VOLUME 6 | ISSUE 4

INTERNATIONAL JOURNAL OF ADVANCED LEGAL RESEARCH

FROM HUMAN TESTIMONY TO MACHINE INTELLIGENCE: THE RISE OF AI EVIDENCE

- Jahanvi Tiwari[1] & Jai Kishan Mishra[2]

Abstract

The way courts collect, analyse, and present evidence is being altered by artificial intelligence (AI). Traditionally, courts relied on human witnesses, document-based evidence, and expert testimony to build their case. Recently, however, the introduction of machine learning (ML), facial recognition (FR), predictive analytics (PA), digital forensics (DF) and automated surveillance (AS) has resulted in AI-based evidence being included in the court system. The research will utilise both doctrinal and analytical methods to determine how AI-based evidence can be admitted, valued as evidence, and given a status under the Bharatiya Sakshya Adhiniyam, 2023 (the Evidence Act). The study also examines how AI is used during criminal investigations; how to detect cybercrime; how to perform forensic examinations; and how to manage evidence. Another reason this study is of great significance is that it contains information on how AI will improve the accuracy, efficiency and reliability of processing complex digital data while minimising the potential for human errors to be made. There are still many barriers to overcome with respect to algorithmic bias, transparency, data privacy, accountability and the integrity of AI-generated evidence. The findings of the study indicate that AI is to be used in conjunction with human judgement in the justice system. The study provides valuable insight into the intersection between law and technology for those studying or otherwise interested in evidence and justice as it relates to the digital world. The study reaches one additional conclusion: that there is a pressing need for a comprehensive legal and ethical framework that will facilitate the responsible integration of AI into the justice system.

[1]Student at School of Legal Studies, Babu Banarasi Das University, Lucknow, Uttar Pradesh, India
[2]Assistant Professor, School of Legal Studies, Babu Banarasi Das University, Lucknow, Uttar Pradesh, India

KEYWORDS: Artificial Intelligence (AI), AI-Generated Evidence, Bharatiya Sakshya Adhiniyam, 2023, Digital Forensics, Admissibility of Evidence, Criminal Investigation, Algorithmic Bias, Administration of Justice.

INTRODUCTION

Traditionally, the judicial system relied heavily on humans to provide information about events, and as a result, the provision of that information in the form of witness testimony was considered the most important way to establish a case legally. However, in most cases, witnesses were biased against one another for various reasons. As a result, they may misremember events because of threatening or intimidating circumstances or their flawed ability to recall what truly occurred.

As courts started to realise the limitations of relying upon human testimony to establish factual determinations and dispute resolution, they became more willing to accept scientific and digital evidence as a way to establish those same factual determinations and provide a resolution to those disputes. Technology, such as computers, cell phones and emails, has made documenting events much easier, which contributes to the courts' acceptance of evidence because it provides the courts with more accurate accounts of events. Therefore, digital evidence plays a critical role in the judicial system's ability to meet its burden of proof to establish a case legally. Due to the increased volume of digital evidence, the courts have begun to develop laws and rules of evidence related to these forms of evidence.

Currently, courts are beginning to accept AI as a form of evidence for establishing fact and providing resolution to a dispute. However, in contrast to traditional digital evidence, AI can access and analyse large volumes of data, identify patterns, assist with facial recognition technology, support digital forensics, and produce more efficient investigations than human technology alone can provide. In addition, courts will be able to use AI to make judicial determinations faster, improve the accuracy of the determinations, and ultimately create a more equitable, fair and just judicial process, which would constitute a major improvement over human testimony or traditional digital evidence.

This study examines the nature, scope, and evidentiary value of AI-generated evidence, focusing on its distinction from traditional forms of evidence such as oral testimony, documentary records, and expert opinions. It critically analyses the legal challenges surrounding the admissibility, reliability, authenticity, transparency, and accountability of AI-generated evidence in judicial proceedings. The research further evaluates whether the

existing Indian legal framework, particularly the Bharatiya Sakshya Adhiniyam, 2023, is adequately equipped to address the growing integration of AI within the justice system and identifies reforms necessary for its effective regulation. Adopting a doctrinal research methodology, the study relies on statutes, judicial precedents, scholarly literature, policy reports, and comparative legal analysis. It examines landmark decisions including Anvar P.V. v. P.K. Basheer[3], Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal[4], and Justice K.S. Puttaswamy v. Union of India[5], while drawing insights from international legal approaches to AI evidence.

The Bharatiya Sakshya Adhiniyam, 2023[6] and the Information Technology Act, 2000[7]recognise electronic records and digital proof, but do not meet the demands created by AI-created proof in the law. Both rules were designed primarily to provide a method of regulating electronic records, confirming their admission into court as well as any lawful establishment of the proof's reliability, value and authenticity. However, there are no guidelines contained within these laws to specify different treatments for the output of AI systems such as machine learning algorithms, predictive analytics, autonomous decision-making systems and deepfakes; therefore,several basic principles regarding how to assess algorithmic transparency, explainability, accountability, reliability and threats of bias associated with these tools have no regulatory framework. The increasing presence of AI in such areas as facial recognition, voice analysis, forensic investigations and assessment of evidence makes the lack of a specific statutory framework surrounding the admissibility, weight of evidence, liability for errors or infringement of fundamental rights difficult to navigate legally. The need to create a regulatory scheme to cover the use of AI-generated proof requires comprehensive reform to ensure a fair, transparent, due process and properly judicially supervised administration of justice.

