IS TRUTH AND INTEGRITY OF JUSTICE ON TRIAL IN THE AGE OF ARTIFICIAL INTELLIGENCE (AI)? RE-ASSESSING DIGITAL EVIDENCE IN THE CONTEXT OF AI-GENERATED EVIDENCE SUCH AS DEEPFAKES*

Abstract

The integration of Artificial Intelligence (AI) into legal systems presents profound challenges and opportunities, particularly the admissibility, reliability, and ethical use of AI-generated evidence. This paper explores the evolving intersection of AI technologies and the standards for digital evidence in Nigeria's legal framework. It also examines how existing laws accommodate, or fail to accommodate, the complexities introduced by AI-generated content. As AI becomes more embedded in surveillance systems, automated decision-making, legal research, and forensic tools, the evidentiary process faces new questions: Can AI-generated content meet the requirements for admissibility under the Nigerian Evidence Act? How do we safeguard against manipulation, hallucination, or opacity in algorithmic systems? What standards or protocols are necessary to authenticate and verify the accuracy of synthetic media or AI-generated evidence in legal cases? The study employs a doctrinal and comparative methodology, analysing Nigerian statutory laws (including the Evidence Act, 2011 (as amended in 2023) and the Nigeria Data Protection Act, 2023), recent judicial trends, and regulatory models from the United States, European Union, and international human rights instruments. The study also incorporates real-world examples and hypothetical scenarios, such as deepfakes, encrypted surveillance footage, and AI-generated court documents, to highlight the technical and legal challenges posed by digital evidence. Central issues addressed include the 'liar's dividend' (where genuine evidence is doubted due to the plausibility of fakes), the opacity of black-box algorithms, and the need for transparency, reliability, and expert validation in courtroom contexts. The researcher is of the view that Nigeria's current legal framework is ill-equipped to handle the rise of AI-generated evidence without targeted reforms. This work recommends a comprehensive strategy involving the enactment of AI-specific evidentiary legislation, or further amendments to the Evidence Act, and the establishment of verification protocols for contested digital content. Further proposals include judicial training, AI impact assessments for legal technology, and adherence to international data protection standards such as the General Data Protection Regulations (GDPR). Ethical concerns were also addressed through recommendations for updating the Rules of Professional Conduct for legal practitioners using AI tools. Ultimately, this research offers a forward-thinking guide for Nigerian legal institutions, integrating legal clarity, technical safeguards, and ethical accountability in upholding fairness, transparency, and resilience in this era of AI.

Keywords: Artificial Intelligence, Digital Evidence, AI-Generated Evidence, Deepfakes, Justice on Trial

1. Introduction

The advent of Artificial Intelligence (AI) has transformed numerous industries, including the legal sector. AI refers to the development of computer systems that can perform tasks that typically require human intelligence, such as learning, problem-solving, and decision-making. In the legal field, AI is increasingly integrated into tasks such as document review, legal research, and even outcome prediction. While these applications promise efficiency and precision, a more pressing concern is emerging: the potential reliance on AI-generated evidence in legal proceedings. AI technologies including machine learning, natural language processing, and facial recognition are being used to generate and analyse digital evidence, such as forensic reconstruction (recreating crime scenes) and analysing security footage. In some jurisdictions, AI tools have been used to flag potential suspects, generate automated reports, and evaluate the credibility of witness statements.[1] As these systems begin to shape what is accepted as 'evidence,' they raise critical questions about truth, authenticity, and admissibility within the court proceedings. In Nigeria, these developments have implications for the provisions of Section 84 of the Evidence Act, 2011 (as amended 2023), which outlines the requirements for admitting electronically generated evidence, including proof that the computer was operating properly, and the information was supplied in the ordinary course of activities. However, AI-generated content, often produced by complex, multi-source systems with limited human supervision, may struggle to satisfy these statutory conditions. The black-box nature of many AI tools, where the logic behind outputs is not easily interpretable, raises serious concerns about compliance with existing evidentiary safeguards.[2] Unlike traditional forms of evidence, which are subject to known rules of admissibility and procedural safeguards such as cross-examination and authentication, computer generated evidence including AI-generated outputs can be difficult to verify or challenged. This creates practical difficulties for judges and legal practitioners in determining whether such evidence meets legal standards. It also raises deeper normative concerns: How can courts ensure accountability when AI-generated content contributes to wrongful convictions or misleading conclusions?

