ARTICLE

ALGORITHMIC EVIDENCE IN THE DOCK: EXAMINING THE ADMISSIBILITY OF AI-GENERATED RECORDS UNDER THE BHARATIYA SAKSHYA ADHINIYAM, 2023 AND THE CASE FOR A DEDICATED NORMATIVE FRAMEWORK

Aniruddh Sachin Bajaj* **& Rishish Singh** +

Abstract

What may a court lawfully treat as established fact, and how ought it arrive at that conclusion? These are old questions, but artificial intelligence has given them a sharper edge than they have had for some time. The Bharatiya Sakshya Adhiniyam, 2023 (BSA) replaced the Indian Evidence Act, 1872 from 1 July 2024 - India's most significant reconfiguration of evidentiary law since the colonial period. Its reforms are real: electronic records are treated more coherently, expert evidence is better structured, and the certification regime is clarified. Yet the BSA was designed to answer a prior generation of problems. When Parliament enacted it, the governing concern was whether a stored or transmitted digital record had been altered in transit. Today's courts face something different: evidence that a machine did not merely store or convey but actively produced. A deepfake reconstructing a conversation, a recidivism score computed from demographic proxies, a DNA mixture resolved by probabilistic software - none of these maps cleanly onto the BSA's certification-based architecture. This article contends that the existing framework is both doctrinally inadequate and practically unworkable for this category of evidence. It draws on comparative law from the United States, the United Kingdom, and the European Union, together with scholarship in AI governance and legal epistemology, to develop a normative framework resting on three claims: that AI-generated evidence warrants a structured judicial reliability inquiry before admission; that Parliament must carve out a discrete statutory category with mandatory disclosure requirements; and that Articles 14, 20, and 21 of the Constitution place substantive limits on the judicial deployment of algorithmic outputs. The article closes with specific proposals for legislative amendment and practice directions.

Keywords – AI-generated evidence; Bharatiya Sakshya Adhiniyam 2023; algorithmic

I. INTRODUCTION

Every legal system rest upon some theory - rarely made fully explicit - of how a court should arrive at findings of fact. The Indian law of evidence, whose modern shape owes most to Sir James Fitzjames Stephen's 1872 codification,[71] proceeded from an assumption of direct human knowledge:

* 2nd Year B.A.,LL.B(Hons.) student at Gujarat National Law University, Gandhinagar
+ 2nd Year B.A.,LLB(Hons.) student at Gujarat National Law University, Gandhinagar

[71] *The Indian Evidence Act*, No. 1 of 1872 (India) (repealed). Drafted by Sir James Fitzjames Stephen and enacted under colonial administration, the Act remained largely intact for over 150 years before being superseded by the BSA.

the witness recounts what she saw, the document preserves what was written, the expert draws on training and observation. Over time, technology complicated this picture. Photographs, telephone recordings, and eventually electronic records each forced doctrinal adjustment. But those adjustments shared a common thread: in each case, the machine's role was custodial. It captured, stored, or transmitted something a human had done or said. The photograph showed what the lens faced; the server log recorded what was sent. Between the human act and the evidential artifact stood only a mechanical process.

Artificial intelligence breaks this pattern. When a generative model fabricates a video of someone uttering words they never spoke, it is not transcribing an event - it is inventing one. When a facial recognition algorithm returns a suspect's name from CCTV footage, it is not identifying a fact but making a probabilistic inference, one shaped by whatever demographic skews happened to characterise its training corpus. When probabilistic genotyping software deconvolutes a mixed DNA sample, the "conclusion" it produces is not the scientist's judgment but the model's output, derived from parameters the scientist may not fully understand. In none of these cases does the reliability of the evidence turn on recording fidelity. It turns on the model's architecture, the composition of its training data, the rigour of its validation, and the assumptions baked silently into its design. The BSA has no tools for examining any of this.[72]

The Bharatiya Sakshya Adhiniyam, 2023[73] is a genuinely improved statute in many respects. The certification jurisprudence that grew - often unsteadily - through *Anwar P.V. v. P.K. Basheer*[74] and *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*[75] is now better anchored in statute, and the expert evidence provisions in sections 45 to 47 offer cleaner structure than the 1872 Act provided. But the BSA is silent on AI-generated evidence. It does not define the category, does not identify what must be disclosed before such evidence is admitted, and does not impose any obligation on those deploying the systems that generate it. Section 63's certification requirement - a mechanism designed to guard against hardware malfunction - simply has nothing to say about whether a machine learning model's outputs are reliable for a given inferential task.

This article is structured in six parts. Part II examines the BSA's existing framework and the specific doctrinal gaps it leaves open. Part III disaggregates AI-generated evidence into three categories - synthetic media, predictive algorithmic outputs, and AI-augmented forensic reports - and considers the distinct challenges each presents. Part IV takes stock of the regulatory and judicial responses in the United States, United Kingdom, and European Union. Part V proposes a normative framework grounded in three principles. Part VI offers conclusions.

---

[72]Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. Davis L. Rev. 399, 418-420 (2017). Calo identifies the "opacity problem" — the inability of even technically sophisticated users to understand why an AI system reaches a particular output — as the central challenge for legal regulation of AI.