2. EVOLUTION FROM THE HUMAN TESTIMONY TO THE MACHINE INTELLIGENCE

The purpose of this report is to outline how courts have used different types of evidence over time. The use of human testimony alone and the problems associated with it (for example, the

[3]Anvar P.V. v. P.K. Basheer, (2014) 10 S.C.C. 473 (India).
[4]Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 SCC 1 (India).
[5]*Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 S.C.C. 1 (India).
[6]Bharatiya Sakshya Adhiniyam, No. 47 of 2023, §§ 61–63 (India).
[7]Information Technology Act, No. 21 of 2000, § 4, India Code (2000) (recognizing the legal validity of electronic records).

case of Ronald Cotton) are contrasted with more recent advancements in forensic science (e.g., fingerprinting and DNA testing) within the field of evidence law. Additionally, we will explore Innovations such as using Artificial Intelligence (AI) both to increase accuracy and reliability in policing today, as well as Potential Issues relating to AI: bias, authenticity, and legal admissibility, respectively, by the Criminal Justice System.

2.1 HUMAN TESTIMONY: THE BASIS OF TRADITIONAL EVIDENCE

Primarily, all justice systems were built on Human Testimony. All courts relied primarily on witnesses (victims, defendants, and/or community members) due to the lack of any scientific techniques or technological tools with which to ascertain the truth. Most evidence was given orally. The decision of a case was generally predicated upon the credibility, truthfulness, and memory of all who appeared in front of the court. Generally speaking, the best type of witness was a person who had "eyes" on the ground when the event occurred, as these individuals were able to provide "firsthand" evidence regarding the actual events in controversy. Thus, all judges heard evidence primarily by way of both examination and cross-examination so that they could "judge" the believability of all witnesses based upon both the content of what was said (i.e., the substance) and the way it was stated (i.e., the demeanour) by all witnesses.

However, as valuable as Human Testimony is, it can never be completely relied upon. Different witnesses may perceive the same event differently, resulting in conflicting testimony. While human testimony remains a crucial part of the legal process, there are many factors that limit its use. Research in psychology has shown that memory is not perfect, nor does it last forever. Elizabeth Loftus, a leading authority on cognitive psychology, has argued that memories are more about how they have been constructed from our perceptions, suggestions and experiences, rather than how they were actually experienced[8]. As a result, the information given by witnesses may be affected by the same kinds of things that impact the eyewitnesses themselves, such as the amount of stress at the time of the crime, the length of time that has passed since the crime, the amount of ambient light present during the time period of the crime, and what other witnesses may say about the event. There is also the possibility that testimony may be false (deliberate), may be a case of mistaken Identity,or may be exerted upon the person providing the testimony. Because the development of Legal Systems (as a society) continued to be developed, and as the need for factual determinations

[8] Elizabeth F. Loftus, Eyewitness Testimony (Harvard University Press 1996).

became more important, the limitations of Human Testimony demonstrated the need to supplement Human Testimony with additional types of evidence that were more objective and, therefore, verifiable.

2.2 THE NEED FOR SCIENTIFIC AND DIGITAL EVIDENCE

Over time, evidence law has changed dramatically, with a shift from prior reliance upon human testimony as evidence to reliance on scientific/digital evidence. This shift resulted from the understanding that human memory and perception cannot be relied upon for all purposes. Ronald Cotton's[9] case shows the problems with eyewitness testimony and how that can affect the justice system. The victim was absolutely sure that Ronald was the attacker, but DNA evidence later proved that he had not committed the crime at all. Years later, and after having spent a lot of time in prison for a crime he did not commit, he was exonerated. This case illustrates how human memory cannot always be trusted, and that this can lead to wrongfully convicted people. This case illustrates the need for testimonial evidence to have physical (scientific) counterpart evidence, so the facts established through forensic & technological means will be accurate and reliable for use by the courts to determine the facts of a case. Due to the limitations of traditional methods, science has taken an increasingly important place in trials. Scientific evidence provides a much more impartial and verifiable way to determine facts since it is based on physical measurements rather than on personal beliefs. The application of various scientific techniques (e.g., fingerprinting, bullets, toxicology, blood testing, DNA analysis) has changed the way we conduct investigations and enhanced the ability of the courts to reach the correct decision.

The first example of the validity of scientific evidence occurred in the Will West trial[10] in 1903. Two inmates who were almost identical in appearance and had similar names were incorrectly identified as being one person using conventional identification methods. The use of fingerprint analysis enabled law enforcement to differentiate between the two, emphasising the advantage of using a scientific means of identifying individuals and promoting the use of fingerprinting in future investigations.

