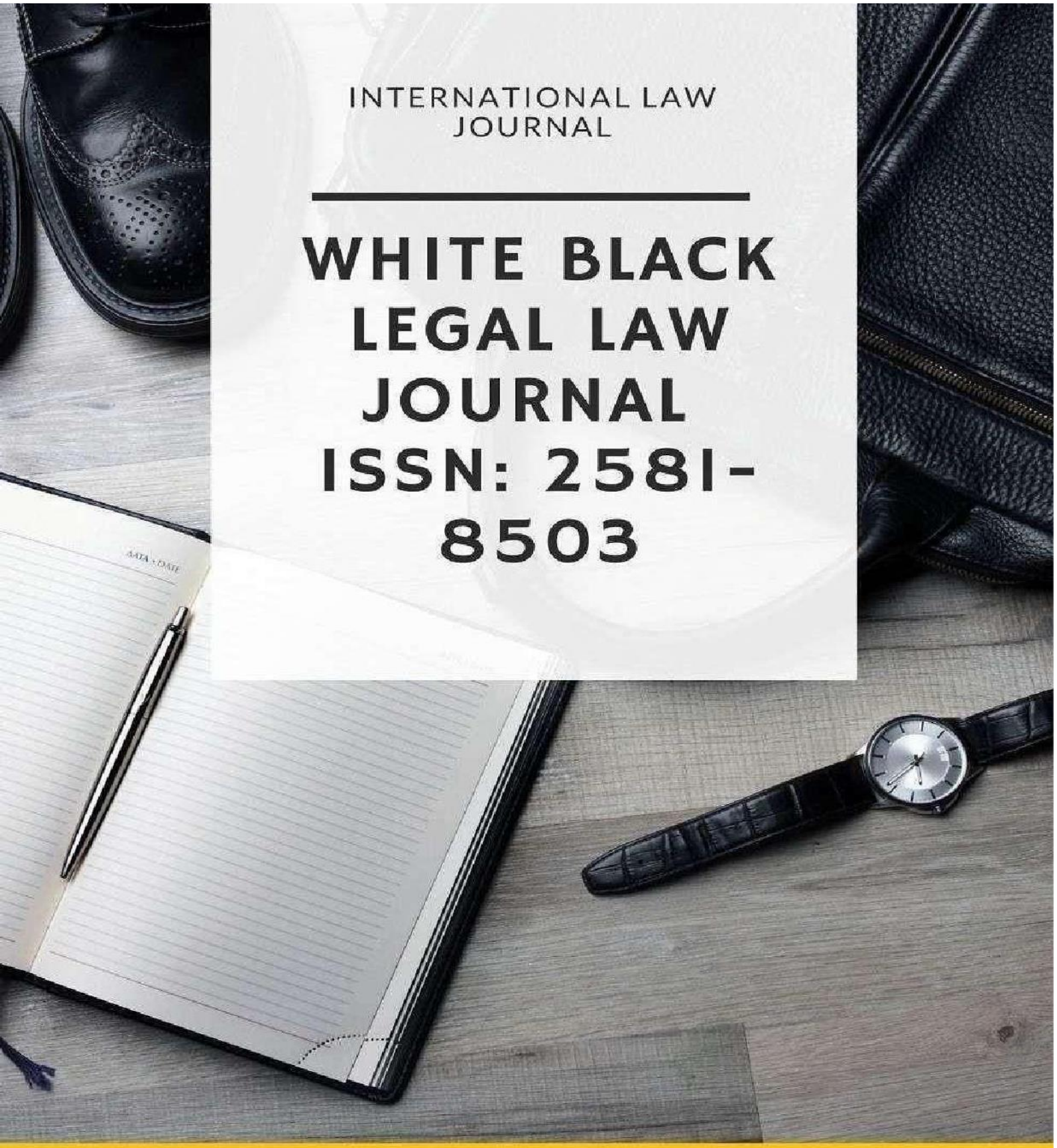


INTERNATIONAL LAW JOURNAL

WHITE BLACK LEGAL LAW JOURNAL

ISSN: 2581-8503

Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

DISCLAIMER

No part of this publication may be reproduced, stored, transmitted, translated, or distributed in any form or by any means—whether electronic, mechanical, photocopying, recording, scanning, or otherwise—without the prior written permission of the Editor-in-Chief of *White Black Legal – The Law Journal*.

All copyrights in the articles published in this journal vest with *White Black Legal – The Law Journal*, unless otherwise expressly stated. Authors are solely responsible for the originality, authenticity, accuracy, and legality of the content submitted and published.

The views, opinions, interpretations, and conclusions expressed in the articles are exclusively those of the respective authors. They do not represent or reflect the views of the Editorial Board, Editors, Reviewers, Advisors, Publisher, or Management of *White Black Legal*.

While reasonable efforts are made to ensure academic quality and accuracy through editorial and peer-review processes, *White Black Legal* makes no representations or warranties, express or implied, regarding the completeness, accuracy, reliability, or suitability of the content published. The journal shall not be liable for any errors, omissions, inaccuracies, or consequences arising from the use, interpretation, or reliance upon the information contained in this publication.

The content published in this journal is intended solely for academic and informational purposes and shall not be construed as legal advice, professional advice, or legal opinion. *White Black Legal* expressly disclaims all liability for any loss, damage, claim, or legal consequence arising directly or indirectly from the use of any material published herein.

ABOUT WHITE BLACK LEGAL

White Black Legal – The Law Journal is an open-access, peer-reviewed, and refereed legal journal established to provide a scholarly platform for the examination and discussion of contemporary legal issues. The journal is dedicated to encouraging rigorous legal research, critical analysis, and informed academic discourse across diverse fields of law.

The journal invites contributions from law students, researchers, academicians, legal practitioners, and policy scholars. By facilitating engagement between emerging scholars and experienced legal professionals, *White Black Legal* seeks to bridge theoretical legal research with practical, institutional, and societal perspectives.

In a rapidly evolving social, economic, and technological environment, the journal endeavours to examine the changing role of law and its impact on governance, justice systems, and society. *White Black Legal* remains committed to academic integrity, ethical research practices, and the dissemination of accessible legal scholarship to a global readership.

WHITE BLACK LEGAL

AIM & SCOPE

The aim of *White Black Legal – The Law Journal* is to promote excellence in legal research and to provide a credible academic forum for the analysis, discussion, and advancement of contemporary legal issues. The journal encourages original, analytical, and well-researched contributions that add substantive value to legal scholarship.

The journal publishes scholarly works examining doctrinal, theoretical, empirical, and interdisciplinary perspectives of law. Submissions are welcomed from academicians, legal professionals, researchers, scholars, and students who demonstrate intellectual rigour, analytical clarity, and relevance to current legal and policy developments.

The scope of the journal includes, but is not limited to:

- Constitutional and Administrative Law
- Criminal Law and Criminal Justice
- Corporate, Commercial, and Business Laws
- Intellectual Property and Technology Law
- International Law and Human Rights
- Environmental and Sustainable Development Law
- Cyber Law, Artificial Intelligence, and Emerging Technologies
- Family Law, Labour Law, and Social Justice Studies

The journal accepts original research articles, case comments, legislative and policy analyses, book reviews, and interdisciplinary studies addressing legal issues at national and international levels. All submissions are subject to a rigorous double-blind peer-review process to ensure academic quality, originality, and relevance.

Through its publications, *White Black Legal – The Law Journal* seeks to foster critical legal thinking and contribute to the development of law as an instrument of justice, governance, and social progress, while expressly disclaiming responsibility for the application or misuse of published content.

SEEING IS NO LONGER BELIEVING: DEEPFAKE EVIDENCE AND THE BHARATIYA SAKSHYA ADHINIYAM, 2023

AUTHORED BY - DR AZAD KUMAR DWIVEDI[1]

Abstract

Artificial intelligence has transformed the evidentiary assumption that photographs, videos and voice recordings provide comparatively objective representations of reality. Deepfakes - synthetic or materially manipulated audio-visual content created through artificial intelligence and machine-learning techniques can convincingly attribute words, conduct and events to individuals who never spoke, acted or appeared as represented. Their emergence therefore creates a dual evidentiary danger: fabricated media may be accepted as genuine, while authentic recordings may be dismissed through an allegation that they are AI-generated, a phenomenon often described as the "liar's dividend." The Bharatiya Sakshya Adhiniyam, 2023 (BSA), which came into force on 1 July 2024, modernises Indian evidence law by expressly recognising electronic and digital records, expanding the conception of primary evidence and establishing a detailed framework under Section 63 for computer outputs. However, a crucial distinction remains between admissibility, integrity and authenticity. A Section 63 certificate or cryptographic hash may demonstrate the manner in which a digital file was produced and whether the submitted copy remained unchanged, but neither necessarily proves that the depicted event actually occurred or that the voice, face or conduct represented in the file belongs to the apparent individual. This paper examines the treatment of deepfake evidence through Sections 2, 39, 57, 61 - 63, 86 and related provisions of the BSA, traces the judicial evolution of electronic evidence authentication, and analyses recent Indian developments concerning synthetic media, personality rights and AI-generated content. Particular attention is given to the Information Technology Amendment Rules introduced in February 2026, which establish duties concerning synthetically generated information, labelling and provenance. The paper argues that India requires a provenance-oriented and technologically neutral evidentiary framework combining chain of custody, native-file preservation, metadata analysis, cryptographic verification, forensic expertise, platform records and independent corroboration.