[73]*The Bharatiya Sakshya Adhiniyam*, No. 47 of 2023 (India). The Act came into force on 1 July 2024 alongside the Bharatiya Nyaya Sanhita, 2023 and the Bharatiya Nagarik Suraksha Sanhita, 2023, constituting a comprehensive legislative overhaul of the three principal pillars of Indian criminal law.

[74]*Anwar P.V. v. P.K. Basheer*, 2014 INSC 645 . The Supreme Court held unanimously that the certificate under s 65B of the Indian Evidence Act, 1872 was a mandatory condition precedent to the admissibility of electronic records and could not be replaced by oral evidence alone.

[75]*Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, 2020 INSC 453 . A three-judge bench confirmed and clarified *Anwar P.V.*, holding that the s 65B certificate must be produced at the time of tendering the electronic record, not at the appellate stage.

II. The Bharatiya Sakshya Adhiniyam, 2023: Architecture, Achievements and Omissions

The Statutory Architecture of Electronic Evidence

Sections 61 to 63 of the BSA represent the statute's principal engagement with electronic records, and by any fair assessment they are an improvement on what came before.[76] The definition of "electronic record" in section 61 is broad - broader, certainly, than the 1872 Act contemplated - and section 63's certification mechanism draws together the scattered jurisprudence that accumulated over two decades of litigation about the old section 65B. The requirement has been the most contested provision in Indian digital evidence law, generating a line of Supreme Court decisions that moved, not always consistently, from *Anwar P.V.*[77] through the partial relaxation attempted in *Shafhi Mohammad v. State of Himachal Pradesh*[78] and back to a firm restatement of the mandatory approach in *Arjun Panditrao.*[79]

The rationale behind the certification requirement is sensible and should not be discarded: before relying on a digital record, a court ought to have some basis for confidence that the record is what it purports to be and has not been corrupted or altered. The trouble is that this rationale was developed with a particular kind of digital evidence in mind - messages, transaction records, server logs - where the technology's job was to store and transmit data that humans had produced. A certificate attesting that a phone or computer was "functioning satisfactorily" in the ordinary course of business tells the court something useful about records of that kind. It tells the court essentially nothing useful about an AI system's outputs. A large language model summarising disputed messages, or a GAN producing a reconstructed video, does not simply store information - it constructs an output from parameters learned during training on datasets that may be vast, demographically skewed, and largely invisible to any certificating officer.[80]

The gap, in other words, is not merely technical. It is conceptual. The section 63 certification framework asks: was the machine working properly? The question that AI-generated evidence actually demands is: was the machine's inference sound for this particular task? Those are fundamentally different inquiries, and the BSA currently provides no mechanism for the second.

Expert Evidence and the Sections 45-47 Problem

The BSA's provisions on expert opinion - sections 45 to 47 - are the obvious candidate for filling the gap left by the certification framework. Section 45 renders relevant the opinion of persons "specially

---

[76] *The Bharatiya Sakshya Adhiniyam*, No. 47 of 2023, §§ 61-63 (India). Section 63 substantially re-enacts the certification mechanism of the repealed s 65B of the Indian Evidence Act, 1872, but widens the definition of "electronic record" to include data stored, generated or transmitted by any information communication technology device.

[77] *Anwar P.V.*, 2014 INSC 645 .

[78] *Shafhi Mohammad v. State of Himachal Pradesh*, 2018 INSC 75 . The court attempted to relax the mandatory certificate requirement where the party producing the record is not in possession of the relevant device. This dilution was subsequently disapproved and overruled in *Arjun Panditrao*.

[79] *Arjun Panditrao*, 2020 INSC 453 .

[80] Kamarrul Bahrin Mustafa, Mohd Zakhiri Md Nor & Sarina Othman, *Artificial Intelligence and Law: Redefining the Legal Framework*, 11 Asian J.L. & Just. 45, 51-55 (2022). The authors identify three modes in which AI intersects with law: as an object of regulation, as a tool deployed by legal actors, and as a source of evidence tendered in judicial proceedings.

skilled" in science, art, handwriting, or foreign law when the court needs to form a view on a question in that domain. In practice, this provision is the primary gateway through which AI-assisted forensic evidence enters Indian courtrooms.[81]

Two difficulties arise. The first is definitional: does expertise in artificial intelligence or machine learning count as expertise in "science" under section 45? The BSA does not say, and there is no settled judicial authority on the point. The practical consequence is significant. The reliability of most AI-generated evidence turns on precisely the kind of technical assessment - model architecture, training data composition, validation methodology, error rate analysis - that requires genuine AI expertise. If courts cannot call on such expertise under section 45, they will be left assessing algorithmic evidence without the intellectual tools to do so.

