of offenses related to the use or presentation of false or falsified evidence in court, the procedural and evidentiary components of the law require enhanced fortification to uphold and ensure justice in the contemporary era of AI unpredictability.