The introduction of DNA analysis to prove the guilt or innocence of people accused of a crime significantly changed the body of evidence in law. An early and pioneering example of

[9]Jennifer Thompson-Cannino, Ronald Cotton & Erin Torneo, *Picking Cotton: Our Memoir of Injustice and Redemption* (2009).
[10]Charles E. Chapel, *Fingerprinting: A Manual of Identification* 48–50 (4th ed. 1941).

this took place in the United Kingdom with the Colin Pitchfork[11] case in 1986. This was the first known use of DNA evidence to both free a wrongly convicted man and to identify a true offender in a criminal case. The success of using DNA analysis as a form of evidence demonstrates that scientific evidence can eliminate mistakes made by witnesses, in confessions, or through investigative assumptions. Therefore, DNA analysis has become one of the most reliable forms of scientific evidence.

Digital evidence can be very important in an investigation, such as during the investigation into the Mumbai Massacre of 2008[12]. Investigators used digital evidence to trace what the attackers did and where they went. Digital evidence also helped identify those involved in the attack. As a result, electronic records are often relied on to provide a more reliable source of information than just a witness's memories of what happened when investigating crimes, like cybercrime and financial fraud, today.

With all the technology that is being used today, there are now machine-generated forms of evidence. Machine-generated forms of evidence include such things as surveillance cameras, biometric machines, automatic transaction records, and devices that record data (surveillance) without humans having to do anything. The result is that this type of evidence is typically considered to be more objective and therefore less prone to errors due to human memory and perception than other types of evidence.

Digital evidence has come under considerable pressure and difficulties throughout the course of time with regard to its use in criminal investigations. Digital evidence now faces challenges relating to data tampering, deepfake technology, the technical complexities of understanding how to collect, store, analyse and retrieve from the massive volume of digital evidence (electronic data), where electronic data is saved and stored on global platforms and/or databases, and how time-consuming the analysis of digital evidence can be. All these issues have led to concern about the authenticity, reliability, and efficient use of digital evidence. Although the Information Technology Act, 2000 and The Bharatiya Sakshya Adhiniyam, 2023, provide that electronic records of evidence may be accepted. However, the provisions for legal recognition do not overcome the inherent limitations of using electronic records of evidence in practice, such as data overload, digital tampering, deepfake technology, technical complexity, and time-consuming data analysis involving large volumes

[11] *R v Pitchfork*, [1988] 1 Crim. L.R. 513 (C.A.)
[12] P. Chidambaram, Statement by the Minister of Home Affairs Regarding the Recent Terrorist Attack in Mumbai, Lok Sabha Debates, Dec. 11, 2008, available at Indian Kanoon

of electronic records, which will require high levels of AI-based tools to improve investigation process efficiency, accuracy and reliability.

As a result of the identified issues with traditional methods of managing electronic records, investigators must turn to more advanced technology, such as Artificial Intelligence, to assist them with the processing and analysis of digital evidence in an efficient manner.

3. INSIGHTS REGARDING ARTIFICIAL INTELLIGENCE (AI) IN THE CRIMINAL JUSTICE SYSTEM

3.1 Definition & Scope of Artificial Intelligence within the Criminal Justice System

Artificial Intelligence (AI) has emerged as one of the most transformational technologies of the 21st Century and is impacting courts and legal systems globally. AI is defined as computer systems that are capable of performing tasks that typically require human-like cognitive functions (*e.g., learning, reasoning, solving problems, recognising patterns, making decisions*)[13]. This reflects the larger digital transformation of modern-day justice systems as the legal profession faces increasingly complex disputes and huge amounts of electronic data. AI provides new solutions for improving judicial processing through enhanced efficiency, accuracy and consistency in the decisions made by courts. Rather than replacing judges/legal practitioners, AI is being used as an assistive or ancillary tool to support the administration of justice (e.g., conducting analyses of physical evidence, conducting legal research, managing cases, and conducting investigations).

3.2 Core AI Technologies That Are Utilized InThe Justice System

Machine Learningis probably the most utilised application of AI within the Justice System. The use of machine learning allows computers to learn from experience (data) and improve their performance without being explicitly programmed. The ability of machine-learning algorithms to analyse enormous amounts of data, detect similarities or differences between those data sets, and ultimately produce valid and accurate results is extremely valuable in the Justice System. The machine learning technologies are utilised more than ever before (with respect to) the use of machine learning within the legal space by way of document review and compliance, fraud detection (and/or) stimulation of document review, legal analytics and verification of evidence. Another significant application of AI within the Justice System is the use of ***Facial Recognition Technology (FRT),*** whereby Facial Features (of an

[13]Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach* 1–35 (4th ed. 2021).

unidentified individual) are compared to pre-encrypted photographs/footage stored within databases; this comparative process results in the identification of an Individual's (facial) features. The third application of AI is through***Natural Language Processing (NLP)***. NLP allows computers to read or write to people who use a particular language. By using NLP, legal professionals can conduct Statutory Analysis, Contract Analysis, Judicial Analysis/Witness Statement Analysis, and Electronic Communication Analysis in a more efficient manner than without Natural Language Processing.