The second difficulty goes deeper. Section 45 treats expert opinion as going to the weight of evidence, not to its admissibility. The expert tells the court what to make of something already before it. But the question raised by AI-generated evidence is anterior to weight: it is whether the methodology that produced the evidence is reliable enough to justify placing the evidence before the court at all. This is the insight that animates the American *Daubert*[82] standard - that courts must act as gatekeepers, evaluating methodological reliability as a pre-admissibility question. Nothing in sections 45 to 47 of the BSA gives Indian courts a comparable tool. The structural consequence is that AI-generated outputs can reach the fact-finder under the endorsement of a human expert who may genuinely be unable to explain why the model produced the result it did.[83]

Relevancy and Reliability: A Doctrinal Gap

Running beneath both of the above problems is a more foundational tension in the BSA's evidentiary architecture. The statute, like its predecessor, organises admissibility around relevancy: section 7 provides that facts in issue and relevant facts are admissible, and everything else is not. Relevancy is a relationship between facts - whether knowing one makes another more or less probable. But reliability, in the sense that matters for AI-generated evidence, is something different: it is a property of the epistemic process that produced the alleged fact. An AI model's output can be highly relevant - it can bear directly on the central issue in a case - while being generated by a process so flawed or biased that it should not be admitted at all.

The present framework handles this mismatch by treating reliability as a matter of weight rather than admissibility. Unreliable evidence comes in and is then discounted. For conventional evidence - eyewitness testimony, expert reports produced by transparent methodology - this approach is workable because the fact-finder has enough to go on when assessing credibility. For opaque AI outputs, it is not. The fact-finder cannot meaningfully discount evidence whose unreliability is embedded in a computational process they cannot access and that even the tendering party may not

---

[81] *The Bharatiya Sakshya Adhiniyam*, No. 47 of 2023, §§ 45-47 (India). Section 45 empowers courts to seek expert opinion in science, art, handwriting, or foreign law. Whether expertise in artificial intelligence or machine learning constitutes expertise in "science" for this purpose has not yet been authoritatively determined.

[82] *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993). The US Supreme Court held that the trial judge must act as a "gatekeeper" for expert scientific testimony, assessing whether the expert's methodology is scientifically valid before it is heard by the jury.

[83] Christopher Slobogin, *Testilying: Police Perjury and What to Do About It*, 67 U. Colo. L. Rev. 1037, 1038 (1996). Slobogin's analysis of structural unreliability in testimonial evidence applies with equal force to algorithmic outputs admitted without rigorous methodological scrutiny.

fully understand.[84] The opposing party, meanwhile, cannot effectively challenge what it cannot examine. The BSA leaves both problems entirely unaddressed.

III. Three Categories of AI-Generated Evidence and Their Distinct Challenges

Generative AI and Synthetic Media: The Deepfake Problem

Generative adversarial networks and diffusion models have advanced to a point where video and audio fabrications are, in many instances, beyond the capacity of unaided human perception to detect. This creates evidentiary difficulties running in opposite directions simultaneously. Fabricated recordings may be tendered as genuine evidence of events that never took place - a confession, a statement, a meeting. Genuine recordings, conversely, may be contested as fabrications, giving parties a technical basis to dispute authentic evidence they find inconvenient. Both possibilities create serious problems for courts that have no reliable statutory framework for resolving them.

The Delhi High Court ran into exactly this problem in *Vijay Kumar v. State of NCT of Delhi.*[85] Competing expert affidavits were placed before the court about whether a recording had been digitally manipulated. Because no statutory standard existed specifying how such a question ought to be resolved - what qualifications the expert must hold, what analysis she must conduct, what disclosure the tendering party must make, what the challenging party can demand - the court was left to manage the dispute as best it could without institutional support. The outcome, though not unreasonable on its facts, offers nothing that future courts can reliably follow.

The underlying problem here is one of provenance - tracing the origin and integrity of an evidential artifact. For an ordinary electronic record, section 63 provides at least a partial answer: a responsible officer certifies that the device was in regular use and functioning as expected. For a deepfake, the provenance question is entirely different in character. Which generative model produced it? On what training data? Under what parameter settings? Is there a technically validated detection methodology capable of distinguishing this output from an authentic recording? The section 63 certificate cannot speak to any of this, and there is no other mechanism in the BSA that can.

There is also a constitutional concern that has received little attention. Article 20(3) protects accused persons from compelled self-incrimination. The Supreme Court's analysis in *Ritesh Sinha v. State of Uttar Pradesh*[86] draws the familiar distinction between testimonial acts, which the provision protects, and physical specimens, which it does not. The construction of a deepfake using a person's biometric likeness - without consent and for deployment against them in criminal proceedings - sits awkwardly within this framework. Whether it constitutes a form of compelled testimonial implication through the person's own likeness is a question current doctrine is not equipped to answer.

---

[84] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. Davis L. Rev. 399, 418-420 (2017).

[85] *Vijay Kumar v. State of NCT of Delhi*, 2023 SCC OnLine Del 3812 (Delhi HC). The court was asked to evaluate competing expert affidavits about whether a video had been digitally manipulated. Absent any statutory framework for assessing AI-generated manipulation, the court improvised, relying on party-appointed expert evidence without any structured reliability inquiry.

[86] *Ritesh Sinha v. State of Uttar Pradesh*, 2019 INSC 770 . While concerning compelled voice samples, the Court reaffirmed that novel forensic techniques are not ipso facto inadmissible; rather, courts must assess their scientific validity before attributing weight to outputs derived from them.