Although Section 84 of the Evidence Act was introduced to modernize evidentiary rules and accommodate electronically produced records, it was primarily designed with conventional digital outputs in mind, such as emails or banking records, where human involvement is traceable. AI, by contrast, produces autonomous outputs that may lack clear human input or a verifiable process. Moreover, the data that informs AI systems may reflect entrenched historical biases, incomplete records, or inaccurate

By David Chukwuebuka MKPO LLB (Hons), LLM, BL, Lecturer, Department of International Law and Jurisprudence, Nnamdi Azikiwe University, Awka Anambra State, Nigeria. dc.mkpo@unizik.edu.ng +2347066347477

[1] Maddox, Ian. 'Artificial Intelligence in the Courtroom: Forensic Machines, Expert Witnesses, and the Confrontation Clause.' *Case W. Res. JL Tech. & Internet* 15 (2024): 416.

[2] Hassija, Vikas, et al. 'Interpreting black-box models: a review on explainable artificial intelligence.' *Cognitive Computation* 16.1 (2024): 45-74.

cultural assumptions. When used without transparency or scrutiny, such data can distort rather than clarify the truth.[3] Even before reaching the courtroom, AI-generated summaries embedded in legal search engines or digital platforms may influence how researchers and lawyers understand legal concepts, often without them realizing the source or structure of that information.

This paper explores the critical question: Is the truth and integrity of justice on trial in the age of AI-generated evidence? It focuses on the implications of integrating AI into evidentiary processes, particularly within Nigeria's legal system, where statutory frameworks such as Section 84 appeared not to be designed with AI in mind. The aim is not to discredit AI's potential, but to interrogate the risks it poses when its outputs are treated as objective or admissible without adequate legal scrutiny. The objectives of this research are to examine the benefits and dangers of AI-generated evidence in justice systems, with emphasis on bias, admissibility, and accountability. It draws on existing research, case studies, and legal frameworks to examine the interaction of these technologies with core principles such as fairness, due process, and the presumption of innocence. This study is significant because it explores a rapidly changing field where technology and law often conflict. While AI offers opportunities for enhancing investigation and decision-making, its use in evidence must be approached with caution. This is because without clear guidelines and strict scrutiny, there is a risk that AI will erode rather than enhance the credibility and fairness of justice. As the legal sector grapples with the implications of digital innovation, it has become important to understand how AI technologies impact evidence, legal interpretation, and judicial decisions. This paper examines how AI-generated evidence may influence and potentially challenge the delivery of fair and credible justice.

2. Conceptual Clarifications

Artificial Intelligence

Artificial Intelligence (AI) has been defined in various ways depending on the disciplinary lens through which it is viewed. Broadly, AI refers to the capability of machines or computer systems to perform tasks that would typically require human intelligence. These tasks include reasoning, problem-solving, learning, perception, and language understanding. According to John McCarthy, one of the founding figures in the field, AI is 'the science and engineering of making intelligent machines' capable of simulating human cognitive processes.[4] This definition emphasizes the scientific ambition to replicate human thought through artificial systems.[5] Stuart Russell and Peter Norvig discuss AI in terms of four categories: thinking humanly, thinking rationally, acting humanly, and acting rationally. They describe AI as intelligent agents that perceive their environment and take actions to achieve goals.'[6] The European Commission defines AI more pragmatically as 'systems that display intelligent behaviour by analysing their environment and taking actions, with some degree of autonomy, to achieve specific goals.' This definition is particularly relevant for governance, including regulatory and ethical discussions, in the legal field.

In the legal context, AI is often understood as software or systems used to assist, augment, or automate legal functions. This includes tools for legal research, predictive analytics, risk assessment, contract review, and increasingly, the generation or analysis of evidence. AI can generate evidence in various forms, such as synthetic data, which enables it to create artificial data that mimics real-world data, like AI-generated images that resemble real-world photos. AI technologies are also used to produce reconstructed crime scenes and automated reports which may eventually be tendered in court. Legal researchers, too, increasingly rely on AI-powered platforms and summarization tools to access case law and legal articles, but these systems may reflect biased training data, fabricate authorities, or distort complex legal doctrines. When accepted without proper scrutiny, such content can mislead legal practitioners and erode both the accuracy and integrity of legal reasoning. Some scholars view legal AI as 'algorithmic systems that contribute to decision-making or legal reasoning, based on data-driven inputs and logical rules.'[7] When used in the generation of legal evidence or courtroom decision-making, AI challenges traditional notions of human judgment, transparency, and accountability. Understanding its nature, capabilities, and limitations is essential for evaluating its compatibility with fundamental legal principles such as fairness, truth, and integrity.