3.3 Using AI-Powered Predictive Analytics and Digital Forensics.

AI uses AI in Predictive Analytics, which is the analysis of past data to project probable future events and trends. AI helps in identifying patterns of criminal behaviour, estimating costs associated with legal actions, providing information about how to best allocate funding for police investigations, and assisting in making strategic decisions about whether to continue or terminate an investigation, among other things. AI is also used as an important tool in Digital Forensics, which involves collecting, preserving, analysing and presenting electronic evidence from the Internet and all forms of digital media. The use of AI tools helps forensic experts recover data, detect hidden data, identify unusual activity and recover pertinent evidence from computers or mobile devices, online storage sites, e-mails, websites, and other forms of digital communication.

3.4 IMPORTANCE OF AI IN CONTEMPORARY LAW

The rising use of AI's in the judiciary is a manifestation of its potential to fortify the evidentiary system and hence improve the general administration of justice. AI gives rise to more rapid investigations, improved evaluation of evidence, more efficient procedures, and improved access to legal information. AI can also decrease errors due to human tiredness, limitations on memory, or unconscious prejudgments, so that the reliability of an evaluation of evidence is increased. However, AI is not a replacement for human judgment. Judges, lawyers and investigators continue to have the duty to evaluate evidence, interpret the law, and comply with procedural safeguards and rights guaranteed by the Constitution.[14] To this end, the successful implementation of AI into the legal system will require balancing the new technology with adequate legal regulation, transparency, accountability, and human oversight.

[14]Richard Susskind, *Online Courts and the Future of Justice* 83–110 (2019).

4. THE USE OF AI-GENERATED EVIDENCE IN COURT UNDER THE BHARATIYA SAKSHYA ADHINIYAM (BSA), 2023:

4.1 LEGAL STATUS OF ELECTRONIC AND DIGITAL EVIDENCE IN COURT

With society's fast-moving digital transformation, how we present evidence to the courts has changed rapidly. Today, there are large quantities of types of digital data, such as electronic records, digital communications (e.g. email), CCTV footage, social media content, computer-generated documents, and everything else that is classed as 'digital evidence' that must be considered in any type of court litigation. With this change from the physical (paper) to the electronic (digital) world, the BSA provides the legal framework recognised under the Indian Evidence Act 1872[15] for using 'electronic evidence' as documentary evidence in a court; however, the BSA does provide some flexibility in how the existing framework is applied because of the many changes in technology today.

The electronic evidence that the BSA recognises that digital records tend to provide a more reliable and objective source of evidence than what is recalled by witnesses in court. Digital evidence, unlike oral evidence from a witness, is not subject to any of the problems associated with the way in which a witness may recall (*like relying on their memory*) or mis recall (e.g. *intentionally give false information about something*), since this digital evidence will always remain in an electronic format that can be accessed, retrieved, and verified a large period of time later. As a direct result of this change in the legal status of electronic evidence, the growing use of electronic evidence has created significant challenges for its use within the judiciary in cases of criminal investigation, cybercrime prosecution, financial fraud, or commercial litigation.

Digital evidence is becoming more prevalent in courtrooms around the country; however, the limitations of the Bharatiya Sakshya Adhiniyam, 2023, are being recognised and addressed due to this increase. Although the Act provides that electronic records shall be admissible as evidence, no parameters or standards exist for assessing the authenticity, reliability, or interpretation of artificially intelligent (AI) generated output. Thus, there is an uncertain evidentiary value, and procedural fairness will exist when using AI in judicial proceedings.

4.2 The Applicability of Artificial Intelligence-Generated Evidence under the Current Legal Framework

[15] Indian Evidence Act, No. 1 of 1872, India Code (1872).

There is considerable debate as to whether AI-generated evidence can be accommodated under the existing legal framework established by the Bharatiya Sakshya Adhiniyam, 2023.

Even though Section 61-63[16]The BSA, 2023 acknowledges electronic/digital records, they do not regulate the use of AI-produced evidence. This is important because evidence produced by AI doesn't just store data; it also analyses data and draws conclusions based on data processed by it using an algorithm. For example, facial recognition technology produced using AI/ML systems is used by law enforcement agencies in India and generates identification matches by an algorithmic process. The current legal framework was designed to govern the creation of electronic records by individuals. The current framework does not define an autonomous output or standards for evaluating the automated outputs of algorithms.