Predictive Algorithmic Evidence: Risk Scores and Facial Recognition

A second category of AI-generated evidence is constituted by the outputs of systems that do not fabricate but predict or classify - tools that process data about an individual and return an inference about their likely behaviour or identity. Two applications matter most in the Indian criminal justice setting: risk assessment scores used to inform bail and sentencing decisions, and facial recognition matches used to identify suspects.

The conceptual problem with risk scores was put clearly by Harcourt[87] some years before such tools reached Indian courts: a group-level statistical probability cannot be applied to an individual without committing a logical error. Saying that members of a certain demographic group reoffend at a particular rate does not tell you anything reliable about whether this particular person will. When the training data incorporates historical patterns of policing and prosecution that are themselves products of systemic inequality - as any dataset drawn from the Indian criminal justice system inevitably will be - the model will reproduce and arithmetically entrench those inequalities. The BSA's weight-based approach, which treats the risk score as evidence to be assessed alongside other material, cannot detect or correct for this structural problem.

Facial recognition presents related but technically distinct difficulties. Buolamwini and Gebru's research[88] established that error rates for commercial facial recognition systems can reach 34.7% for darker-skinned women - groups systematically underrepresented in training datasets - while approaching zero for lighter-skinned men. Indian police forces have deployed facial recognition at significant scale across Delhi, Hyderabad, and Bengaluru,[89] in the absence of any regulatory framework governing minimum accuracy standards, required demographic validation, or mandatory disclosure of error rates. A system with a materially elevated false positive rate for any identifiable group, used to generate criminal identification evidence, creates a risk of wrongful identification that cannot be undone through judicial weight-assessment after the fact.

The constitutional dimension is not subtle. Article 14's equal protection guarantee is, at its core, a prohibition on the state treating similarly situated persons differently, and on the state deploying tools that produce systematically less accurate results for one group than another.[90] Deploying a demonstrably biased facial recognition system in criminal proceedings is difficult to square with this guarantee even without proof that any particular identification was erroneous. The right to privacy established in *K.S. Puttaswamy v. Union of India*[91] adds a further layer: the collection and processing

---

[87]Bernard Harcourt, *Against Prediction: Profiling, Policing and Punishing in an Actuarial Age* 3-10 (2007). Harcourt's central argument is that actuarial tools systematically disadvantage individuals who share statistical characteristics with high-risk groups, regardless of individual culpability.

[88]Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 Proc. Mach. Learn. Res. 1, 1-15 (2018). This study found error rates of up to 34.7% for darker-skinned women in commercial facial recognition systems that performed near-perfectly for lighter-skinned men.

[89]S. Sunder Rajan, *The Datafication of Criminal Justice in India: Structural Risks and Constitutional Concerns*, 54 Econ. & Pol. Wkly. 34, 37-39 (2022). The author documents deployment of facial recognition technology by police forces in Delhi, Hyderabad and Bengaluru, noting the absence of regulatory framework governing accuracy thresholds or demographic validation.

[90]V.N. Shukla's Constitution of India 88-92 (M.P. Singh ed., 13th ed. 2017) (commentary on Art. 14). Article 14 guarantees equality before the law and equal protection of the laws. The deployment of a biased AI system in criminal adjudication constitutes a form of systemic discrimination incompatible with this guarantee.

[91]*K.S. Puttaswamy v. Union of India*, 2017 INSC 801 . The nine-judge bench unanimously held that privacy is a fundamental right under Art. 21 of the Constitution of India, with implications for AI systems that process biometric and behavioural data in generating judicial evidence.

of biometric data as a step in generating criminal evidence engages proportionality analysis,[92] and the current BSA framework does not require courts to conduct that analysis at any stage.

AI-Augmented Forensic Reports: The Attribution Problem

A third and less publicly visible category involves forensic science. AI tools have become standard in a number of forensic disciplines: probabilistic genotyping software interprets complex mixed DNA profiles; machine learning models compare ballistic striations; digital forensics platforms use algorithmic triage to recover and classify data from seized devices. In each of these settings, the resulting report carries a human expert's name and signature. The substantive analytical work, however, has been performed by the AI system. This mismatch between nominal authorship and actual epistemic agency is what this article calls the attribution problem.

The law of expert evidence resolves this mismatch by treating the human signatory as the author of the opinions expressed. She must hold those opinions genuinely, explain their basis, and submit to cross-examination. But where the basis of the opinion is an opaque machine learning model's output, the expert may simply not know why the system reached the conclusion it did - and may not be in a position to find out, particularly where the software is proprietary and the source code is a trade secret. Pasquale's account of the "black box" problem[93] identifies this difficulty precisely: the internal reasoning of complex machine learning models is frequently inaccessible not only to lay users but to the researchers who developed them. Cross-examination of a human expert who lacks access to the system's decision pathway is, in any practical sense, pointless. The adversarial process - upon which Indian evidentiary law relies for quality assurance - simply cannot function as designed.[94]

This is not a hypothetical concern. The probabilistic genotyping software deployed in DNA casework has been the subject of sustained legal challenge in the United States and United Kingdom precisely because of the tension between the defendant's right to challenge the evidence and the developer's intellectual property rights in the software. Indian courts will face the same tension. The BSA provides no framework for navigating it.

