The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

FABRICATED REALITY: DEEPFAKES AND CRISES OF DIGITAL EVIDENCE

By Sai Sahasra Sarvadevabhatla

ABSTRACT:

In this era, technology is taking reins in every domain. It is the driving force of development that is influencing the society in more ways than one. As we all are well aware every aspect in life has both advantages and disadvantages. We weigh these pros and cons to determine if it's worth pursuing and then we try to brainstorm for ideas and formulate plans and safeguards that will attenuate the negative impacts.

In this context, even technology has drawbacks along with advantages. Though technology has a stronghold on all domains, in this paper, we will focus on the impact on law specifically focusing on the topic deepfakes and its effect on the credibility of digital evidence.

Here digital evidence is the upside of technology i.e. it presents more credible and solid evidence to the courts, aiding them in accurately determining whether to give a verdict of guilty or not guilty. Digital evidence refers to any probative information or data that is stored, transmitted, or collected in an electronic format which can be used or presented in a court of law as proof. It includes text messages, emails, GPS logs, browser history, etc. They are commonly extracted from sources like Computers, laptops, tablets, etc.

The existence of DEEPFAKES spirals all the advantages of law integrating technology into the process of giving justice. This concept shattered the assumption of "seeing is believing" that has been persisting in the courtrooms for generations.

Deepfake refers to any form of digital media such as videos, images, or voice audio clips which have been altered realistically or have been generated entirely using advanced artificial intelligence (AI) to make someone appear as if they said something or done something that they never actually did.

Deepfakes pose a profound and a structural challenge to the acceptance of digital evidence and the conviction of criminals based on such evidence. This concept fundamentally shifted how judges evaluate admissibility and how juries perceive reasonable doubt.

In this article we are going to explore all the problems, advantages, disadvantages and the solutions in relation to the topic deepfakes and its effect on the digital evidence acceptability and credibility.

KEYWORDS:

Digital evidence, Deepfake, The "Liar's Dividend" (Plausible Deniability), AI-generated Content, Law of Evidence, Digital Forensics, AI Detection Tools, Digital Verification, Forensic Authentication, Bharatiya Sakshya Adhiniyam, Epistemic Crisis

INTRODUCTION

Law has always been evolving according to the times and technology to be relevant with the social change so that the laws would never be outdated and protect the people from the contemporary issues that they face rather than be struck in time. Law has integrated technology into various aspects including accepting digital evidence. Though this has multitudinous advantages but it also poses abundant chances of manipulation and manifolds the chance of giving the wrong verdict. Few of the significant impacts of DEEPFAKES on the acceptability of digital evidence are the "Liar's Dividend" (Plausible Deniability), Weaponized Authentication Costs and Evidentiary Infiltration (The Risk of Wrongful Conviction) among many others.

The courts have been taking proactive steps to counter attack the major problems of accepting digital evidence to upheld the abundance of trust that people have on the courts and law to protect and emanate justice.

Let’s unravel all aspects of deepfakes and the crisis of digital evidence piece by piece gradually in detail.

DIGITAL EVIDENCE

In this generation, everyone has a smartphone even if it is just a small one and all the communication is done through this phone or through any other electronic devices like laptop, tablets, etc.

The calls, texts, emails, photos or recordings that can potentially reveal the intentions of a suspect, aid in tracking the chronological activity of the suspect which helps in tracing his/her activities in order, checking the suspect's whereabouts tells about his/her location during the happening of the event and check out his/ her alibi which can be produced in a court of law to prove or disprove something is called as digital evidence.

Hence, digital evidence is any data stored or transmitted in digital form that can support or reject a hypothesis in court of law. It is essential in modern forensic investigations which allows law enforcement agencies in establishing a timeline, intent and relationships.

Examples include emails, CCTV footage, chats, call recording, GPS location, etc.

Digital evidence includes a wide range of electronic devices and formats. They are as follows:

- Communication: Text messages (SMS/MMS), emails, and call logs.
- Files & Media: photos, videos, audio, and documents.
- Location Data: GPS coordinates, navigation history, and cell tower connection records.
- Online Activity: browser histories, IP addresses, search queries, and social media interactions.
- System Data: Metadata, hidden files, deleted files (if not overwritten), and system registry keys

These sorts of evidences can be extracted from any device that processes or stores data, which includes a wide range of devices such as Computers, laptops, tablets, Smartphones and smartwatches, Smart home appliances (e.g., smart TVs, connected thermostats), Cloud storage servers and online accounts.