The level of weight that can be assigned to evidence presented in any court or tribunal is based on the reliability, authenticity, accuracy, and explainabilityof the evidence. As AI systems can quickly process massive files of data, they can help identify and create patterns, detect anomalies, and support digital forensic investigations with great accuracy. However, the results produced by AI are reliant on both the quality of data used to train them and the way in which the algorithms themselves are designed. Because AI systems operate consistently, they are less prone to human error compared to people who may suffer from fatigue or have lapses in memory. Nonetheless, there are wide-ranging issues surrounding programming errors in machines and the presence of bias in the datasets used to develop the algorithms. In addition, it is critical that any evidence produced using AI has not been tampered with or altered over time, technology, or other methods. Courts must also provide opportunities for parties to challenge evidence produced by AI, and, therefore, they must ensure that evidence produced by AI systems can be sufficiently explained to ensure that it provides a transparent and reliable basis upon which judicial decisions can be made.

4.3 JURISPRUDENTIAL ADVANCEMENT FOR DIGITAL EVIDENCE

The ongoing development of digital evidence through decisions by Indian Courts has provided and emphasised the importance of following established processes to guarantee authentic digital evidence.

[16]Bharatiya Sakshya Adhiniyam, No. 47 of 2023, §§ 61–63, India Code (2023) (Section 61 recognizes electronic and digital records as documents; Section 62 provides special provisions relating to evidence of electronic records; and Section 63 governs the admissibility of electronic records).

Following Anvar P.V. v. P.K. Basheer (2014),the Supreme Court of India determined the criteria for admissibility of digital evidence and stressed the requirement for an electronic evidence certification under **Section 65B of the Indian Evidence Act, 1872**[17] and now under **Section 63(4)(c) of Bharatiya Sakshya Adhiniyam, 2023.** The Supreme Court determined that digital evidence is a unique kind of evidence and can be admissible provided it meets the legal criteria.

The Supreme Court of India reaffirmed the guidelines for determining the authenticity and reliability of electronic evidence established by the case of Anvar P.V. in its judgement, Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal[18]AIR 2020 SC 4908. This decision upholds and bolsters traditional safeguards for electronic evidence; however, it creates a regulatory deficiency due to a lack of specific standards in relation to AI-produced documents or algorithms under the BSA, 2023, for the purpose of evaluating the accuracy, explainability, and accountability of AI-created evidence and algorithmically produced outputs.

The establishment of the electronic evidence jurisprudence has provided additional assurance of the authenticity of digital records, while the introduction of AI technology creates new avenues for the completion of the investigative process and forensics. As a result, reliable authentication processes must be established to demonstrate that AI-generated evidence will be suitable for use in court.

5. AI-GENERATED EVIDENCE: AUTHENTICATION

For any evidence to be admitted into evidence, it must first be authenticated. Before relying on AI-generated evidence, courts must ensure that the evidence is genuinely, accurately and sufficiently reliable - and that the evidence has not been tampered with.[19] Traditional electronic forms of evidence are normally authenticated using metadata, digital signaturesor through certification procedures; whereas, AI-generated forms of evidence are much more difficult to authenticate because they can be created, modified or enhanced without the involvement of a human.

[17] 4(c) dealing with any of the matters to which the conditions mentioned in sub-section (2) relate, and purporting to be signed by a person occupying a responsible official position in relation to the operation of the relevant device or the management of the relevant activities (whichever is appropriate) shall be evidence of any matter stated in the certificate; and for the purposes of this subsection it shall be sufficient for a matter to be stated to the best of the knowledge and belief of the person stating it.

[18] *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, (2020) 7 S.C.C. 1 (India).

[19] Richard Susskind, *Online Courts and the Future of Justice* 94–110 (Oxford Univ. Press 2019).

Through the rapid development of generative artificial intelligence, it is now possible to create highly realistic 3D images, video, audio recordings and written text.The differences between AI-generated forms of evidence and authentic forms of evidence can be difficult to detect, making it extremely difficult to determine the origin of any AI-generated output and to assess its reliability. Therefore, a court may have significant difficulty assessing whether an AI-generated output reflects an accurate copy of reality or has been algorithmically manipulated. Therefore, to properly authenticate AI-generated evidence, a court will need to utilise technical means such as forensic digital analysis, metadata verification, expert analysis and disclosure of the AI system that produced the AI-generated evidence to provide sufficient assurance of the credibility of any AI-generated evidentiary material produced by the AI system before allowing such evidentiary material to be admitted in any court proceeding.[20]

5.1 BURDEN OF PROOF AND RESPONSIBILITY

The use of AI-generated evidence raises complex issues surrounding the burden of proof and the accountability of those who create, utilise, or rely upon such evidence. Although evidentiary rules require the party introducing evidence to establish its relevance and reliability, AI-generated evidence adds a level of complexity because it may not always be possible to understand the rationale that produced the evidence.

Machine learning (ML) systems can yield inaccurate results as a result of programmed errors, bias in the training data, and invalid or faulty input. Thus, ML raises questions of reliability and accountability. The potential for wrongful identification presents us with an example of the problem presented by the use of AI facial recognition tools by Indian law enforcement officers. Wrongful identification has the potential to infringe on the rights of the individual, and it is unclear whether the developers, the deploying agency, or the end-users will be held accountable if an erroneous identification occurs.