IV. Comparative Perspectives: Lessons for a Reform Agenda

The United States: Gatekeeping and Its Discontents

The most developed framework for novel scientific evidence in any common law jurisdiction is the American *Daubert*[95] framework, as subsequently elaborated in *Joiner* and *Kumho Tire*[96] and codified

---

[92] *K.S. Puttaswamy*, 2017 INSC 801 . The proportionality standard articulated in *Puttaswamy* requires that any limitation on the right to privacy — including the use of biometric data to generate criminal evidence — must have a legitimate aim, be rationally connected to that aim, and not go beyond what is necessary.

[93] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* 8-14 (2015). Pasquale coins the term "black box" for algorithmic systems whose internal decision processes are opaque to those affected by them, a description applying with particular force to deep learning models deployed in forensic contexts.

[94] *Harjinder Singh Kalra v. State of Punjab*, CRM-M-57458 of 2022 (India). The Supreme Court reiterated that in matters of digital evidence, credibility is always a question for the court and expert evidence, however persuasive, is never conclusive.

[95] *Daubert*, 509 U.S. 579.

[96] *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997); *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999). These decisions extended the *Daubert* gatekeeping function to non-scientific expert testimony and confirmed that the inquiry ultimately concerns reliability of methodology rather than formal categorisation.

in Rule 702 of the Federal Rules of Evidence. Under this framework, before expert testimony may be placed before a jury, the trial judge must determine that it is grounded in sufficient factual material, reflects reliable principles and methodology, and represents a genuine application of those methods to the facts of the case. The judge, in short, acts as gatekeeper - not to exclude legitimate science, but to ensure that jurors are not placed in the position of having to evaluate claims they cannot assess.

The translation of this framework to AI-generated evidence has been neither straightforward nor uniform. Courts that have applied it rigorously have required proponents to establish error rates, demonstrate peer validation, and show that the methodology enjoys genuine expert consensus. Courts more reluctant to engage have effectively reverted to the pre-*Daubert* standard articulated in *Frye*[97] - asking simply whether the technique is generally accepted, a question that proves hospitable to AI systems that have achieved commercial ubiquity without rigorous independent validation. The most persistent difficulty is the proprietary system problem: a methodology-based reliability inquiry is structurally compromised when the methodology is a trade secret.

What the American experience establishes, despite its inconsistencies, is that reliability-based gatekeeping is not only possible but necessary for AI-generated evidence. It also reveals where the approach breaks down without accompanying disclosure obligations. A reliability inquiry conducted without access to training data, architecture documentation, and validation studies is an inquiry conducted in the dark. India cannot simply import *Daubert* - but it cannot afford to ignore the structural insights it embodies.

The United Kingdom: Caution Without a Framework

English law's engagement with computer-generated evidence is, in retrospect, a cautionary tale about the dangers of premature legislative retreat. Section 69 of the Police and Criminal Evidence Act 1984 imposed a positive obligation to prove that a computer producing an evidential output had been working correctly. Parliament repealed it in 1999, on the reasoning that computers had become sufficiently commonplace and dependable that a presumption of reliability could replace the positive requirement. Subsequent courts applied - and continue to apply - that presumption without pausing to consider whether it remains appropriate as the "computers" in question evolve into deep learning systems whose designers cannot fully explain their own outputs.

The Law Commission of England and Wales confronted this problem squarely in its 2021 consultation paper,[98] recommending that courts develop a structured reliability inquiry for novel or complex algorithmic evidence rather than continuing to rely on a blanket presumption rooted in the 1990s understanding of computing. The recommendation awaits legislative response. Judicial practice in the meantime has been case-by-case: in *R v. Atkins*[99] the Court of Appeal accepted facial

[97]*Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923). The "general acceptance" test requires a novel scientific technique to be generally accepted by the relevant scientific community before evidence produced by it is admissible. The test has been criticised for entrenching established paradigms at the expense of newer, potentially more reliable methodologies.

[98]Law Commission of England and Wales, *Digital Evidence and Expert Reports in Criminal Proceedings* (Consultation Paper No. 245, 2021), https://www.lawcom.gov.uk/. The Commission expressed concern that courts continue to apply a presumption of computer reliability to complex AI systems that do not justify such a presumption.

[99]*R v. Atkins*, [2009] EWCA (Crim) 1876. The Court of Appeal held facial mapping evidence based on photographic comparison admissible, but emphasised that the reliability of the specific methodology employed must be established independently for each case.

mapping evidence in principle while insisting that the reliability of the particular methodology must be demonstrated afresh for each case. The result is doctrinal progress without institutional consistency. India would be wise not to replicate England's mistake of repealing a reliability requirement before building a replacement, but it should also avoid the English court's slow-moving common law approach when a more deliberate legislative solution is available.