[1] Principal, City Academy Law College, Lucknow

Such an approach can preserve the probative potential of digital evidence without permitting either technological deception or unsupported allegations of manipulation to undermine the administration of justice.

Keywords: Deepfakes; Bharatiya Sakshya Adhiniyam, 2023; Electronic Evidence; Section 63; Digital Forensics; Synthetic Media; Authentication.

1. Introduction: The Collapse of the "Seeing Is Believing" Presumption

The history of evidence law has repeatedly reflected the technology through which facts are recorded and communicated. Oral testimony was supplemented by documentary evidence, paper records were supplemented by photographs and tape recordings, and the widespread use of computers and smartphones eventually required courts to develop legal principles for electronic records. Artificial intelligence now presents a qualitatively different evidentiary problem. Earlier questions generally concerned whether an existing recording had been edited, altered, copied or improperly produced. Generative artificial intelligence makes it possible to create realistic speech, facial expressions, movements and entire audio-visual scenes that may never have existed in the physical world at all.[2] Consequently, a court confronting a video can no longer safely begin with the intuitive proposition that the camera must at least have recorded some version of the depicted event. The video itself may be synthetic.

The expression "deepfake" broadly refers to synthetic or manipulated audio, images or video generated or substantially modified through artificial intelligence techniques so that the resulting material falsely but convincingly depicts a person, object, statement or event.[3] Early deepfakes were commonly associated with face-swapping technologies based on deep learning. Contemporary systems are considerably more sophisticated. Generative adversarial networks, diffusion models, voice-cloning systems, lip-synchronisation technologies and multimodal generative models can reproduce facial appearance, vocal characteristics, mannerisms and environmental details with remarkable realism. A single photograph may be sufficient to animate a face, while comparatively short voice samples can facilitate synthetic speech. The evidentiary implications are especially serious because the same technologies are increasingly

[2] Setty Hrudai Tej, "Regulating AI-Generated Deepfake Evidence Under the Bharatiya Sakshya Adhiniyam, 2023," 16(3) *The IUP Journal of Law Review* 5–21 (2026).
[3] Mika Westerlund, "The Emergence of Deepfake Technology: A Review," 9(11) *Technology Innovation Management Review* 39–52 (2019).

available through consumer applications rather than being confined to specialist laboratories.

The problem is not merely that courts might admit fabricated evidence. Deepfakes undermine confidence in authentic evidence as well. Once society becomes aware that almost any recording can theoretically be fabricated, a litigant confronted with damaging genuine material may simply allege that it was created through artificial intelligence. This phenomenon, generally described as the "liar's dividend," converts the existence of deepfake technology into a defensive weapon. A genuine CCTV recording, telephone conversation, intimate communication, political speech or confession may be attacked not through concrete proof of alteration but through the abstract proposition that synthetic media are technologically possible.[4] Evidence law must therefore avoid two opposite errors. Excessive credulity risks admitting fabricated media; excessive scepticism risks making reliable digital recordings practically unusable.

The Bharatiya Sakshya Adhiniyam, 2023 represents India's principal statutory response to the broader transformation of evidence in a digital environment. Section 2 expressly brings electronic and digital records within the concept of a "document", while Sections 57, 61, 62 and 63 address primary evidence, electronic records and computer outputs. The legislation therefore acknowledges that contemporary evidence exists across smartphones, computer systems, servers, cloud environments and digital networks rather than exclusively in conventional physical form.[5]

Yet the BSA was not drafted as a specialised "deepfake statute". Its central electronic evidence provisions continue the foundational logic that had developed under Sections 65A and 65B of the Indian Evidence Act, although with important modifications. The central question for synthetic media is therefore not simply whether an electronic record can enter the court room. It is whether the law possesses sufficiently sophisticated mechanisms to decide what that record actually proves. A perfectly preserved file can still contain a perfectly manufactured falsehood. Similarly, a certificate establishing that a video was copied correctly from a computer says little about whether the video was generated by that computer through an AI model before being copied. Evidentiary integrity and historical authenticity are related but distinct concepts.

[4] Robert Chesney and Danielle Keats Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security," 107 *California Law Review* 1753, 1757–1820 (2019).
[5] The Bharatiya Sakshya Adhiniyam, 2023 (Act 47 of 2023), ss. 2, 39, 57, 61 – 63

The distinction becomes especially important in criminal adjudication. A synthetic recording purporting to show an accused committing an offence may influence bail, investigation, public opinion and ultimately guilt. Conversely, the accused may allege manipulation against genuine surveillance footage. In matrimonial disputes, a cloned voice could be presented as evidence of cruelty or admission; in sexual offences, synthetic intimate imagery could create both criminal harm and evidentiary confusion; in corruption prosecutions, fabricated conversations could simulate demands for illegal gratification; and in electoral or communal contexts, deepfake speeches can generate consequences before forensic verification is even possible.[6] These examples demonstrate that deepfakes are not a narrow cyber law issue. They touch relevance, proof, authentication, expert evidence, constitutional fairness, privacy, reputation and procedural integrity.

This paper argues that the BSA provides a useful statutory foundation but not a complete solution. Its recognition of digital records, detailed certification process, expanded conception of primary evidence and recognition of electronic evidence experts constitute significant improvements. At the same time, the reliability problem generated by AI requires courts to move beyond a binary enquiry into whether Section 63[7] has technically been complied with. Authenticity must become a distinct evidentiary enquiry addressing provenance, generation history, metadata, chain of custody, source device information, cryptographic integrity, platform records, forensic indicators and external corroboration. Recent legislative developments concerning synthetically generated information strengthen this emerging architecture, but procedural evidence law must integrate these developments into a coherent judicial methodology.

WHITE BLACK LEGAL

2. Deepfakes as an Evidentiary Technology: Fabrication, Manipulation and the Liar's Dividend

A proper legal approach to deepfakes begins by recognising that synthetic media encompass different technological processes with different evidentiary implications. Some material is wholly generated: the person, voice or event appearing in the recording may never have existed in the represented form. Other material is partially manipulated: a real video may be retained while the face is replaced, a genuine speech may be altered to change particular words, or

[6] Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* 1–35, 119–170 (Institute of Advanced Legal Studies, University of London, London, 4th edn., 2017).
[7] The Bharatiya Sakshya Adhiniyam, 2023 (Act 47 of 2023), s. 63

authentic footage may be synchronised with a cloned voice. Still other content may involve ordinary editing, enhancement, compression or restoration that changes technical characteristics without falsifying the underlying event. Evidence law therefore cannot treat every digitally modified file as fraudulent. The legally significant question is whether the manipulation materially affects the proposition for which the evidence is tendered.

This distinction has become particularly important because the evidentiary authenticity of digital media is layered. At the first layer lies file integrity: whether the exhibit before the court is the same sequence of digital data that was collected at an earlier stage. Cryptographic hashing is highly valuable for answering this question because even a small subsequent change normally alters the hash value. At the second layer lies source integrity: whether the file was actually produced by the claimed device, account, camera, platform or person. Metadata, device identifiers, server logs and account records become important here.[8] The third layer is content authenticity: whether the recording accurately represents the real-world event that it appears to depict. Deepfakes are especially dangerous because a file may satisfy the first two layers while failing the third. A synthetic video generated on a particular laptop can be perfectly preserved, hashed, copied and certified. Every evidentiary step after generation may be impeccable, yet the video remains false as a representation of reality.

Traditional digital forensic techniques remain essential but are consequently insufficient in isolation. Examination may reveal inconsistent frame structures, unusual compression patterns, irregular lighting, facial boundary anomalies, mismatched reflections, spectral irregularities in audio, discontinuities in electrical network frequency, or inconsistencies between lip movement and speech. Metadata may reveal unexpected editing software or creation dates. However, sophisticated generative models increasingly reduce these artefacts, while social media platforms routinely recompress material and strip metadata.[9] A video downloaded from an online messaging application may therefore lose precisely the information that would otherwise have helped authenticate it. This means that courts should prefer examination of the native or earliest reasonably obtainable version rather than relying only on a forwarded clip, screen recording or social media download.

[8] Luisa Verdoliva, "Media Forensics and DeepFakes: An Overview," 14(5) *IEEE Journal of Selected Topics in Signal Processing* 910–932 (2020).
[9] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales and Javier Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," 64 *Information Fusion* 131–148 (2020).