As digital data is very fragile and alarmingly prone to manipulation, professionals use sophisticated techniques called DIGITAL FORENSICS to collect and analyse the data safely and protect its legal integrity at the same time.

DEEPFAKES:

Deepfakes are what the name suggests they are AI generated audios, videos or photos that are inconceivably realistic. They exhibit people doing or saying things that they never actually did, they alter the photos, change voices and even accurately and convincingly replace a person or their voice a video or audio with another for example there are a lot of trending ai generated videos which are highly realistic that are flooding social media platforms such as Instagram, twitter and YouTube. One such example is the fabricated video where president Zelenskyy was depicted telling his troops lay down arms.

Deepfakes are essentially media (video, audio, images) that are AI-generated or manipulated and which can be presented as if they are genuine recordings. In digital evidence, this raises serious concerns regarding the authenticity, reliability, and admissibility.

This poses a challenge as courts rely on digital evidence to present facts and deliver verdicts. Deepfakes fabricate or alter reality shattering the phenomenon of seeing is believing that the court have been following for centuries.

Deepfakes in digital evidence may lead to detrimental outcomes. Few of the examples are:

- False Accusations: A deepfake video could lead to a wrongful conviction or acquittal.
- Fabricated Confessions: AI-generated audio can mimic voices for creating fake admissions which also violate the fundamental rights of citizens in addition to misleading the courts which s deemed as an action of contempt of court.
- Erosion of Trust: Even genuine evidence may be doubted because of the possibility of deepfakes this is called the “liar’s dividend.” Due to this many prosecutors drop charges as even a genuine video may be claimed as deepfake by the defence and it also costs a fortune to prove them to be authentic leading to waste of resources and time adding to already a mountain of backlog of cases.
- Authenticity Test: Under the rules of evidence (like the BSA or Federal Rules of Evidence in the US) digital evidence must be proven to be authentic. Deepfakes

complicate this because they can pass casual inspection. To prevent this a lot of safeguards are implemented.

- Chain of Custody: Lawyers must show that digital files were preserved without tampering. Deepfakes present a loophole to the defence, which they can use to exploit weak chains of custody.

DIGITAL FORENSICS:

Think of the process of digital forensics as the guardian and protector of the integrity of digital evidence ensuring that it is not tampered and admissible in the court of law. Its ultimate goal is to uncover the evidence hidden in digital information and present it in a way that is legally acceptable.

Digital forensics is, at its heart, the science of keeping digital evidence trustworthy. It's the careful process of finding, preserving, and examining information stored on computers, phones, networks, or even the cloud, so that it can stand up in court or in an investigation. In today's world, where deepfakes can make people appear to say or do things they never did, digital forensics has become more than just a technical discipline — it's a guardian of authenticity.

Think of the forensic process as a journey. It starts with identifying where the evidence might be — a laptop, a server, or a mobile device. Next comes acquisition, where experts make an exact copy of the data without changing the original. Preservation ensures that this copy stays safe from tampering. Then comes the analysis, where specialists dig into files, logs, and metadata to uncover hidden details or inconsistencies. Finally, the findings are presented in reports or testimony, translated from technical language into something judges, lawyers, and juries can understand.

One of the most important safeguards is the chain of custody. This is like a diary that records every step the evidence takes — who collected it, when it was handled, and how it was stored. In a world of deepfakes, this record is crucial because it proves that the evidence hasn't been swapped or altered along the way. Even the way evidence is packaged and transported matters: sealed containers, tamper-proof labels, and secure storage prevent interference, while Faraday bags block wireless signals that could remotely change data.

When deepfakes enter the picture, digital forensics becomes the frontline defence. Experts use advanced tools to spot subtle signs of manipulation — mismatched pixels, odd lighting, or

metadata that doesn't add up. These techniques allow them to separate genuine recordings from synthetic ones. By following strict procedures and documenting every step, forensic specialists give courts confidence that what they are seeing or hearing is real.

Hence, Digital forensics is a disciplined process of uncovering, preserving, and interpreting information which is stored in electronic systems in such a way that it can serve as credible evidence in investigations and in the courts. It focuses on examining computers, mobile devices, networks, cloud environments to trace activities, recover hidden or deleted data, and demonstrate whether digital records have been altered. The aim is to maintain the integrity of the evidence while translating complex technical findings into clear, legally admissible proof.

Digital forensics is not just about recovering files or tracing activity. It's about protecting truth itself. As deepfakes grow more sophisticated, the discipline ensures that justice remains grounded in facts, not illusions.