The European Union: A Risk-Based Framework with Evidentiary Implications

The EU Artificial Intelligence Act, 2024[100] is the most architecturally comprehensive statutory response to AI governance yet enacted. It works by classifying AI systems according to their risk profile. Systems that assist courts in fact-finding are designated high-risk under Annex III, with the consequence that their providers must maintain detailed technical documentation, build in automatic logging sufficient for post hoc auditing, and communicate clearly to deployers the system's known limitations.[101] These are not primarily evidentiary rules - the EU AI Act is not an evidence statute - but their practical effect on AI evidence is profound. A system subject to these obligations can, at least in principle, be examined after the fact: its error rates can be established from logs, its design can be reviewed in documentation, and its limitations can be disclosed to courts and parties before they are asked to rely on its outputs.

Wachter and colleagues have noted the limits of transparency requirements in this context.[102] Knowing a system's architecture does not always explain a particular output, and a description of a model's general capabilities may give little guidance about its reliability for a specific inferential task in a specific population. These are genuine cautions, but they argue for strengthening transparency requirements rather than abandoning them. Without disclosure of design and training, no meaningful reliability inquiry is possible at all. India's emerging governance architecture - the NITI Aayog's responsible AI principles[103] and the Digital Personal Data Protection Act, 2023[104] - provides some foundation for movement in this direction, though neither instrument addresses the evidentiary context directly.[105]

---

[100]Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence (EU AI Act), 2024 O.J. (L 2024/1689), http://data.europa.eu/eli/reg/2024/1689/oj. AI systems used in the administration of justice are classified as high-risk under Annex III and subject to requirements of transparency, documentation, human oversight, and auditability.

[101]EU AI Act, *supra* note 24, arts. 11-13. Article 11 requires technical documentation; Article 12 mandates logging capabilities enabling post hoc audit; Article 13 imposes transparency obligations requiring deployers to be informed of the system's capabilities and limitations.

[102]Sandra Wachter, Brent Mittelstadt & Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 Int'l Data Privacy L. 76, 77-80 (2017). The authors caution that even robust transparency requirements may not adequately resolve the opacity of complex machine learning models.

[103]NITI Aayog, *Responsible AI for All: Approach Document for India* 18-23 (2021), https://www.niti.gov.in/. The document identifies accountability and explainability as two of seven core principles but stops short of prescribing legally binding obligations.

[104]*The Digital Personal Data Protection Act*, No. 22 of 2023 (India). The Act establishes a framework for processing personal data but does not specifically address the use of personal data in the generation or evaluation of judicial evidence.

[105]Ministry of Electronics and Information Technology, Government of India, *National Strategy for Artificial Intelligence* 12-18 (2018), https://www.niti.gov.in/. The strategy acknowledges AI's transformative potential but does not address the evidentiary implications of AI systems deployed in law enforcement or judicial contexts.

V. Towards a Normative Framework: Legislative and Judicial Responses

Three Foundational Principles

Building a workable normative framework for AI-generated evidence in Indian law requires principles that can be operationalized by courts and translated by legislators into concrete rules. Three principles, each grounded in existing legal commitments and responsive to distinct dimensions of the problem identified above, are proposed here.

The first principle is explainability. Before an AI system's output can properly be used as evidence, the party tendering it must be able to give the court a coherent account of how the output was produced: what task the system was designed to perform, what data it learned from, how its outputs were validated against known outcomes, and what the principal uncertainties are. That account cannot be generic. A system validated for one population may not be explainable as reliable for another; a tool calibrated on one jurisdiction's criminal justice data may not be explainable as applicable in a different social and institutional context. The cross-examination rights preserved in sections 138 to 140 of the BSA already implicitly presuppose that the basis of expert opinion can be exposed to adversarial testing; explainability makes this aspiration real rather than formal for AI-generated evidence. The epistemological case for treating explainability as a precondition rather than a factor going to weight draws on Ferrando's analysis of algorithmic evidence[106] and Goldman's social epistemology:[107] evidence that cannot be explained cannot meaningfully be weighed against competing evidence, because the fact-finder has no rational basis for doing so.

The second principle is auditability. Explainability without independent verification is merely the tendering party's account of its own system's reliability. For the account to carry epistemic weight, the system must be amenable to examination by someone other than its developer or operator. This requires technical documentation to be preserved and made available; training data to be accessible, at minimum to a court-appointed expert under appropriate safeguards; and the conditions of the specific output in question to be documented with enough precision to permit a meaningful audit. The NIST framework for identifying and managing bias in AI systems[108] provides a useful methodological template: it asks about training data composition, validation methodology and conditions, known error rates disaggregated by demographic variables, and post-deployment monitoring. An audit framework modelled on these questions would give Indian courts a practical basis for the reliability inquiries the evidence demands.

The third principle is proportionality. The degree of scrutiny an AI-generated output must survive before being admitted ought to track the gravity of the proceeding and the significance of the evidence within it. Where an AI output is the primary or only basis for a criminal conviction, the

---

[106]Tomaso Ferrando, *Evidence in the Time of Algorithms: Epistemic Foundations and Procedural Challenges*, 34 Leiden J. Int'l L. 891, 898-902 (2021). Ferrando argues that algorithmic evidence challenges the foundational premise of judicial fact-finding: that facts are independent of the method used to identify them.

[107]Alvin Goldman, *Knowledge in a Social World* 79-85 (1999). Goldman's social epistemology provides theoretical grounding for the proposition that the reliability of an AI system cannot be assessed in isolation but must be evaluated against the epistemic practices of the community in which it is deployed.