Deepfake detection software introduces another difficulty. AI-based detectors generally infer manipulation by identifying statistical patterns associated with known forms of synthetic generation. Their outputs may be probabilistic rather than categorical. A detector might conclude, for example, that a recording has a high probability of artificial generation, but the legal meaning of such a percentage depends upon the system's validation, dataset, false-positive rate, false-negative rate and ability to identify the particular generation technique at issue. A model trained primarily on particular languages, facial types, codecs or generation architectures may perform differently in Indian evidentiary environments.[10] Recompression, poor lighting or ordinary video enhancement can also interfere with automated analysis. Courts should therefore resist treating an AI detector as an electronic equivalent of an infallible truth machine.

The problem is intensified by the "detection-generation arms race". Detection tools learn to identify artefacts left by existing generators; newer generators then reduce or eliminate those artefacts. A forensic conclusion that is well supported against one class of models may not generalise to another. The correct judicial question is not merely whether a forensic tool calls a file "fake", but how the tool reached that conclusion, what data it was validated against, whether the relevant error rates are known, whether the method has been independently tested, and whether other evidence supports the inference.[11] Digital forensic expertise must consequently be subjected to the same evidentiary discipline expected of scientific evidence generally.

The liar's dividend complicates the question in the opposite direction. If the mere assertion "this could be a deepfake" were sufficient to exclude digital evidence or create reasonable doubt, the development of generative AI would perversely benefit anyone whose conduct happens to have been captured electronically. Courts therefore require a threshold methodology. A bare allegation of manipulation should not automatically displace properly authenticated evidence. The party challenging authenticity should ordinarily identify some concrete basis for the challenge, such as discontinuities in the file, inconsistencies in provenance, evidence of account compromise, conflicting metadata, expert analysis, inability to locate the original, or factual impossibility within the recording. Once a credible basis is

[10] Luciano Floridi, "Artificial Intelligence, Deepfakes and a Future of Ectypes," 31(3) *Philosophy & Technology* 317–321 (2018).
[11] *Supra* Note 6

demonstrated, however, a stronger evidentiary foundation should be expected from the proponent.

Provenance becomes critical in this context. Provenance refers to information describing the origin and history of digital content: who or what created it, when and where it was created, what device or software was involved, whether it was modified, and how it moved between systems. Emerging technological standards such as the Coalition for Content Provenance and Authenticity framework seek to attach cryptographically verifiable information concerning the source and modification history of digital media.[12] Such systems cannot themselves guarantee substantive truth - a person can record a staged event authentically - but they can materially improve the capacity to distinguish known-origin media from anonymously circulating content.

The evidentiary response to deepfakes must accordingly be technologically neutral. A statute tied to the artefacts of today's generative models would rapidly become obsolete. The more sustainable legal approach is to require demonstrable provenance, reliable preservation, transparent forensic methodology and contextual corroboration. Such an approach does not depend upon determining whether a particular generation architecture was used. It asks a more durable question: has the proponent supplied enough reliable information about the origin, integrity and factual context of the digital object to justify treating it as evidence of the asserted event?

The importance of this approach can be understood through a hypothetical criminal prosecution. Suppose a video appears to show an accused threatening a victim. If the prosecution possesses the victim's original smartphone, the native file, consistent creation metadata, contemporaneous cloud backup, matching location data, testimony from a person who witnessed the recording, platform transmission logs and forensic confirmation that no signs of synthetic manipulation are detectable, the evidentiary foundation is strong even though theoretical deepfake technology exists. By contrast, if the prosecution possesses only a heavily compressed forwarded clip downloaded from an unidentified social media account, with no native source, no traceable uploader, no metadata and inconsistent audio characteristics, the visual realism of the clip should not substitute for authentication. The difference illustrates why

[12] Coalition for Content Provenance and Authenticity, *Content Credentials: C2PA Technical Specification*, Version 2.2 (2025).

"seeing" can no longer be equated with "believing"; courts must increasingly examine how the thing seen came into existence.

3. The Bharatiya Sakshya Adhiniyam, 2023 and the New Architecture of Digital Evidence

The BSA significantly modernises the statutory language through which Indian courts deal with electronic information. Section 2(1)(d) defines a "document" in a manner expressly encompassing electronic and digital records.[13] The statutory illustrations include electronic records on computers, laptops and smartphones, e-mails, server logs, documents, messages, websites, locational evidence and voice-mail messages stored on digital devices. The importance of this drafting choice lies in the normalisation of digital evidence: an electronic record is no longer conceptually exceptional simply because it lacks a paper form. It is part of the ordinary documentary universe of contemporary adjudication.

Section 57 is particularly important in the deepfake context because the BSA expands the meaning of primary evidence for electronic and digital records.[14] Where an electronic or digital record is created or stored simultaneously or sequentially in multiple files, each such file may constitute primary evidence. An electronic or digital record produced from proper custody is treated as primary evidence unless it is disputed. Where a video recording is simultaneously stored, transmitted, broadcast or transferred to another location, each stored recording may be primary evidence. The Act also addresses automated storage, including temporary files. These provisions respond to the distributed nature of digital information, where insisting upon a single physically identifiable "original" may be technologically artificial.

For deepfake disputes, however, the phrase "unless it is disputed" assumes considerable importance. A party alleging synthetic manipulation is precisely disputing the authenticity of the digital object. The expanded primary evidence rule should therefore not be interpreted as creating a blanket presumption that a file emanating from proper custody is substantively genuine. Custody can support provenance, but it cannot establish that content was not synthetically manufactured before it entered that custody. The provision is best understood as resolving questions concerning the status of digital originals and copies, not as abolishing the requirement to prove that the material represents what its proponent claims.

[13] The Bharatiya Sakshya Adhiniyam, 2023 (Act 47 of 2023), s. 2
[14] *Id*. s. 57

Sections 61, 62 and 63 form the core statutory architecture for electronic records.[15] Section 61 provides that an electronic or digital record cannot be denied admissibility merely because it exists in electronic form and gives it the same legal effect as other documents subject to Section 63. Section 62 directs that the contents of electronic records may be proved in accordance with Section 63. Section 63 then establishes the conditions under which information contained in an electronic record, when printed, stored, recorded or copied in specified forms as a computer output, is deemed to be a document and may be admitted without production of the original computer or device, provided the statutory conditions are met. These conditions substantially continue the functional approach familiar from Section 65B of the Indian Evidence Act.

The statutory conditions focus on reliability of the computer process. The computer or communication device must have been regularly used to create, store or process information during the relevant period by a person having lawful control over it; information of the relevant kind must have been regularly fed into the system in the ordinary course; the system must have been operating properly or any malfunction must not have affected the accuracy of the record; and the information reproduced in the computer output must be derived from data regularly supplied to the system. Section 63(3) adapts these principles to modern computing by recognising computers, systems, networks, computer resources used for creation, processing or storage, and intermediaries rather than assuming a single isolated desktop machine.

Section 63(4) significantly formalises certification. A certificate is required to identify the electronic record, describe the manner in which it was produced, give particulars of the device involved and address the statutory conditions. The Schedule to the BSA provides a structured certificate. It seeks information regarding source devices including computers, storage media, digital video recorders, mobile devices, flash drives, CDs, servers and cloud environments. It further contemplates identification particulars such as make, model, serial number, IMEI, unique identifiers, MAC address or cloud identifier, and specifically calls for hash values such as SHA-1, SHA-256 or MD5. Part B is designed for completion by the expert. These requirements are significant because they move the legislation closer to a forensic conception of electronic proof rather than treating certification as a purely formal affidavit.

Nevertheless, Section 63 must be interpreted carefully. The certificate primarily addresses how

[15] *Id*. ss. 61. 62, 63

an electronic record was produced and the reliability of the relevant computer process. A hash value can establish that two digital files are identical at the bit level or that a file has not changed after the hash was calculated. It cannot establish that the underlying content is historically true. If a deepfake video is generated on 1 October, hashed on 2 October and tendered in court on 10 October with an unbroken chain of custody, the matching hash proves that the same deepfake travelled through that evidentiary chain. It does not transform the fabricated event into reality. Consequently, Section 63 compliance should be treated as an essential component of admissibility and integrity where applicable, not as conclusive authentication of the depicted event.

The BSA also contains provisions relevant to expert analysis. Section 39(2) provides that when the court has to form an opinion concerning information transmitted or stored in a computer resource or in electronic or digital form[16], the opinion of an Examiner of Electronic Evidence referred to under Section 79A of the Information Technology Act, 2000[17] is a relevant fact, and such examiner is treated as an expert. This provides an important statutory bridge between evidence law and forensic expertise. Deepfake disputes will increasingly require courts to utilise such expertise not only to determine whether a file was copied correctly but also to examine generation artefacts, metadata, encoding structures, audio characteristics, facial consistency and provenance.

Section 86[18] is equally revealing. It permits a presumption concerning the integrity of a "secure electronic record" in the specified circumstances, but the statutory structure does not create a general presumption that every electronic record is authentic merely because it exists in digital form. This caution is particularly appropriate in an era of synthetic media. Section 90[19] concerning electronic messages, similarly allows the court to presume correspondence between the message received and the message fed into a computer for transmission, but does not permit a presumption regarding the person by whom the message was sent. These differentiated presumptions show that the BSA already recognises the need to distinguish technological transmission from human attribution.