Courts should require parties relying on AI-generated evidence to establish the reliability of the underlying technology and the safeguards adopted during its development and use. Such scrutiny would enhance accountability and reduce the risk of erroneous outcomes. These concerns become particularly significant with the emergence of deepfakes, synthetic media, and AI hallucinations.

5.2 DEEPFAKES, SYNTHETIC MEDIA, AND AI HALLUCINATIONS

[20]Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach* 28–35 (4th ed. 2021).

The advancement of AI has the capability to provide a greater level of evidence gathering, digital forensics research, and judicial resolution through quicker, more advanced analytical methods applied to large amounts of data. However, the emergence of deepfakes, synthetic media, and AI hallucinations creates serious challenges to the trustworthiness of digital evidence. Robert Chesney and Danielle Keats Citron both agree that the increasing sophistication of deepfake technologies creates serious risks to privacy, democratic institutions, and the integrity of information systems[21]. Additionally, as generative AI systems create hallucinations, they produce credible-looking but factually incorrect, misleading, or completely made-up information. For example, they can produce manipulated audio-visual content, fabricated legal authorities, and/or contradictory factual claims that can contribute to errant judicial outcomes. In light of these things, it is imperative that strong authentication mechanisms, adequate verification tools, and increased scrutiny from judges will assist in preserving the trustworthiness and integrity of AI-generated evidence.

Generative AI has many advantages when it comes to conducting legal research and analysing data, but it can produce authoritative answers that can be inaccurate due to hallucinations created by the AI. This increases the need for thorough verification of these AI-generated responses before using them as evidence in court proceedings. However, there is not always a clear way to verify these outputs, as a court may have no way to see the reasoning being used by an AI to produce an output. Furthermore, these issues illustrate the greater problem posed by black-box algorithms and the increasing demand for more transparency and explanation for AI-generated evidence.

5.3 BLACK-BOX ALGORITHMS AND EXPLAINABILITY

As discussions concerning AI-generated evidence continue to evolve, so too has explainability grown to become one of the most pressing issues. The principles that govern judicial systems are rooted in ideals of fairness, transparency, and procedural justice; all parties involved in adjudicating a matter should be afforded the right to fully comprehend and effectively contest evidence introduced against them. Currently, however, many sophisticated AI technologies are considered black-box models in that they produce outputs but do not provide any clear insight into their underlying reasoning capabilities.[22]

[21]Robert Chesney & Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 **Calif. L. Rev.** 1753, 1758–78 (2019).
[22]Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 Fordham L. Rev. 1085, 1088–1103 (2018).

The operation of many sophisticated AIs is like a "black box"; they yield results or important decisions, but do not leave much information about how they came to such conclusions. This makes it very difficult for lawyers, the judiciary or litigants to determine whether AI-generated results are valid and trustworthy. When there is no meaningful access to the methodology, training data or decision-making process of an AI, prosecutors, defenders, or plaintiffs may doubt whether the results will produce fair and accurate outcomes in a trial. In addition to providing judges with an understanding of how AI reached its conclusions, explainability will aid in the establishment of trust in the use of technical data by the judicial system and uphold the core principles of natural justice. However, the issues arising from the use of AI as an evidence collection tool cannot be resolved through transparency alone; therefore, as AI use continues to grow, comprehensive regulatory safeguards will need to be developed to control the authentication, evaluation, and admissibility of AI-produced evidence in court.

5.4 REGULATORY SAFEGUARD TO AI-DRIVEN EVIDENCE

Many of the current types of evidence used to evaluate whether to admit a piece of evidence at trial rely on classic parts of how we store things electronically. However, using a piece of evidence generated by an AI poses different challenges compared to prior methods used with traditional forms of digital evidence. As such, however, AI can create more efficient ways for investigatory agencies to process the investigation of crimes, conduct forensic analyses, and process cases in the judicial system using technology.

The Ethics Guidelines for Trustworthy AI created by the European Commission identify the key components of trustworthy AI as transparency, accountability, and human oversight[23]. The legal systems will need to have an evidentiary standard that is specifically designed to ensure responsible, transparent, and trustworthy use of evidence generated by an AI system in the judicial process while still taking into consideration the fairness and reliability of AI-generated evidence.

The increased reliance on artificial intelligence in the evidentiary process means that legal systems need to re-evaluate their existing laws regarding admissibility and reliability. The outputs produced by artificial intelligence pose significant challenges in terms of accountability, transparency and fairness, and cannot be addressed through the framework of

[23]High-Level Expert Group on Artificial Intelligence, European Commission, *Ethics Guidelines for Trustworthy AI* 13–22 (2019).

traditional electronic evidence, which is governed by the Bharatiya Sakshya Adhiniyam, 2023. There are some examples of judges across Australia taking steps in this direction; however, legal safeguards like independent oversight and verification may not be sufficient to address the challenges presented by the rapidly changing and global nature of artificial intelligence technology.