Truth and Integrity

In the legal field, the concept of truth extends far beyond mere factual correctness. It lays more emphasis on authenticity, verifiability, and procedural fairness. Legal truth is not simply about what happened, but about what can be established through evidence and due process.[8] It is rooted in standards that ensure the information presented is both credible and lawfully obtained. In court proceedings, truth is often arrived at through adversarial testing of evidence, governed by established rules of admissibility and guided by the presiding judge. This approach ensures that conclusions are based on information that can be scrutinized and verified. As such, verifiability which is the ability to confirm the origin, accuracy, and reliability of evidence, is a foundational aspect of legal truth.[9] In the context of computer-generated evidence, this raises concerns about hyper-realistic videos created through AI, which could be mistaken for real-world videos. This is more pressing when the decision-making

[3] Emma, Lawrence. 'The Ethical Implications of Artificial Intelligence: A Deep Dive into Bias, Fairness, and Transparency.' (2024).

[4] John McCarthy, 'What is Artificial Intelligence?' Computer Science Department, Stanford University (Revised November 12, 2007) <https://www-formal.stanford.edu/jmc/whatisai.pdf> accessed 20th June, 2025

[5] El Samad, Mahmoud, Ghalia Nasserddine, and Ahmad Kheir. *Introduction to artificial intelligence.' Artificial intelligence and knowledge processing.* CRC Press, 2023. 1-14.

[6] Russell, Stuart J., and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 2016)

[7] Harry Surden, 'Artificial Intelligence and Law – An Overview of Recent Technological Changes in Large Language Models and Law' *96 Colorado Law Review pp. 376 – 411 (2025)* <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5135305> accessed 24th June, 2025

[8] Summers, Robert S. 'Formal legal truth and substantive truth in judicial fact-finding—their justified divergence in some particular cases.' Law and philosophy 18 (1999): 497-511.

[9] Patterson, Dennis Michael. *Law and truth*. Oxford University Press, 1999.

process or data source of an algorithm cannot be fully explained, understood, or reproduced, a challenge often referred to as the 'black box' problem. Authenticity, another important aspect of truth, refers to the genuineness and integrity of evidence. Under Nigerian law, for instance, Section 84 of the Evidence Act 2011(as amended) requires parties to establish that electronic or computer-generated documents are authentic before they can be admitted in court. This includes proving that the computer was operating properly and that the document was generated in the ordinary course of use. AI-generated outputs, particularly those from autonomous or composite systems, may pose significant hurdles in satisfying this requirement, especially where the chain of custody, data sources, or processing logic are opaque or unverifiable.[10]

Integrity, closely tied to truth, relates to the moral and procedural backbone of justice. In legal terms, integrity reflects consistency, impartiality, and strict adherence to both ethical and legal standards, especially in the handling, admission, and evaluation of evidence. A legal system with integrity ensures that processes are transparent and accountable, that all parties are treated fairly, and that outcomes are not distorted by bias, manipulation, or undue influence. When AI is introduced into justice systems, particularly in evidence creation or interpretation, these values come under strain. Any legal system seeking to incorporate AI must ensure that these values are preserved, particularly where the evidence in question plays a decisive role in determining guilt, liability, or rights.

Justice

Justice is a foundational principle of every legal system, which encompasses fairness, equality, and the proper administration of the law. At its core, justice involves not only the fair resolution of disputes and protection of rights but also the preservation of public confidence in legal institutions.[11] In this sense, justice is both a procedural and a moral concept which ensures that the legal system operates impartially, transparently, and with respect for human dignity. In the context of this study, justice is closely tied to how truth is uncovered and presented, particularly through the handling of evidence. The legitimacy of any judicial decision depends heavily on the integrity of the fact-finding process and adherence to due process. As AI systems significantly influence how evidence is generated, analysed, and interpreted, they may affect core elements of justice. The title of this paper uses the phrase 'on trial' not in the narrow sense of court proceedings, but as a metaphor for scrutiny and pressure. The phrase implies that the values of truth and integrity, which underpin justice, are being tested and re-evaluated in the face of emerging technologies. The rise of AI challenges traditional assumptions about what constitutes reliable evidence, who holds responsibility for errors, and whether fair legal processes remain obtainable in a digital age. This metaphorical framing provides the lens through which this paper interrogates the evolving relationship between AI and justice.