[16] *Id*. s. 39
[17] The Information Technology Act, 2000 (Act 21 of 2000), s. 79A
[18] The Bharatiya Sakshya Adhiniyam, 2023 (Act 47 of 2023), s. 86
[19] *Id*. s. 90

The transitional provision contained in Section 170 is also relevant to contemporary case law. Although the BSA came into force on 1 July 2024, proceedings that were already pending immediately before commencement continue to be governed by the Indian Evidence Act, 1872. This explains why decisions delivered after July 2024 may still analyse Section 65B rather than Section 63 of the BSA. Such decisions should not automatically be dismissed as obsolete. Much of the Supreme Court's electronic evidence jurisprudence concerns underlying principles of authenticity, reliability and certification that continue to illuminate the BSA, even though the new legislation changes statutory details.

The overall architecture is therefore technologically progressive but conceptually incomplete for deepfakes. The BSA answers the question "when and how may an electronic record be legally proved?" more comprehensively than the nineteenth-century statute it replaced. Deepfakes add another question: "what must a court require before concluding that the reality portrayed by the record actually occurred?" That second question requires development through judicial doctrine, forensic standards and procedural rules.

4. Section 63, Authenticity and Chain of Custody: Why Admissibility Is Not Truth

Section 63 occupies a central position in any contemporary discussion of digital evidence because it establishes the statutory route through which computer outputs are proved. Yet one of the most important distinctions in the deepfake era is the distinction between admissibility and evidentiary weight. Meeting a technical condition that allows an electronic record to be received by the court does not compel the court to believe the content of that record.[20] The proposition is familiar in ordinary evidence law: a document may be admissible yet unreliable, a witness may be competent yet untruthful, and a photograph may be relevant yet misleading. Artificial intelligence makes the distinction exceptionally important because a synthetic file may satisfy every technical requirement concerning copying and preservation.

The Supreme Court's electronic evidence jurisprudence under the earlier Act laid the foundation for this understanding. In *Anvar P.V. v. P.K. Basheer*[21], the Court treated Section 65B as a special provision governing electronic records and emphasised compliance with the statutory certification requirements where secondary electronic evidence was sought to be

[20] *Supra* Note 4
[21] *Anvar P.V. v. P.K. Basheer*, (2014) 10 SCC 473

produced. The decision rejected the suggestion that ordinary secondary evidence provisions could simply bypass the specialised electronic evidence regime. In *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*[22], a Constitution Bench reaffirmed the mandatory character of the certificate for computer outputs governed by Section 65B(4), while clarifying procedural mechanisms through which a party unable to obtain a certificate from a third-party custodian could seek the court's assistance. The conceptual legacy of these decisions remains significant because Section 63 of the BSA substantially carries forward the specialised statutory treatment of computer outputs.

What these cases did not decide, because the technological environment was different, was whether a statutory certificate proves that the content was not synthetically generated. It does not. A certificate is evidence concerning the production of the electronic output and operation of the relevant system. A deepfake allegation concerns the representational truth of the content. The former asks whether the copy faithfully reproduces a particular digital file; the latter asks whether the digital file faithfully reproduces reality. Conflating these questions would create what may be described as a "certification fallacy": assuming that because a record is formally admissible under Section 63, its substantive authenticity has also been established.

A recent and particularly useful illustration is *N. Mahesh v. G.N. Nanjundaswamy*[23]. Applying the BSA, the court scrutinised electronic material rather than treating its formal tender as conclusive. It noted deficiencies relating to device particulars, metadata, browser or application records, the manner of transfer, chain of custody, platform evidence and independent technical support. The decision reflects an important emerging proposition: admission of an electronic exhibit is not synonymous with proof of authenticity or probative value. A Section 63 certificate cannot operate as a magic formula where the surrounding evidentiary foundation remains unreliable.

A similar insistence upon statutory compliance appears in *Rajesh Kumar v. State of Haryana and Others*[24]. The High Court treated Section 63 of the BSA as substantially corresponding to the earlier Section 65B framework and reiterated the importance of certification for copied forms of electronic material such as CCTV footage and other computer outputs. The case is

[22] *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, (2020) 7 SCC 1
[23] *N. Mahesh v. G.N. Nanjundaswamy*, C.C. No. 29859/2025
[24] *Rajesh Kumar v. State of Haryana and Others*, CRR No. 1642 of 2026 (O&M)

important not because it involved a sophisticated deepfake dispute, but because it confirms that post-BSA courts continue to insist upon a structured legal foundation for electronic evidence rather than permitting mere marking of a digital exhibit to substitute for proof.

In a genuine deepfake dispute, chain of custody should begin as close as possible to acquisition. Investigators should identify the earliest available version, record who obtained it, preserve the native file without unnecessary conversion, calculate and record cryptographic hashes, create forensic working copies, document each transfer and preserve the source device whenever lawful and proportionate.[25] If the material originated online, investigators should seek platform-side records where available, including upload timestamps, account information, server copies and relevant metadata. Where content was received through messaging applications, the difference between the sender's source file, the recipient's downloaded copy and subsequent forwarded versions should be preserved rather than collapsed into a single "video".

This matters because digital content changes through ordinary technical processes. Messaging services may compress video, strip metadata or rename files. A social media platform may transcode an uploaded video into several resolutions. Screenshots and screen recordings create new files whose metadata relates to the act of capturing the screen rather than the underlying content. An investigator who repeatedly exports, converts or forwards a file can unintentionally destroy clues needed for later authenticity analysis.[26] Section 63's emphasis on device particulars and production methodology should therefore be understood as part of a broader forensic chain rather than as post hoc paperwork prepared when litigation begins.

The requirement of hashing is valuable but should similarly be understood within its technical limits. A cryptographic hash is analogous to a highly sensitive digital fingerprint. If a file changes, the hash normally changes. Matching hashes between acquisition and trial can therefore establish that the exhibit remained unchanged during that period. But the evidentiary significance begins only at the moment of hashing. If a manipulated recording entered police possession before the first hash was calculated, the hash faithfully preserves the manipulated

[25] Marie-Helen Maras and Alex Alexandrou, "Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos," 23(3) *The International Journal of Evidence & Proof* 255–262 (2019).
[26] Thanh Thi Nguyen et al., "Deep Learning for Deepfakes Creation and Detection: A Survey," 223 *Computer Vision and Image Understanding* 103525 (2022).

recording. For deepfake cases, courts should therefore ask not only "does the hash match?" but "when was the first hash created, from what source, by whom, and how close was that source to the alleged real-world event?"

Metadata should also be treated neither as infallible nor irrelevant. Creation timestamps, encoding information, camera model, GPS data, software tags and editing histories can strongly support or undermine authenticity, but metadata can be stripped or deliberately modified. The absence of metadata is not automatically proof of falsification because ordinary platform processing may remove it. Conversely, the presence of apparently consistent metadata is not conclusive because sophisticated manipulation can reproduce plausible values. Metadata should therefore operate as one component of a cumulative authentication enquiry.[27]

The evidentiary burden requires calibrated handling. The initial proponent should establish a basic foundation demonstrating the origin, possession and production of the media. A party opposing the evidence should not be able to obtain exclusion through speculative reference to "AI" alone. Where the challenger identifies credible reasons for doubting authenticity, however, the court should require heightened proof proportionate to the evidentiary significance of the record.[28] If the recording is merely corroborative of extensive independent evidence, a moderate foundation may suffice. If the prosecution seeks conviction substantially on the basis of a disputed video or cloned voice recording, rigorous forensic authentication should be indispensable.

This approach aligns with the larger principle that the criminal standard of proof concerns the reliability of the conclusion, not merely procedural admissibility. Section 63 is therefore best understood as the beginning of the deepfake enquiry rather than its end. It provides a statutory gateway and a mechanism for preserving technical accountability. The judicial task remains to determine whether the electronic representation can safely be connected to the underlying human conduct or event.

5. Forensic Expertise, Detection and Provenance Under Sections 39 and 86

Deepfake disputes will often present questions beyond ordinary judicial perception. A judge can compare visible features or listen to voices, but sophisticated synthetic media are

[27] Brian Dolhansky et al., "The DeepFake Detection Challenge (DFDC) Dataset," *arXiv:2006.07397* (2020).
[28] *Supra* Note 3

specifically designed to defeat human perception. Section 39(2) of the BSA accordingly assumes special importance because it recognises the opinion of an Examiner of Electronic Evidence under Section 79A of the Information Technology Act, 2000 as expert evidence where the court must form an opinion concerning information stored or transmitted electronically. The provision supplies the legal foundation for specialised technical assistance, but the quality of that assistance must itself be judicially evaluated.