[108]Nat'l Inst. of Standards & Tech., U.S. Dep't of Commerce, NIST Special Publication 1270, *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence* 8-12 (2022), https://doi.org/10.6028/NIST.SP.1270. NIST recommends a lifecycle approach to bias identification encompassing data curation, model training, deployment, and post-deployment monitoring.

inquiry must be exhaustive: demonstrated reliability for the specific task, known and acceptable error rates, and a genuine adversarial opportunity to challenge. Where AI evidence appears as one thread of corroborating material in a civil proceeding, something lighter may suffice. This graduated approach reflects the constitutional imperatives of Articles 14, 20, and 21, all of which bear on the procedural and substantive quality of evidence in criminal proceedings.

Legislative Amendments: What the BSA Requires

Three targeted amendments to the BSA would substantially remedy the doctrinal deficiencies identified in Part II above.

First: Parliament should insert a definition of "AI-generated record" as a distinct sub-category within the BSA's treatment of electronic evidence. The definition should capture any record whose substantive content is generated, materially reconstructed, or inferentially produced by a machine learning system - as distinguished from records that are merely stored or transmitted electronically. Critically, the definition should be framed in technology-neutral terms, so that it covers not only current generative AI and predictive systems but future architectures that may work quite differently. The Law Commission's 2023 recommendations[109] identified the need for definitional clarity in this area but drew back from recommending a discrete statutory category. This article argues that the step the Commission declined to take is the necessary one: the evidentiary challenges posed by AI-generated evidence are sufficiently distinct from those posed by conventionally produced electronic records to require separate statutory treatment.[110]

Second: Parliament should enact an AI evidence disclosure certificate as a mandatory condition of adducing any AI-generated record in judicial proceedings. Unlike the section 63 certificate - which a responsible officer of a device produces - this certificate should come from a qualified AI systems expert and should contain: the identity and version of the system; the purpose for which it was built and the purpose for which it was actually deployed; a description of the training data and its known demographic limitations; the validated error rate and the conditions under which validation was conducted; any material discrepancy between those conditions and the conditions of deployment; and available post-deployment monitoring data. The certificate is not conclusive of reliability. It is the evidentiary foundation - the basic factual record - on which the judicial reliability inquiry described below can be conducted. It should be open to cross-examination like any other exhibit.

Third: sections 45 and 61 of the BSA should be amended to provide expressly that expertise in artificial intelligence or machine learning constitutes expertise in "science" within the meaning of section 45, and to empower courts to appoint a neutral technical expert to conduct an independent audit of any AI system whose output has been tendered in evidence where the reliability of that output is genuinely in dispute. The audit should be funded by the party tendering the evidence. The neutral expert's report should be made available to all parties and open to examination. This mechanism is consistent with the Supreme Court's inherent powers under Article 142 and would

[109]Law Commission of India, *Report No. 288: Reforms in the Law of Evidence with Reference to Electronic Evidence* 34-41 (2023), https://lawcommissionofindia.nic.in/. The Commission recommended amendments to the BSA to clarify the evidentiary treatment of AI-assisted tools but stopped short of recommending a distinct statutory category for AI-generated records.

[110]*Shreya Singhal*, 2015 INSC 257 . The absence of a statutory definition of "AI-generated record" produces precisely the kind of legal uncertainty the Court condemned in *Shreya Singhal*: parties and courts are left to reason by analogy from provisions designed for a categorically different technological context.

give courts the institutional capacity to conduct reliability inquiries without depending exclusively on party-funded experts whose independence may be compromised.

Judicial Guidelines: A Structured Reliability Inquiry

Legislative change cannot resolve every case the courts will encounter, and the speed at which AI technology develops means that any statutory framework requires regular revisiting. It is accordingly important that alongside legislative reform, the Supreme Court of India issue practice directions establishing how courts should approach AI-generated evidence - both under the existing framework pending legislative action, and under any revised framework once enacted. The analogy most apt to the Indian context is the Court's guidance on video-conferencing evidence in *State of Maharashtra v. Praful B. Desai,*[111] which gave courts a structured and workable approach to novel technology-mediated evidence.[112]

Stage one: classification. The court asks whether the evidence before it is an AI-generated record within the statutory definition. Expert assistance is available at this stage on application. If the evidence falls within the definition, the subsequent stages are engaged. If it does not, ordinary admissibility principles apply.

Stage two: threshold reliability. The tendering party must satisfy the court that the AI system's methodology is sufficiently reliable to justify the evidence being considered. The standard is not proof beyond doubt but demonstrated validity: the system must have been tested against known outcomes, its error rate must be established, and any significant sources of systematic bias must be identified and disclosed. The court has discretion to reject the evidence at this stage if the showing cannot be made, regardless of how probative the evidence might otherwise be.

Stage three: contextual validity. A system that is generally reliable may nonetheless be unreliable for the particular inferential task it has been asked to perform in this case, or for the particular population to which it has been applied. The tendering party must address this specifically, demonstrating that the system's reliability extends to the application at hand. This is the Indian equivalent of the "fit" requirement in *Daubert,*[113] adapted to the institutional context of Indian courts.