AI-Generated and Computer-Generated Evidence

The Evidence Act defines a computer as any device for storing and processing information, where outputs are derived from input data through processes like calculation, comparison, or other automated operations. Consequently, computer-generated evidence refers to evidence generated through these automated processes.[12] On the other hand, AI-generated evidence refers to evidence created or processed using artificial intelligence technologies, such as machine learning algorithms, that can generate, analyse, or manipulate data.[13] This type of evidence is a subset of computer-generated evidence, as it relies on automated systems. The outputs of these systems, which can include automated analytical reports, facial recognition matches, predictive assessments, and reconstructed crime scenes, highlight the complexity and versatility of computer-generated and AI-generated evidence.[14] However, the increasing reliance on such technologies raises significant concerns regarding authenticity, accuracy, and accountability, not only in terms of the evidence itself, but also the processes through which it is generated. AI systems often function as black boxes, meaning their internal logic is obscure and difficult to interpret or explain, even by their developers. This lack of transparency complicates efforts to verify how certain outputs were produced. Additionally, bias may be integrated in the training data or algorithms, which can lead to outputs that reflect or enhance existing social or institutional prejudices.[15] Understanding these concepts is essential for evaluating how AI technologies may shape the future of legal evidence and its role in the justice system.

Deepfakes and Synthetic Media

Synthetic media refers to digital content, such as images, videos, or audio, that is generated or manipulated using AI or machine learning techniques.[16] Unlike the conventional media captured from real-world events, synthetic media is either entirely artificial or significantly altered to convey something that did not actually occur. Common forms include deepfakes, which are AI-generated videos or audio clips that make it appear as though someone said or did something they never did, and AI-generated images, which are realistic visuals created by AI without being based on specific real people or existing footage. While synthetic media has legitimate applications in fields such as film production, advertising, and academic research, its increasing sophistication has raised urgent concerns in the legal sphere. Deepfakes in particular, present serious risks when

[10] Singhal, Amisha. 'Social Challenges of AI Governance.' <https://burnishedlawjournal.in/wp-content/uploads/2024/05/Social-Challenges-of-AI-Governance-by-Amisha-Singhal-.pdf> accessed 1st July 2025.

[11] Pound, Roscoe. 'Justice according to law.' *Colum. L. Rev.* 13 (1913): 696.

[12] Evidence Act, 2011 (as amended 2023) s.258

[13] Grimm, Paul W., Maura R. Grossman, and Gordon V. Cormack. 'Artificial intelligence as evidence.' *Nw. J. Tech. & Intell*. Prop. 19 (2021): 9.

[14] Cardenuto, João Phillipe, et al. 'The age of synthetic realities: Challenges and opportunities.' *APSIPA Transactions on Signal and Information Processing* 12.1 (2023).

[15] Von Eschenbach, Warren J. 'Transparency and the black box problem: Why we do not trust AI.' *Philosophy & Technology* 34.4 (2021): 1607-1622.

[16] Millière, Raphaël. 'Deep learning and synthetic media.' *Synthese* 200.3 (2022): 231.

introduced into evidentiary processes. They can be used to fabricate false evidence, such as doctored security footage, fake confessions, or altered witness recordings.[17] These false representations may appear convincingly real and can mislead investigators, influence judicial reasoning, or wrongfully sway public opinion. The growing prevalence of such content raises complex legal questions: How can courts detect and verify manipulated media? Who bears the burden of proving its authenticity or falsity? And can traditional methods ensure the integrity of digital evidence when it can be easily altered?

These issues undermine truth and reliability in judicial proceedings. Courts and law enforcement agencies often lack the forensic tools or technical expertise needed to confidently distinguish between real and manipulated content. As a result, the legal system risks admitting or relying on deceptive material, which may affect both procedural fairness and public trust. This concern is particularly pressing in jurisdictions where evidentiary rules were developed without anticipating the emergence of such technologies. Thus, the emergence of deepfakes as a category of AI-generated content is not merely a technological development, but a **fundamental change t**hat challenges existing legal principles in obtaining evidence.[18]