An effective deepfake examination should ordinarily be multimodal. Visual analysis may examine frame-level artefacts, facial boundaries, inconsistent blinking, texture, lighting, shadows, reflections and physiological signals. Audio analysis may examine waveform discontinuities, spectral patterns, background noise, speaker characteristics and evidence of splicing or synthesis.[29] File analysis can examine containers, codecs, timestamps, quantisation patterns and software signatures. Metadata may identify the source device or editing software. Network or platform information can establish when and by whom content was uploaded. Where the media depicts a known location or event, contextual analysis may compare weather, shadows, physical geography, clothing, clocks, other cameras or witness accounts.[30] A reliable conclusion emerges from convergence among several evidentiary sources rather than a single detector score.

Expert evidence should therefore explain methodology rather than merely announce an outcome. A statement that "the video is 92 per cent likely to be AI-generated" is of limited forensic utility unless the court knows what the percentage means. The expert should explain the model used, the version of that model, the class of manipulations it can detect, the validation dataset, known error rates, whether the input was recompressed or degraded, whether comparable Indian-language or demographic material was part of validation, and what alternative explanations exist. The opposing party should have a meaningful opportunity to test those assumptions. Judicial acceptance of opaque AI output without methodological disclosure would merely replace one black box - the deepfake generator, with another - the deepfake detector.

[29]Yisroel Mirsky and Wenke Lee, "The Creation and Detection of Deepfakes: A Survey," 54(1) *ACM Computing Surveys* 1–41 (2021).
[30] Haiying Guan, James Horan and Andrew Zhang, *Guardians of Forensic Evidence: Evaluating Analytic Systems Against AI-Generated Deepfakes*, Forensics@NIST 2024 (National Institute of Standards and Technology, 2025).

The possibility of model bias is not theoretical. Audio detection may perform differently across languages, dialects and recording conditions. Indian courts may encounter Hindi, Bengali, Tamil, Telugu, Marathi, Punjabi, Urdu and multilingual conversations recorded through inexpensive devices and heavily compressed messaging platforms. Detection systems trained mainly on clean English-language datasets may not generalise reliably. The development of Indian datasets such as the IndiaAI-linked "IndicSynth" initiative, involving large-scale synthetic speech across multiple Indian languages, illustrates the growing recognition that deepfake detection requires locally representative data.[31]

The distinction drawn by Section 86 between secure electronic records and ordinary electronic records also offers an important normative clue. The BSA does not instruct courts to presume the authenticity of all digital content. Where a secure electronic record qualifies for statutory treatment, certain presumptions may arise; otherwise, the record must earn evidentiary confidence through ordinary proof. In deepfake litigation this means that courts should resist visual realism as a substitute for authentication. Modern synthesis can produce precisely the kinds of facial detail, natural movement and vocal familiarity that once reassured fact finders.[32]

Provenance technology may strengthen this framework. Cryptographically signed content credentials can record information concerning creation, editing and publication. The C2PA technical standard, for example, is designed to establish verifiable provenance and content history. Such credentials should not be elevated into an absolute presumption of truth, because they authenticate a content history rather than the truthfulness of the depicted scene.[33] A genuine camera can record a staged event; a verified account can upload a false statement. Nevertheless, provenance evidence can materially reduce uncertainty concerning whether a file came directly from a particular device, whether it passed through recognised editing software and whether later modifications occurred.

Forensic practice should therefore distinguish "content credentials" from "content truth". The former can prove technical history; the latter remains an evidentiary inference from the entire record. Courts should also recognise that absence of provenance credentials does not prove

[31] *Id.*
[32] Britt Paris and Joan Donovan, *Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence* 5–15 (Data & Society Research Institute, 2019).
[33] Deepfakes Analysis Unit, *Deepfakes Analysis Unit's Sixth Quarter Report* (Misinformation Combat Alliance, Jan. 19, 2026).

falsity, particularly because legacy devices and platforms may not implement emerging standards. During a transitional period, provenance systems are best treated as strengthening evidence rather than mandatory prerequisites.

Another issue concerns preservation of models and forensic reproducibility. If an expert reaches a conclusion through a proprietary detection system that is later updated, the defence or opposing party may be unable to replicate the result. Courts should therefore encourage documentation of the specific model version, parameters, input file hash and output generated at the time of examination. Where feasible, the tested forensic copy should be preserved along with logs.[34] This principle mirrors traditional forensic science, where another competent examiner should be able to evaluate the methodology and, where possible, reproduce the analysis.

Human expertise remains particularly important because synthetic content may be technically sophisticated yet contextually impossible. A voice may match the apparent speaker while the wording contradicts demonstrable contemporaneous facts. A video may depict a politician at one location when verified travel records place the person elsewhere. A purported CCTV recording may contain accurate visual synthesis but lack any corresponding DVR entry.[35] Conversely, a detector may flag authentic content because platform compression created unusual artefacts. Forensic technology must therefore function within the judicial process rather than replacing it.

Ultimately, Sections 39 and 86 permit Indian courts to develop a sophisticated authenticity jurisprudence without waiting for the BSA to be amended each time AI changes. Section 39 supplies access to expertise; Section 86 demonstrates legislative caution regarding presumptions. The doctrinal opportunity is to combine these provisions with Section 63 and ordinary principles of relevance and proof to create a layered method in which technical integrity, source attribution, forensic analysis and contextual corroboration jointly determine evidentiary reliability.

[34] *Supra* Note 10
[35] Hany Farid, *Photo Forensics* 1–20, 187–216 (The MIT Press, Cambridge, 2016).

6. Judicial Evolution from Tape Recordings to Synthetic Media

Indian courts began confronting technologically recorded evidence decades before artificial intelligence entered ordinary public life. Those earlier cases remain valuable because they developed authentication principles capable of adaptation to synthetic media. In *Yusufalli Esmail Nagree v. State of Maharashtra*[36], the Supreme Court accepted the potential evidentiary value of tape-recorded conversations while recognising the danger arising from alteration and erasure. The Court's focus upon accurate recording, identification and reliability anticipated the central dilemma of digital evidence: technology may preserve powerful evidence, but its persuasive appearance cannot eliminate the need for authentication.

In *R.M. Malkani v. State of Maharashtra*[37], the Supreme Court further recognised tape-recorded conversations as admissible where the material was relevant, the voice was properly identified, and the accuracy of the recording was established. The concern that the possibility of tampering must be adequately excluded remains directly applicable to synthetic audio. A cloned voice may sound more convincing than the altered tapes contemplated by twentieth-century courts, but the underlying legal principle is identical: resemblance cannot substitute for proof of origin and integrity.

The authentication requirements were developed more systematically in *Ram Singh v. Col. Ram Singh*[38], where the Court stressed identification of the speaker, accuracy of the recording, exclusion of possibilities of tampering, safe custody and audibility. These requirements are strikingly contemporary when translated into digital terms. "Safe custody" corresponds with chain of custody and hash verification; "accuracy" requires examination of native digital files and source systems; "identification" increasingly requires more than a witness saying "that sounds like him" where voice-cloning tools can convincingly imitate familiar speakers.

In *Tukaram S. Dighole v. Manikrao Shivaji Kokate*[39], the Supreme Court acknowledged both the evidentiary utility of recorded material and its susceptibility to manipulation, emphasising the importance of satisfactory proof of authenticity. The observation assumes greater force today. Traditional editing often left discontinuities detectable through careful examination;

[36] *Yusufalli Esmail Nagree v. State of Maharashtra*, AIR 1968 SC 147
[37] *R.M. Malkani v. State of Maharashtra*, (1973) 1 SCC 471
[38] *Ram Singh v. Col. Ram Singh*, 1985 Supp SCC 611
[39] *Tukaram S. Dighole v. Manikrao Shivaji Kokate*, (2010) 4 SCC 329

generative models may create internally coherent material without an original recording ever existing. Courts should therefore understand earlier warnings about tampering as minimum safeguards rather than exhaustive standards.

The statutory jurisprudence concerning electronic records developed significantly in *State (NCT of Delhi) v. Navjot Sandhu*[40], where the Supreme Court had permitted electronic material to be proved through general principles of secondary evidence despite non-compliance with Section 65B in certain circumstances. That approach was substantially displaced by *Anvar P.V. v. P.K. Basheer*[41], which treated the specialised provisions governing electronic records as controlling and emphasised the certificate required for computer outputs. *Anvar* recognised source and authenticity as important concerns in electronic evidence and marked a doctrinal move towards greater technical discipline.

In *Tomaso Bruno v. State of Uttar Pradesh*[42], the Supreme Court emphasised the significance of scientific and electronic evidence and addressed the consequences of withholding potentially important CCTV evidence. The case remains relevant in the deepfake era for a different reason: authenticity enquiries should not be confined to the disputed clip itself when more reliable source material exists. Where the original CCTV system, DVR, server copy or platform data is available but not produced, the court should ask why a litigant relies upon a derivative recording rather than the superior source.

Procedural objections were considered in *Sonu @ Amar v. State of Haryana*[43], where the Supreme Court distinguished objections relating to the mode or method of proof and emphasised the importance of raising curable objections when the evidence is tendered. Deepfake challenges, however, may involve more than a formal defect in mode of proof. If the allegation is that the content itself is synthetic, the objection goes to authenticity and probative reliability, and in a criminal case may directly affect whether guilt has been established beyond reasonable doubt.