Stage four: weight. Once the evidence has cleared the threshold reliability and contextual validity assessments, the court determines what weight to give it alongside all other material in the case. Practice directions should alert courts explicitly to the risk of "automation bias" - the documented tendency to over-defer to algorithmic outputs relative to competing human judgment - and should direct courts to calibrate weight with particular care to the limitations disclosed at earlier stages.

Constitutional Constraints and the Fair Trial Imperative

The proposals above rest on constitutional foundations that are independent of the statutory arguments. The right to a fair trial - grounded in Articles 20 and 21, elaborated across decades of Supreme Court jurisprudence - is not merely a procedural guarantee. It is a substantive constraint on

---

[111] *State of Maharashtra v. Praful B. Desai*, 2003 INSC 205 . The Supreme Court held that recording evidence by video-conferencing was as valid as recording evidence in the physical presence of the accused, establishing judicial receptiveness to technology-mediated evidentiary processes.

[112] *Praful B. Desai*, 2003 INSC 205 . The Court's endorsement of technology-mediated judicial processes in *Praful B. Desai* was premised on the capacity of the court to evaluate the reliability of the technology employed, a premise that does not hold where the technology is an opaque AI system.

[113] *Daubert*, 509 U.S. 579.

the quality of the evidence on which criminal adjudication rests. A court that convicts on evidence generated by an unreliable algorithmic process does not simply make a factual error: it violates the accused's right to be judged on a rational, contestable, and reliable evidentiary foundation. The epistemic quality of fact-finding is a constitutional matter, not merely an administrative one.

The constitutional argument reaches Parliament as well. Maintaining the evidentiary status quo - in which AI tools of demonstrably uneven accuracy may be deployed against criminal defendants without any obligation of disclosure or reliability inquiry - is not a position of neutrality. It is a choice to expose persons from socially and economically marginalised groups to a disproportionate risk of wrongful conviction, given the documented demographic disparities in AI system performance. Article 14 requires not merely formal procedural equality but substantive equity in the operation of adjudicative tools. The NHRC's engagement with structural dimensions of state accountability[114] provides the constitutional anchor for that argument. The Supreme Court's reasoning in *Shreya Singhal*[115] - that legal ambiguity produces chilling effects on the exercise of constitutional rights - extends, by structural analogy, to the evidentiary domain: uncertain admissibility standards for AI evidence create their own chilling effects, distorting the conduct of investigations and prosecutions in ways that may be difficult to detect and correct.

VI. Conclusion

Indian evidentiary law has a long history of adapting - haltingly, imperfectly, but ultimately - to new technologies. The photograph, the audio recording, the electronic message: each arrived before the law was ready for it and each eventually prompted the courts or the legislature, often in dialogue with each other, to develop workable rules. AI-generated evidence is the current iteration of that recurring challenge, and it is a harder one than previous iterations because the epistemic structure of the problem is qualitatively different: the machine does not just record, it infers, and its inferences may be wrong in ways that are invisible, systematic, and correlated with existing social inequalities. The Bharatiya Sakshya Adhiniyam, 2023 did not address this problem. A framework capable of processing a section 63 certificate for a WhatsApp message will struggle badly when confronted with a facial recognition match, a recidivism score, or a generative video tendered to prove what someone said or did. The three-part framework developed in Part V of this article - a statutory definition coupled with mandatory disclosure, a court expert mechanism, and a four-stage reliability inquiry - is designed to give courts the institutional means to manage these encounters rationally, and to give Parliament a clear legislative agenda if it chooses to act.

Two things should be said clearly by way of conclusion. The first is that this framework is not sceptical of AI. A facial recognition identification that has been properly validated for the relevant population, disclosed in accordance with the proposed certificate requirements, independently audited, and subjected to genuine adversarial testing is, if it survives all of that, a more reliable piece of evidence than one admitted with none of those assurances. The framework promotes rational admission, not systematic exclusion. The second is that the moment for deliberate legislative

[114] *National Human Rights Commission v. State of Arunachal Pradesh*, 1996 INSC 38 . While the substantive holding concerned a different matter, the case affirms the principle that the state must exercise its investigative powers in a manner accountable to fundamental rights guarantees.

[115] *Shreya Singhal v. Union of India*, 2015 INSC 257 (India). While principally a free speech case, the Court's reasoning about the chilling effect of vague legal standards applies by analogy to the evidentiary context: uncertain admissibility standards for AI evidence create perverse incentives for both prosecution and defence.

intervention is passing. Case law is accumulating around an inadequate framework, habits of practice are forming, and institutional inertia is building. To act before those patterns solidify is the difference between shaping the law and being shaped by it.

How Indian courts come to know the facts of cases before them will increasingly depend on algorithms - systems whose workings are opaque, whose errors are demographic, and whose reliability, without the kind of framework this article proposes, will be taken on trust rather than demonstrated by evidence. The law of evidence, which is at root the law of what courts may claim to know, must answer to this. The question before the Indian legal system is not whether AI-generated evidence will reach its courts. It already has. The question is what standards that evidence must meet before it may be acted upon.