BSA PROVISIONS

The Bharatiya Sakshya Adhiniyam (BSA), 2023 is India's modern law of evidence, designed to keep pace with the digital age. Although the Act does not directly mention deepfakes, its provisions are central to protecting courts from the dangers of manipulated digital material. In a time when artificial intelligence can create videos or audio that look and sound real, the BSA provides the legal scaffolding to test whether electronic records are genuine or deceptive. One of its most important updates concerns electronic records. The old Section 65B of the Indian Evidence Act has now been replaced by Section 63(4), which continues the requirement of certification but presents it in clearer, more practical language.

One of the most promising aspects of the BSA is its inclusive definition of electronic records under Section 2(1)(d). This ensures that everything from CCTV footage and WhatsApp chats to cloud-stored files and digital photographs falls within the evidentiary net. By recognizing electronic records as documentary evidence under Section 61, the law acknowledges that digital material is now as important as paper documents in proving facts.

Sections 62 and 63 add further clarity by distinguishing between primary and secondary evidence. Primary evidence refers to the original device or storage medium, while secondary evidence includes copies such as printouts or screenshots. This distinction matters because deepfakes often circulate as "copies" — manipulated versions of genuine recordings. By

requiring stricter compliance for secondary evidence, the law builds a safeguard against casual acceptance of potentially fake material.

The real backbone of the system is Section 63(4), an electronic record can be admitted as evidence only if it is accompanied by a certificate. This certificate is not just a formality — it must identify the record, explain how it was produced, and be signed by someone in a responsible position. Crucially, it must also include the hash value of the file. A hash value is like a digital fingerprint: a unique code generated by algorithms such as SHA-256. Even the smallest change in a file alters its hash completely, which makes it a powerful way to prove that a piece of evidence has not been tampered with. By requiring hash values, the BSA adds a strong safeguard against manipulation, ensuring that courts can check whether a video, audio clip, or document is truly original.

Despite these strengths, the BSA faces serious challenges when confronted with deepfakes. Certification under Section 63(4)B was designed for traditional electronic records, not synthetic media created by AI. A deepfake video can be generated with convincing metadata, making it appear authentic even when it is not. This means that certification alone may not be enough to detect manipulation.

Another limitation is the absence of explicit provisions for synthetic media. While the law covers electronic records broadly, it does not provide specific guidelines for identifying or excluding deepfakes. Courts must therefore rely heavily on forensic experts, who analyse pixel inconsistencies, audio distortions, or metadata trails to prove manipulation. This reliance increases costs, delays, and the burden on the judicial system.

Even the chain of custody, though vital, is vulnerable to human error. If documentation is incomplete or improperly maintained, the credibility of evidence can be questioned. In high-profile cases, this gap could allow deepfake material to slip through unnoticed.

RECOMMENDATIONS

To strengthen the BSA framework against deepfakes, scholars and practitioners have suggested additional safeguards. AI watermarking can embed invisible authenticity markers in genuine recordings, making it easier to distinguish them from synthetic content. Blockchain-based chain of custody can provide tamper-proof tracking of evidence handling, ensuring transparency at every stage. Mandatory forensic audits for suspicious digital evidence would

further enhance reliability, giving courts confidence that manipulated content has been filtered out. These recommendations highlight the need for a dynamic approach. As technology evolves, the law must adapt to ensure that justice remains anchored in verifiable facts. These measures would complement Section 63(4), reinforcing trust in digital truth.

In short, Section 63(4) of the BSA provides a modern safeguard for electronic records, with hash values acting as a digital fingerprint to protect authenticity. While this strengthens the law's ability to filter out tampered files, deepfakes expose its limits, reminding us that legal rules must evolve alongside technology. The future of digital evidence lies in combining statutory safeguards with advanced forensic techniques, ensuring that truth remains verifiable even in a world of synthetic deception.

EPISTEMIC CRISIS AND DIGITAL EVIDENCE

The rise of deepfakes does not only challenge the technical side of law; it creates what scholars call an epistemic crisis — a breakdown in our collective ability to trust what we see and hear. In simple terms, an epistemic crisis occurs when people begin to doubt whether knowledge itself can be verified. In the courtroom, this translates into a dangerous situation: if judges, lawyers, or the public lose confidence in digital evidence, even genuine recordings may be dismissed as fake, while fabricated ones may slip through as real.