The temporary doctrinal uncertainty created by *Shafhi Mohammad v. State of Himachal*

[40] *State (NCT of Delhi) v. Navjot Sandhu*, (2005) 11 SCC 600
[41] *Anvar P.V. v. P.K. Basheer*, (2014) 10 SCC 473
[42] *Tomaso Bruno v. State of Uttar Pradesh*, (2015) 7 SCC 178
[43] *Sonu @ Amar v. State of Haryana*, (2017) 8 SCC 570

Pradesh[44], which relaxed certification where the party did not possess the relevant device, was resolved by the Constitution Bench in *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*[45]. *Arjun Panditrao* reaffirmed the mandatory statutory certification framework for secondary electronic evidence governed by Section 65B while recognising that courts can assist parties in obtaining certification from third-party custodians. The principle is particularly useful for platform-hosted deepfake disputes because the relevant server or account records may be controlled by intermediaries rather than the litigating party.

The character of electronic storage as documentary evidence received further attention in *P. Gopalkrishnan @ Dileep v. State of Kerala*[46], where the Supreme Court treated a memory card containing digital material as a document and sought to balance fair-trial access with privacy and victim-related concerns. Deepfake litigation will require comparable balancing. Defence access to forensic copies, metadata and expert material may be necessary to challenge authenticity, yet unrestricted circulation of intimate, sexual or otherwise sensitive media can itself cause severe harm. Courts will need controlled-access procedures capable of preserving both adversarial fairness and privacy.

Recent judicial developments concerning recordings also remain relevant. In *Vibhor Garg v. Neha*[47], the Supreme Court dealt with secretly recorded material in matrimonial proceedings and recognised that the relevance of a recording cannot automatically be defeated by a general privacy objection, although questions of reliability and evidentiary assessment remain. In the age of synthetic media, the case illustrates why admissibility, privacy and authenticity must remain separate inquiries. A recording can be privately obtained yet authentic, or consensually obtained yet artificially manipulated.

In *Himanshu Chordia v. State of Rajasthan*[48], the Supreme Court again confronted evidentiary issues involving electronic material in a proceeding governed by the older Act because of transitional rules. The Court's discussion reflected continuing concern with the legal treatment of new forms of privately generated investigative material and the need for legal mechanisms capable of addressing evolving evidentiary circumstances. Such observations are significant

[44] *Shafhi Mohammad v. State of Himachal Pradesh*, (2018) 2 SCC 801
[45] *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, (2020) 7 SCC 1
[46] *P. Gopalkrishnan @ Dileep v. State of Kerala*, (2020) 9 SCC 161
[47] *Vibhor Garg v. Neha*, 2025 INSC 829
[48] *Himanshu Chordia v. State of Rajasthan*, 2026 INSC 778

because artificial intelligence increasingly enables private actors not only to gather evidence but to manufacture apparently evidentiary material at scale.

The most significant post-BSA development is *N. Mahesh v. G.N. Nanjundaswamy*[49]. The court's scrutiny of missing source-device details, transfer methodology, metadata, technical corroboration and chain of custody demonstrates the direction in which deepfake jurisprudence should evolve. The central lesson is that electronic evidence must "inspire confidence" through its surrounding technological foundation. Formal certification cannot cure an absence of reliable provenance. Similarly, *Rajesh Kumar v. State of Haryana and Others*[50], confirms the continuing importance of statutory certification under the new evidence regime. Together, these decisions indicate that Indian courts are beginning to confront the BSA not merely as a renumbered Evidence Act but as a framework requiring technologically informed judicial scrutiny.

Deepfakes have also entered personality-rights litigation, providing courts with direct exposure to synthetic impersonation. In *Anil Kapoor v. Simply Life India & Ors.*[51], the Delhi High Court protected the actor's name, image, voice and other indicia of personality against unauthorised commercial and technological exploitation, including misuse facilitated through artificial intelligence. The case is important because it recognises that synthetic reproduction of a person's identity can cause legally cognisable harm even where the dispute is not about evidentiary admissibility.

The Bombay High Court followed a similar trajectory in *Arijit Singh v. Codible Ventures LLP & Ors.*[52], where the unauthorised use of AI technologies capable of imitating or converting voices formed part of the personality-rights dispute. The decision is especially relevant to evidentiary law because voice cloning undermines traditional assumptions that familiar speech can be authenticated simply by a listener recognising the speaker. Likewise, the Delhi High Court's decision in *Jaikishan Kakubhai Saraf v. Peppy Store*[53], forms part of the growing judicial response to unauthorised exploitation of celebrity persona through digital technologies.

[49] *N. Mahesh v. G.N. Nanjundaswamy*, C.C. No. 29859/2025 (2026)
[50] *Rajesh Kumar v. State of Haryana and Others*, CRR No. 1642 of 2026
[51] *Anil Kapoor v. Simply Life India & Ors.*, 2023 SCC OnLine Del 6914
[52] *Arijit Singh v. Codible Ventures LLP & Ors.*, 2024 SCC OnLine Bom 2445
[53] *Jaikishan Kakubhai Saraf v. Peppy Store*, 2024 SCC OnLine Del 3664

A further development occurred in *Aishwarya Rai Bachchan v. Aishwaryaworld.com*[54], where the Delhi High Court dealt with unauthorised digital and deepfake uses of the actor's persona. Significantly from the perspective of evidence law, the proceedings also reflected the practical importance of Section 63 certification for electronic material placed before the court. In *Jaya Bachchan v. Bollywood Bubble Television & Ors*[55], the Court similarly restrained misuse involving AI-generated or manipulated representations of personality.

The trend continued in *Akkineni Naga Chaitanya v. www.sexvid.xxx*[56], where the Delhi High Court addressed unauthorised exploitation of persona, including AI/deepfake-associated misuse in a highly invasive context. Although personality-rights orders do not themselves establish criminal evidentiary standards, they demonstrate a rapidly developing judicial awareness that faces, voices and apparent conduct can be technologically replicated. That recognition should influence evidentiary common sense. A witness's assertion that "I recognise this face" or "I know this voice" remains relevant, but in a genuinely contested synthetic-media case it cannot always be the end of authentication.

The judicial trajectory from *Yusufalli Esmail Nagree* to the 2026 cases reveals remarkable doctrinal continuity. Courts have consistently required relevance, identification, reliability and protection against tampering. What has changed is the magnitude of technological deception. The modern task is not to discard earlier jurisprudence but to reinterpret its safeguards for an environment in which manipulation may occur before any conventional "recording" exists.

7. The Wider Indian Regulatory Framework and the 2026 Synthetic-Media Rules

The law governing deepfakes in India cannot be understood through the BSA alone. Evidence law determines how material is proved in judicial proceedings, while the Information Technology Act, intermediary rules, criminal law, intellectual property principles, privacy doctrine, personality rights and election regulation govern the creation, dissemination and consequences of synthetic media. The most significant recent development is the formal incorporation of "synthetically generated information" into India's intermediary-governance framework in 2026.

[54] *Aishwarya Rai Bachchan v. Aishwaryaworld.com*, 2025 SCC OnLine Del 5943
[55] *Jaya Bachchan v. Bollywood Bubble Television & Ors.*, CS(COMM) 1194/2025
[56] *Akkineni Naga Chaitanya v. www.sexvid.xxx*, 2026 SCC OnLine Del 4619

On 10 February 2026, the Central Government notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 through G.S.R. 120(E)[57], with the relevant amendments taking effect from 20 February 2026. The amended framework defines "synthetically generated information" in substance as audio, visual or audio-visual information artificially or algorithmically created, generated, modified or altered through a computer resource in a manner that appears real, authentic or true and depicts an individual or event in a manner likely to be indistinguishable from a natural person or real-world occurrence. Importantly, the definition excludes various routine good-faith technical modifications, such as certain forms of formatting, enhancement, noise reduction, transcription and compression, where they do not materially falsify the content.

This distinction mirrors the evidentiary analysis required under the BSA. Not every modified digital file is a deepfake. Courts should distinguish ordinary processing from material synthetic alteration affecting the proposition asserted by the evidence. A video whose brightness has been adjusted for visibility is fundamentally different from one in which a person's face has been replaced or words have been generated in the person's cloned voice. The 2026 definition provides useful regulatory vocabulary even though evidentiary admissibility remains governed by the BSA.

The amended Rules impose due-diligence obligations relating to synthetic media. Among other matters, intermediaries providing technological resources capable of generating or modifying synthetic information are required to adopt reasonable technical measures addressing specified unlawful uses.[58] The regulatory framework also requires prominent disclosure or labelling for relevant synthetic content and, to the extent technically feasible, the embedding of persistent metadata or provenance-related information, including identifiers associated with the generating intermediary. Measures are also directed against removal or suppression of required labels or provenance information.