3. Legal and Institutional Frameworks

Legal Framework

Constitution of the Federal Republic of Nigeria 1999 (as amended 2023): The Constitution provides the foundation of the legal authority for all laws in the country, including those regulating evidence and the administration of justice. Section 1(1) declares the Constitution as supreme, which is binding on all persons and authorities, while subsection (3) establishes that any other law inconsistent with its provisions shall be void to the extent of the inconsistency. This supremacy clause is critical when evaluating the admissibility and use of AI-generated or synthetic evidence such as deepfakes. If AI-generated evidence or related legal processes violate constitutional principles, such as the right to fair hearing or right to privacy, they may be deemed unconstitutional regardless of their alignment with statutory laws. This also emphasises the need to ensure that emerging technologies align with fundamental rights guaranteed by the Constitution. Section 36 of the Constitution, which guarantees the right to a fair hearing within a reasonable time before a court or tribunal that is independent and impartial, is the provision that supports the integrity of the justice system and assumes that parties will have the opportunity to challenge evidence presented against them. The use of AI-generated evidence, particularly when derived from vague or unverifiable systems, poses a unique challenge to this right. Deepfakes and automatically generated forensic evidence can influence judicial decisions while evading thorough examination due to their complex nature, thereby undermining the fairness of the proceedings. Furthermore, Section 36(2) allows administrative authorities to make decisions affecting civil rights only where affected persons have had the opportunity to make representations. This procedural safeguard is difficult to maintain where evidence is shaped by automated systems that cannot be cross-examined or fully explained. Section 39 of the Constitution, which guarantees the right to freedom of expression, also has implications for AI-generated content. It affirms the right to hold opinions, receive, and impart information without interference. However, the same technologies that facilitate this freedom, such as generative AI tools, can also be used to produce and circulate false content. Deepfakes, for instance, may be used to fabricate evidence that appears convincingly real but is fundamentally deceptive. While subsection (3) of Section 39 permits restrictions on speech necessary to maintain the authority and independence of the courts, striking the right balance remains difficult in practice. Courts must weigh the need to protect due process and the integrity of trials against the constitutional commitment to open access to information. In addition, the Constitution reserves the regulation of evidence exclusively for the federal government. The Second Schedule, Part I of the Constitution places 'evidence' under the Exclusive Legislative List, which means only the National Assembly is empowered to legislate on matters relating to the admissibility and handling of evidence in Nigeria. This provision limits the ability of individual states to independently respond to emerging challenges posed by synthetic and AI-generated content, and underscores the importance of a unified national framework. Given that existing law on evidence, that is the Evidence Act, was not designed with AI or deepfakes in mind, the need for legislative reform at the federal level becomes even more pressing. Collectively, these constitutional provisions serve as a reminder that while technological innovations such as artificial intelligence offer new tools for legal investigation and information sharing, they must not undermine the fundamental guarantees of fairness, privacy, transparency, and accountability enshrined in the Constitution. Any reliance on AI-generated evidence within Nigeria's legal system must therefore be carefully assessed to ensure full compliance with these constitutional standards.

Evidence Act 2011 (as amended 2023): The 2011 revision of the Nigerian Evidence Act marked a significant shift in the country's legal response to digital innovation, particularly with respect to the admissibility of electronic and computer-generated evidence. Prior to this amendment, electronically generated documents were not expressly recognized, making their admissibility highly problematic. The reform was therefore crucial, as it expanded the definition of a 'document' to include materials produced by electronic devices, enabling the admission of digital records in court proceedings. Under Section 84 of the Evidence Act (as amended in 2023), the law provides detailed conditions under which electronic records can be admitted as evidence. Section 84A recognizes that information required by law to be in writing may be rendered electronically, provided it is accessible for future reference. Section 84B extends this by deeming information stored or copied on electronic, magnetic, optical media, or cloud computing platforms admissible as documentary evidence, if certain conditions are satisfied. These conditions are particularly stringent and reflect the importance of authenticity and reliability in the admissibility of digital

[17] Pfefferkorn, Riana. 'Deepfakes' in the Courtroom.' *BU Pub. Int. LJ* 29 (2019): 245.
[18] Wang, Yifei. *Synthetic realities in the digital age: Navigating the opportunities and challenges of ai-generated content.* Authorea Preprints (2023).