As a result, there is a substantial need for recommendations to develop international standards to support the use of artificial intelligence and related technology. One of the significant guiding norms is provided by the UNESCO Recommendation on the Ethics of AI, which contains fundamental principles such as accountability, transparency, fairness, and human oversight needed for AI systems to be trustworthy.[24]These principles provide a normative basis to determine the admissibility of artificial intelligence-produced evidence and go beyond existing domestic legal frameworks. Thus, it is necessary to undertake an analysis of the various countries developing their own standards as a way of creating a more cohesive and globally informed regulatory environment for the admissibility and evaluation of AI-produced evidence.

6. INTERNATIONAL APPROACHES TO AI-GENERATED EVIDENCE

As artificial intelligence (AI) systems become more prevalent in law and the courts, many jurisdictions have begun to create various kinds of regulatory frameworks regarding the use of AI systems in law. At this time, however, no jurisdiction has enacted a comprehensive law specifically addressing the admissibility of evidence produced by AI; however, there have been several important developments in the European Union, United States and United Kingdom that offer some helpful guidance with respect to how these jurisdictions are evolving their evidentiary standards for AI evidence.

Among all of these developments, the most significant is the AI Act adopted by the European Union which represents the first comprehensive legal framework for regulating AI systems anywhere in the world. The AI Act will adopt a risk-based regulatory framework for regulating AI and will require compliance with obligations designed to ensure a level of transparency, human oversight, data governance, record-keeping and accountability for high risk AI systems. In addition to that, the AI Act also presents some helpful concepts with respect to establishing standards for admissibility of AI-generated evidence by emphasizing the requirement for transparency of the outputs of AI systems as well as providing

[24]UNESCO, *Recommendation on the Ethics of Artificial Intelligence* ¶¶ 125–140 (2021).

mechanisms for documenting and creating the ability to trace back the inputs of the AI systems. Therefore, the AI Act creates additional reliability and verifiability for AI-generated evidence that will be presented in court so that the courts will have the ability to justify accepting or denying admissibility of such AI-generated evidence.[25]

In the **United States**, AI regulation will not be done using one Federal statute but will be based on various state levels, case law, ethical rules, and guidance from the agencies. Courts have reinforced the necessity of verifying any legally generated AI content from someone other than the AI, following the sanctions imposed on lawyers who submitted false AI-generated quotes to courts. Courts continue to enforce strict accountability requirements for professionals and require verification of accurate information that is created from AI technology through multiple levels of human verification[26].

The **United Kingdom** presents a reasonably pragmatic stance regarding the use of AI throughout the course of a case. It has published guidance from the court system which suggests that there could be positive implications for enhancing efficiency through AI-supported activities however there are also potential concerns regarding hallucination or bias, confidentiality breaches and inaccuracies. Each judge maintains personal responsibility for any AI-based evidence that he or she collects and utilises in making judicial mistakes.[27]

These international developments have shown that there is a trend at play: while AI might support the legal process, it cannot replace the need for human oversight, transparency, and accountability.

AI-generated data being accepted by the courts as electronic evidence has become commonplace in various parts of the world, yet there is no overall basis of determination for the courts on how to assess it. This has led to uncertainty about whether the data is authentic and reliable. Currently, courts generally assess evidence under the rules of evidence used only for traditional electronic documents (e.g. emails and word processing documents). With the growing prevalence of audio, video and still image data being generated by artificial intelligence upon invasion of venue prior to presenting evidence, this also increases the difficulty for courts in assessing whether the original source of the material was electronic or artificially edited. This disconnect between the technological advances in data generation and

[25]Regulation 2024/1689, of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689).
[26] See, e.g., Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023).
[27]Courts & Tribunals Judiciary, *Artificial Intelligence and the Courts and Tribunals* (2023).

the legal system is an area that will require formulating regulatory guidelines, especially evident in India's position today.

6.1 LEGAL GAPS EXIST IN REGULATING AI-GENERATED EVIDENCE IN INDIA

Evidentiary challenges have arisen in India as a result of the increasing incorporation of artificial intelligence (AI) into criminal investigations by law enforcement agencies. Although the Bharatiya Sakshya Adhiniyam, 2023 acknowledges that electronic records may be admissible as evidence, it does not specifically address whether or not material created or processed using autonomous AI systems will be included, thus leaving a doctrinal gap with respect to algorithmic outputs. ***Richard Susskind*** states that this lack of clear legal definition creates ambiguity with regard to classification; specifically, courts will have to determine if AI-created materials meet traditional criteria for relevance, reliability, and probative value (even though they were created by independent machine learning methods).[28]

Additionally, the current framework does not adequately address the authentication issues that are inherent in dealing with AI systems. AI outputs are generated based on a multitude of complex variables, including the quality of the datasets used, how the model was trained, the integrity of the algorithms, and the frequency with which systems receive updates; none of these variables are presently being evaluated pursuant to Indian evidentiary rules. Conversely, the ***National Institute of Standards and Technology (NIST)*** identifies continuous testing, validation, and risk management throughout the lifecycle of an AI system as critical components of trustworthy AI.[29] The lack of equivalent protections subverts the assessment of the reliability of AI-generated evidence in a judicial forum.