These obligations are highly significant for evidence law. A provenance marker embedded when content is generated can later become evidence concerning origin. A platform label

[57] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E), Gazette of India, Extraordinary, Part II, sec. 3(i), Feb. 10, 2026.
[58] Ministry of Electronics and Information Technology, Government of India, *Frequently Asked Questions: Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026* (2026).

indicating that an audio-visual file was synthetically generated does not automatically determine every legal question, but it can provide an evidentiary lead. Conversely, intentional removal of provenance information may become relevant to consciousness of manipulation, depending upon the circumstances. The most productive legal development would therefore be institutional interoperability between the regulatory framework governing synthetic media and the BSA framework governing proof.

The intermediary rules also contain preservation requirements. Records associated with content removed or disabled by an intermediary may be required to be preserved for the stipulated period, including circumstances where longer retention is legally required. This has obvious evidentiary value because disputed deepfake content may disappear rapidly after takedown. Preservation of platform-side records can permit courts to reconstruct upload history, account association, timestamps and technical metadata even where the public post itself is no longer available.

Election regulation provides another important field of development. The Election Commission of India has issued directions and advisories requiring political parties to responsibly use artificial intelligence and to clearly label specified AI-generated or synthetically altered campaigning content. In January 2025, the Commission called for prominent labels such as "AI-Generated", "Digitally Enhanced" or "Synthetic Content" in appropriate political material.[59] The Commission has subsequently reiterated concerns about misleading AI-generated or manipulated campaigning material and rapid action against unlawful content, including during the 2026 electoral cycle.

The election context illustrates why deepfake regulation cannot be reduced to subsequent courtroom proof. Synthetic material can cause irreversible consequences before adjudication. A fabricated speech released hours before voting may influence public behaviour even if a forensic report proves it false days later. Nevertheless, provenance and labelling rules created for preventive governance can later assist judicial fact-finding. Election authorities, intermediaries and investigating agencies should therefore preserve not only copies of impugned content but associated provenance and account records.

[59] Election Commission of India, *Advisory for Labelling Synthetic/AI-Generated Content Used by Political Parties for Election Campaigning*, Jan. 16, 2025.

Personality rights jurisprudence supplies a complementary civil law response. *Anil Kapoor v. Simply Life India & Ors*[60], *Arijit Singh v. Codible Ventures LLP & Ors.*[61], *Aishwarya Rai Bachchan v. Aishwaryaworld.com*[62], and *Akkineni Naga Chaitanya v. www.sexvid.xxx*[63], demonstrate growing judicial willingness to restrain the unauthorised technological appropriation of a person's name, image, voice and likeness. Their significance extends beyond celebrities. The underlying harm - false technological attribution of identity - is the same harm that can transform synthetic media into false evidence.

The criminal law implications are likewise distributed across different offences. Depending upon facts, deepfake conduct can engage provisions concerning cheating, personation, intimidation, sexual exploitation, obscenity, forgery-related conduct, defamation, privacy violations and cyber offences. Synthetic intimate imagery can cause particularly acute harm because its falsity does not necessarily reduce reputational or psychological consequences. Yet when such material becomes evidence, courts must avoid an ironic secondary harm: a victim should not be required to prove falsity by allowing unrestricted circulation of intimate synthetic material among multiple actors. Controlled forensic access and sealed procedures may therefore be essential.

India's emerging framework is moving from a reactive "remove harmful deepfakes" approach toward a provenance-oriented model. The February 2026 amendments are important because they recognise labelling and embedded metadata rather than relying solely upon after-the-fact detection. Detection asks whether a finished object looks synthetic; provenance asks where it came from and what happened to it. The latter is often more stable as a legal foundation.

Still, the regulatory and evidentiary systems are not yet completely integrated. The BSA does not expressly specify the evidentiary consequences of synthetic-media labels, missing provenance metadata, platform-generated authenticity credentials or evidence that a required label was removed. Nor does it create a specialised procedure for emergency preservation of suspected deepfake evidence. These gaps do not make the existing framework ineffective, but

[60] *Anil Kapoor v. Simply Life India & Ors.*, 2023 SCC OnLine Del 6914
[61] *Arijit Singh v. Codible Ventures LLP & Ors.*, 2024 SCC OnLine Bom 2445
[62] *Aishwarya Rai Bachchan v. Aishwaryaworld.com*, 2025 SCC OnLine Del 5943
[63] *Akkineni Naga Chaitanya v. www.sexvid.xxx*, 2026 SCC OnLine Del 4619

they indicate where future rules, judicial guidelines or legislative amendments can improve coherence.

8. Constitutional Fair-Trial Risks and Comparative Lessons for Deepfake Adjudication

Deepfake evidence engages constitutional values because erroneous authentication can affect liberty, dignity, privacy, reputation and the fairness of adjudication. Article 21 of the Constitution protects life and personal liberty through a procedure that must satisfy constitutional standards of fairness. In *K.S. Puttaswamy (Retd.) v. Union of India*[64], the Supreme Court recognised privacy as a constitutionally protected right connected with dignity and autonomy. Synthetic media can threaten that right in an unusual manner: they do not merely disclose genuine private information but can manufacture apparently private conduct that never occurred.

Fair trial concerns operate for both sides. An accused confronted with a sophisticated synthetic video must receive a meaningful opportunity to challenge the technology, inspect forensic material and test expert methodology. If the State relies upon an AI-generated probability score without disclosure of the underlying basis, adversarial testing becomes difficult. At the same time, complainants and victims have a legitimate interest in preventing unnecessary reproduction and dissemination of intimate or humiliating recordings. Courts should therefore develop controlled-access mechanisms, forensic clones, confidentiality orders and restrictions on copying where required. The principle reflected in *P. Gopalkrishnan @ Dileep v. State of Kerala*[65], concerning balancing access to digital evidence with sensitive privacy interests, provides a useful doctrinal starting point.

The liar's dividend also has constitutional consequences. If every prosecution based upon electronic material faces a vague claim that "AI could have made this", fact-finding may become paralysed. Reasonable doubt is not equivalent to imaginable doubt. Courts must insist upon a rational evidentiary basis for synthetic-media objections. At the same time, because criminal liability requires proof beyond reasonable doubt, a credible forensic indication of manipulation cannot be brushed aside merely because the recording looks realistic. A calibrated

[64] *K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1
[65] *P. Gopalkrishnan @ Dileep v. State of Kerala*, (2020) 9 SCC 161

burden mechanism is therefore preferable to either extreme.[66]

Comparative developments support an emphasis on transparency and provenance. Article 50 of the European Union Artificial Intelligence Act imposes transparency requirements concerning certain AI-generated and manipulated content. Providers of relevant AI systems must facilitate detection of synthetically generated outputs through machine-readable formats, while deployers of deepfake systems are subject to disclosure duties concerning artificially generated or manipulated content, subject to specified qualifications. The relevant transparency obligations become applicable within the EU Act's staged implementation framework from 2026.

The European approach is instructive not because India should reproduce it mechanically, but because it recognises that provenance should be created at the time of generation rather than reconstructed only during litigation. India's February 2026 synthetic-information rules move in a broadly compatible direction by emphasising labels and persistent metadata. The evidentiary opportunity is to treat these regulatory records as components of an authentication ecosystem.[67]

Technical standardisation can also reduce judicial inconsistency. NIST and other scientific bodies have increasingly examined forensic methods for AI-generated media, reflecting the wider international recognition that deepfake detection must be evaluated systematically rather than through intuitive visual assessment. Indian forensic institutions would benefit from published validation standards covering minimum datasets, laboratory procedures, error-rate reporting, documentation and peer review. A court should be able to ask not merely whether an expert is qualified, but whether the particular forensic method used in the case meets accepted scientific standards.

The constitutional principle of equality also counsels against assuming that all litigants possess equivalent technological resources. Wealthy parties may commission several private forensic reports while ordinary litigants depend upon overloaded public laboratories. If sophisticated

[66] Jan Kietzmann, Linda W. Lee, Ian P. McCarthy and Tim C. Kietzmann, "Deepfakes: Trick or Treat?," 63(2) *Business Horizons* 135–146 (2020).
[67] Vijaykumar Meti, "Practical Impact and Challenges of Electronic Evidence under the Bharatiya Sakshya Adhiniyam in 2025: A Study," 5(10) *Indian Journal of Legal Review* 492–505 (2025).

authentication becomes essential to justice, forensic capacity must be treated as judicial infrastructure rather than a luxury service. Government laboratories, Section 79A examiners and independent accredited experts will require specialised training and computational resources. Legal aid may need to include access to digital forensic assistance in serious cases.[68]

Judicial training is equally important. The challenge is not that judges must become machine-learning engineers. Rather, judges must understand the questions that should be asked: Was the native file examined? What is the earliest known copy? Is the hash traceable to acquisition? Was metadata preserved? Does the detector have known limitations? Was the purported speaker independently located at the scene? Does another camera corroborate the event? Was the content originally labelled as synthetic? Does platform data support the account attribution? These questions translate complex technology into ordinary evidentiary reasoning.