Deepfakes intensify this crisis because they blur the line between truth and illusion. A video of a confession, a photograph of a crime scene, or an audio clip of a conversation can all be convincingly forged. Once society knows that such manipulation is possible, the credibility of all digital evidence is weakened. This is sometimes described as the "liar's dividend" — the advantage gained by dishonest actors who can now claim that authentic evidence is fake, simply because deepfakes exist.

The Bharatiya Sakshya Adhiniyam (BSA), 2023 attempts to guard against this crisis through provisions like Section 65B certification, which requires parties to prove how electronic records were produced and preserved. While this safeguard is valuable, it cannot fully resolve the deeper problem of trust. Certification may confirm the technical process, but it does not guarantee that the content itself has not been synthetically generated. This is where digital forensics plays a vital role, using advanced tools to detect anomalies in pixels, metadata, or

audio signals. Yet even forensic science faces limits, as AI becomes more sophisticated in masking traces of manipulation.

The epistemic crisis therefore forces us to think beyond procedure. It is not enough to rely on legal rules alone; courts must embrace new safeguards such as AI watermarking, blockchain-based chain of custody, and mandatory forensic audits. More importantly, the justice system must cultivate public confidence by showing that it can adapt to technological threats. Without this, the crisis of trust could erode the very foundation of evidence law, leaving truth vulnerable to doubt and deception.

CASE LAWS

ANVAR P.V. V. P.K. BASHEER (2014)

This landmark Supreme Court case reshaped the way Indian courts treat electronic evidence. The dispute revolved around whether digital records such as CDs and printouts could be admitted without proper certification. The Court held that electronic evidence must comply strictly with statutory requirements, then under Section 65B of the Evidence Act, now Section 63(4) of the BSA. The judgment emphasized that authenticity must be backed by a certificate explaining how it was produced and preserved. It shows that courts already recognized the fragility of digital material and insisted on procedural safeguards long before AI-generated media became widespread. Certification adds accountability, but it cannot alone detect synthetic manipulation. Thus, Anvar P.V. v. Basheer stands as a reminder that while legal rules can filter evidence, technology must evolve alongside them to protect truth in the age of deepfakes.

ARJUN PANDITRAO KHOTKAR V. KAILASH KUSHANRAO GORANTYAL (2020)

In this case, the Supreme Court revisited the issue of electronic evidence admissibility. The dispute concerned whether video recordings and digital files could be admitted without certification. The Court reaffirmed the principle laid down in Anvar P.V. v. Basheer, making it clear that certification is not optional but mandatory. By insisting on certification, the Court sought to ensure that only records with a traceable origin and integrity are admitted. However, the case reveals a limitation: certification can confirm process but not content. This shows why

forensic tools, hash values, and advanced safeguards must complement the statutory framework.

CONCLUSION

Deepfakes have created an epistemic crisis by eroding trust in digital truth. The BSA's provisions, supported by case law across jurisdictions, provide a framework for authenticity, but the future demands more. AI watermarking, blockchain custody, and forensic audits must complement statutory safeguards. Ultimately, the fight against fabricated reality is not just about law or technology — it is about preserving trust in justice itself. Courts must adapt, experts must innovate, and society must remain vigilant, ensuring that truth remains verifiable even in a world where illusion can be engineered with a click.

REFERENCES

1. *Anvar P.V. v. P.K. Basheer*, (2014) 10 SCC 473 (India).
2. *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, (2020) 7 SCC 1 (India).
3. Bharatiya Sakshya Adhiniyam, No. 23 of 2023, Acts of Parliament, India.
4. Indian Evidence Act, No. 1 of 1872, Acts of Parliament, India (repealed).
5. Law Commission of India, *Report on Digital Evidence and Admissibility*, Government of India (2023).
6. Supreme Court of India, *Compilation of Judicial Pronouncements on Electronic Evidence*, Registry Publication (2021).
7. Aparna Chandra, *Technology and Evidence Law in India: Challenges of Digital Authenticity*, 12 Indian J.L. & Tech. 45 (2021).
8. Prashant Reddy T., *Deepfakes and the Indian Legal System: Evidentiary Concerns in the Age of AI*, 8 NALSAR L. Rev. 112 (2022).
9. Rituparna Das, *Digital Forensics and Chain of Custody in Indian Courts*, 15 NUJS L. Rev. 89 (2020).

10. Shruti Bedi, *Epistemic Crisis and the Law of Evidence: Deepfakes in Indian Context*, 6 Jindal Global L. Rev. 134 (2023).

11. Indian Cyber Law Journal, *Authenticity of Electronic Records under the Bharatiya Sakshya Adhiniyam*, Vol. 18 (2024).