The evidentiary framework in India clearly has gaps with respect to AI-generated evidence. The two principal reasons for this are transparency and explainability issues as the majority of AI systems that exist are “black boxes” and doing so limits any meaningful review by judges.[6] The need for transparency and explainability regarding legitimate AI governance is recognised by the OECD’s AI Principles.[30] Furthermore, with no defined liability framework, there is uncertainty as to who has liability for inaccurate, biased, or manipulated AI output. With these evidentiary issues, it is evident that current legal standards are inadequate for

[28]Richard Susskind, *Online Courts and the Future of Justice* (2019).
[29]Nat’l Inst. of Standards & Tech., *AI Risk Management Framework 1.0* (2023).
[30]OECD, *OECD Principles on Artificial Intelligence* (2019).

addressing emerging technology risks and as a result of this, the adoption of legislative measures will be necessary. Therefore there is a clear need to establish a specific regulatory framework through which there are minimum standards established with respect to authentication, traceability, transparency, accountability and structured judicial review in order to achieve reliability and fairness when admitting into evidence AI-generated evidence.

7. RECOMMENDATIONS

As a result of AI's growing inclusion into evidentiary and investigative practises, AI has substantially raised difficulties for India's legislative and institutional responses. The BSA, 2023, is currently the governing framework; nonetheless, defining the term "AI-generated evidence" as an expressly defined type of electronic record, as well as distinguishing AI-generated evidence from conventional electronic records, is necessary to eliminate ambiguities about how admissible, classified and construed, as well as the evidentiary worth of AI-generated evidence.

Moreover, there must be a framework for authenticating AI-generated results. Courts must have the authority to compel the disclosure of particular relevant technical information about the AI system that produced the AI-generated evidence, including, but not limited to, the types of datasets used in training and testing the model and the forms of architecture used to produce the model's output, how the model was trained, and whether the model has been changed or updated since production. Additionally, requirements for mandatory forensic audits, the preservation of the associated metadata, and algorithmic traceability are essential to achieving the highest standards of AI-generated evidence reliability, authenticity, and verifiability.

India needs to establish guidelines that outline how courts should evaluate the validity of AI evidence. These guidelines will be based on principles such as transparency, explainability, accountability, and human oversight. The purpose of these guidelines is to create an environment in which the courts can rely upon evidence derived from black-box systems. A liability framework must also be created to assign liability among AI tool developers, deploying/commissioning agencies, and end-users according to their level of control over the use and outcome of the AI tool. This will prevent the diffusion of responsibility when errors, bias, or manipulation occur in the use of AI tools by unjustly shifting responsibility from one party to another. Additionally, ongoing training opportunities should be made available to judges, prosecutors, defence counsel, and investigators to enhance the technical knowledge of

these individuals regarding the use of AI tools in the justice system (for example: deepfakes, algorithmic bias, forensic uses). Finally, there is also an opportunity for the establishment of an independent regulatory or oversight body that would establish technical standards for AI tools used in evidential processes, conduct audits of these tools to ensure compliance with ethical and legal standards, and regulate the use of these tools to ensure due process, reliability, fairness, and transparency in judicial decision-making in India's justice system, thus facilitating the responsible introduction of AI to the Indian justice system.

8. CONCLUSION

The addition of AI technology in the field of evidence signifies the fundamental change-of-structure occurring in today's legal system. The courts are moving away from their reliance on oral testimonies, documented records, and expert testimony to using machine-generated evidence created and processed through algorithms. This shift reflects law and technology becoming more interconnected than they have ever been, as courts must now engage with and create meaning from vast digital datasets that often require AI-powered tools to create value from.

The overall findings indicate that AI has improved investigative capabilities in each area by providing increased efficiency, accuracy, and the ability to detect patterns, such as with digital forensics and the management of cases and evidence. The result is a significant decrease in the amount of human error while also allowing rapid processing of complicated data, thus providing law enforcement and the courts with the ability to address high-tech forms of crime. In addition to the benefits of using AI, there also exist significant legal issues regarding authenticity, reliability, transparency, explanation for the evidence, and ultimately the accountability of the use of AI for producing evidence. New challenges such as deepfakes, synthetic media, and the use of non-transparent algorithms have created even more doubt about the integrity of evidence.

The Bharatiya SakshyaAdhiniyam previously known as the Indian Evidence Act,1872), 2023 allows for electronic records to be admissible in a court of law, but it does not specifically create requirements for AI-produced or processed evidence, nor create standards by which AI evidence will be evaluated. This lack of regulation creates a normative gap when the evidence has been produced/processed with minimal human input.

The future of evidence law does not include machine intelligence for human judgment; instead, it needs a balanced approach where AI will be a tool that works under the

supervision of judges and that has some form of constitutional protection. As such, any potential use of AI-generated information to improve crime investigation efficiency and the analysis of evidence would require establishing clear legal standards for checking these new types of information prior to their use in the criminal justice system, such as legal foundation, authentication, transparency, accountability, and human oversight.