A particularly important comparative lesson is that provenance should complement, not replace, traditional corroboration. Even the most advanced cryptographic credential cannot determine whether a recorded statement was coerced, sarcastic, staged or taken out of context. Likewise, a deepfake detector cannot establish mens rea or legal causation.[69] The objective of a modern evidentiary framework should therefore be modest but vital: determine with reasonable technical confidence whether the digital object has the origin and representational relationship claimed for it. The ordinary law of evidence must then determine what legal inference follows.

Deepfake adjudication therefore sits at the intersection of technology and constitutional procedure. The appropriate response is neither a presumption against digital evidence nor technological enthusiasm. Courts should demand transparency proportionate to the stakes. Where life, liberty, reputation or intimate privacy may turn upon an electronic recording, authenticity must be demonstrated through a process capable of being understood, challenged and independently evaluated.

[68] Chanchal Meena and Pooja Jain, "Admissibility of Electronic Evidence under the Bharatiya Sakshya Adhiniyam, 2023: A Critical Analysis of Evidentiary Standards and Judicial Challenges," 11(7) *International Journal of Research and Innovation in Applied Science* 1128–1134 (2026).
[69] Cristian Vaccari and Andrew Chadwick, "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News," 6(1) *Social Media + Society* 1–13 (2020).

9. A Deepfake-Specific Evidentiary Protocol for India: Reform Proposals and Conclusion

The BSA already contains sufficient doctrinal building blocks for Indian courts to begin constructing a deepfake-specific authentication protocol. What is required is an integrated procedure linking Sections 39, 57, 61–63 and 86 with digital forensic standards, intermediary preservation duties and emerging provenance mechanisms. The objective should not be to create a new exclusionary rule for every AI-generated object. Rather, the law should establish a structured enquiry capable of determining whether a disputed electronic record reliably represents the fact for which it is offered.

The first element should be preservation at the earliest possible point. Investigating agencies and litigants should preserve the native or earliest obtainable version of the disputed media, document the acquisition source and calculate a cryptographic hash immediately. Where the source is a smartphone, camera, CCTV system, cloud account or social media platform, relevant device and account information should be recorded. Investigators should avoid unnecessary conversion, editing, screen recording or forwarding. A forensic working copy can be created while the preserved original remains untouched.

The second element should be provenance. Section 63 certificates should not be completed as generic forms stating merely that "a computer was working properly". The specific source device, relevant identifier, method of extraction, date and time of acquisition, software used, hash values and transfer process should be accurately recorded. The BSA Schedule's demand for device identifiers and hash information provides a strong statutory basis for this approach. Where an intermediary or platform possesses upload records, server copies or synthetic-content metadata, courts should facilitate timely preservation and production subject to lawful process.

The third element should be a meaningful threshold for deepfake objections. A party alleging synthetic manipulation should ordinarily identify the basis of the challenge rather than rely upon technological possibility alone. Relevant grounds could include unexplained breaks in custody, absence of a purported original, inconsistent metadata, forensic anomalies, contradictory location evidence, platform records indicating synthetic generation, evidence that the account was compromised or expert analysis raising a genuine question. This prevents the liar's dividend from transforming every electronic case into unlimited forensic litigation.

Once a credible challenge is established, the fourth element should be enhanced authentication

by the proponent. The exact level should depend upon the centrality of the evidence. If a disputed clip is cumulative and supported by multiple independent witnesses, less extensive technical examination may be sufficient. If the digital object is the decisive evidence linking the accused to an offence, stronger authentication should be expected. This may include native-file examination, source-device inspection, platform records, expert evidence, metadata consistency, cryptographic chain of custody and independent circumstantial corroboration.

The fifth element should concern expert methodology. Examiners should disclose what they analysed, which tools and model versions they used, the hash of the tested file, the relevant reliability limitations, and whether the finding was based on a single automated detector or multiple forensic techniques. Where AI-assisted detection is used, an unexplained numerical output should not itself be treated as decisive expert opinion. The court must be able to evaluate the reasoning connecting the data to the conclusion.

The sixth element should integrate synthetic-media provenance created under the 2026 IT Rules. Where content carries a statutory label or persistent provenance metadata indicating synthetic generation, that information should be preserved as part of the evidentiary record. Where such metadata appears to have been deliberately removed, the circumstances surrounding removal may be relevant, though absence alone should not create an automatic adverse presumption. Legacy content and ordinary technological processes can lack provenance information for innocent reasons.

The seventh element should preserve human and contextual corroboration. Deepfake adjudication should never become entirely detector-centric. Courts should ask whether witnesses saw the event, whether independent CCTV footage exists, whether mobile location data accords with the recording, whether other communications reference the incident, whether background details are consistent with the claimed time and place, and whether the apparent speaker's conduct before and after the event is consistent with the recording. A convincing evidentiary mosaic is substantially more resistant to synthetic fabrication than a single isolated media file.

The eighth element should involve procedural safeguards for sensitive material. Where the disputed content is intimate, sexually explicit, medical, confidential or otherwise dignity-sensitive, forensic examination should occur through controlled copies and limited-access

mechanisms. Courts should prevent the authentication process itself from becoming a means of further dissemination. At the same time, the accused or opposing party must retain a genuine opportunity to challenge the evidence through an expert of choice where fairness requires it.

The ninth element should be institutional development. Standard operating procedures for deepfake evidence should be prepared for police agencies, prosecutors, forensic laboratories and courts. Section 79A electronic evidence examiners should receive specialised capacity in synthetic audio, video and image analysis. Laboratory accreditation should include validation of tools and preservation of reproducibility records. Judicial academies should incorporate synthetic-media evidence into continuing education. The BSA's new certificate Schedule should be supported by practical instructions explaining device identification, hashing, cloud acquisition and preservation of metadata.

Legislative amendment may eventually be desirable, but the immediate need is for authoritative judicial or executive guidelines. A specific rule could clarify that compliance with Section 63 establishes the statutory foundation for the relevant computer output but does not by itself establish the substantive authenticity of the depicted content when synthetic generation is genuinely disputed. Such clarification would eliminate the dangerous tendency to treat certification as synonymous with truth. A complementary rule could empower courts to direct preservation and production of platform-side provenance information at an early stage.

Indian courts should also expressly recognise three separate concepts: integrity, attribution and authenticity. Integrity asks whether the exhibit has changed after acquisition. Attribution asks whether the file came from the person, account, system or device alleged. Authenticity asks whether the content truthfully represents the purported event. Deepfakes demonstrate why these propositions cannot be merged. A hash may establish integrity without attribution; platform records may establish attribution without substantive authenticity; and eyewitness evidence may strongly corroborate authenticity even where the electronic file has undergone benign technical conversion.

The developing jurisprudence already points toward this distinction. *Anvar P.V. v. P.K. Basheer*, and *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, created procedural discipline around computer outputs. *Ram Singh v. Col. Ram Singh*, and *Tukaram S. Dighole v. Manikrao Shivaji Kokate*, emphasised accuracy and protection against tampering. *N. Mahesh*

v. G.N. Nanjundaswamy, illustrates the need to look beyond formal admission to source, metadata, custody and technical reliability. The growing personality rights cases illustrate judicial awareness that artificial intelligence can convincingly appropriate identity. Together, these strands can be developed into a coherent jurisprudence of synthetic-media authentication.

The BSA should therefore be viewed neither as inadequate merely because it does not repeatedly use the word "deepfake", nor as complete merely because it recognises electronic evidence. Its greatest strength lies in technological neutrality. Sections 39, 57 and 61–63 are sufficiently flexible to accommodate changing forms of digital proof. Its principal limitation lies in the absence of an express distinction between reliable reproduction of a digital file and reliability of the reality represented within that file. That distinction must now be developed doctrinally.

As of October 2026, India is moving toward a more coherent synthetic-media framework. The BSA has modernised evidentiary treatment of digital records; the 2026 intermediary amendments have formally addressed synthetically generated information, labels and provenance; election authorities have introduced AI-content disclosure expectations; and courts are increasingly protecting identity against deepfake misuse. The next phase must connect these developments.

The traditional expression "seeing is believing" reflected an era in which fabricating convincing audio-visual reality was difficult, expensive and unusual. Generative AI has reversed that assumption. Yet the appropriate legal response is not to replace belief with universal disbelief. Digital evidence remains indispensable to contemporary justice. CCTV recordings, smartphones, call data, platform communications and electronic transactions often provide evidence more precise than human recollection. The challenge is to distinguish trustworthy digital evidence from manufactured appearance.

The central evidentiary principle for the AI era should therefore be: the more convincingly technology can imitate reality, the more carefully law must establish provenance. A court should believe a digital recording not because the face looks familiar, the voice sounds convincing or a Section 63 certificate is attached, but because the entire evidentiary chain - from creation or acquisition through preservation, forensic examination and corroboration, provides a rational basis for concluding that the recording reliably represents the event alleged.

The Bharatiya Sakshya Adhiniyam, 2023 provides the foundation for such an approach. Judicial interpretation, forensic capacity and provenance-based regulation must now complete the architecture.