evidence. Section 84(2) of the Evidence Act sets out four cumulative requirements that must be satisfied before a computer-generated document can be admitted as evidence. First, it must be shown that the document was produced by a computer during a period when it was regularly used to store or process information. Second, the kind of information contained in the document must have been regularly fed into the computer in the ordinary course of the activities of the person or body responsible for its use. Third, the computer must have been operating properly throughout the relevant period, or, if not, any malfunction must not have affected the accuracy of the information contained in the document. Lastly, it must be established that the information reproduced or derived from the computer was supplied to it in the ordinary course of its activities. These requirements ensure that electronic records are not only generated by a functioning and reliable system but also reflect data that was regularly and appropriately entered. These requirements are complementary, meaning that failure to satisfy any one of them renders the electronic evidence inadmissible. This was affirmed in *Kubor v. Dickson*,[19] where the Supreme Court held that mere tendering of an electronic document from the bar is insufficient; the party must lay the proper evidentiary foundation in accordance with Section 84(2). The amended Act also introduced digital authentication mechanisms. Section 84C allows a person to authenticate electronic records using digital signatures, provided they are reliable. Reliability is determined by two key factors: the signature data must be exclusively linked to the signatory, and any alterations made after authentication must be detectable. This ensures the digital signature's authenticity and integrity. Section 84D provides that digital signatures must be proven genuine unless they qualify as secure signatures – those that meet specific requirements ensuring authenticity and integrity – in which case they're presumed authentic. Section 258 of the Act provides definitions crucial to understanding how these provisions operate. A computer is broadly defined as any device capable of storing and processing information, including mobile phones. Terms like digital signature, electronic signature, electronic record, and audio-visual communication are also defined, reflecting the diverse technologies now relevant to legal processes.

Importantly, the use of audio-visual records as evidence is particularly relevant in the context of AI-generated and manipulated media. Section 15(4) of the Administration of Criminal Justice Act (ACJA) complements the Evidence Act, by permitting confessional statements to be electronically recorded through video or other audio-visual means. Similar provisions appear in several state laws. This is notable in the context of deepfakes or synthetic media, as the legal system must now grapple with distinguishing genuine audio-visual confessions from convincingly fabricated ones. Given the complex nature of AI-generated content, many of these outputs, such as predictive assessments or facial recognition matches, may be derived from autonomous systems that collect data from multiple sources. These systems often operate without clear human supervision or traceable processes, and may not meet Section 84(2)'s requirements for admissibility of computer-generated evidence, given concerns about their operation, input, and reliability. Moreover, the digital signature framework, while promising, may be difficult to apply in practice when dealing with AI-generated evidence that lacks a clearly attributable author or originator. Where no identifiable human input exists, questions arise about how to satisfy the legal requirement of authorship or how to authenticate such evidence in line with statutory conditions. In essence, the Evidence Act, particularly Sections 84A–84D and 258, reflects Nigeria's attempt to modernize evidentiary standards in light of technological advancement. However, these provisions were primarily designed for conventional computer-generated evidence, not the autonomous and often opaque processes involved in AI systems. As such, while the Act provides a necessary foundation, it may require further reform or interpretative clarity to address the unique challenges presented by AI-generated and manipulated digital evidence.

Nigeria Data Protection Act 2023: While Nigeria has not yet enacted a specific law dedicated exclusively to AI, the Nigeria Data Protection Act, 2023 (NDPA) offers a foundational legal framework that governs the use of AI systems, particularly where such systems involve the processing of personal data. The NDPA adopts a risk-based approach to data processing, and this applies equally to AI technologies deployed for data analysis, decision-making, or automation. Significantly, Section 37 of the NDPA places limits on the use of automated decision-making, including profiling. It provides that data subjects shall not be subjected to a decision based solely on automated processing of personal data, especially where such decisions produce legal or similarly significant effects, unless there is meaningful human intervention, and the logic behind the decision is explainable and contestable. This provision is critical in the context of AI-driven decision-making, as it aligns with international standards on algorithmic transparency, fairness, and the right to explanation. In addition, Section 29 of the NDPA imposes obligations on data controllers and processors to implement appropriate technical and organizational measures to ensure the confidentiality, integrity, and availability of personal data. In practice, this may involve the adoption of AI-based security tools for tasks such as anomaly detection, data encryption, and risk prediction. Thus, even in the absence of a comprehensive AI statute, developers of AI systems in Nigeria must ensure their activities are in compliance with the NDPA's requirements on data protection and accountability. Therefore, until a dedicated AI regulatory framework is enacted, existing laws like the NDPA serve as the primary legal instruments guiding the creation, deployment, and use of AI systems in Nigeria, particularly in contexts involving personal data. Compliance with these provisions is essential to uphold data privacy rights, ensure algorithmic fairness, and mitigate risks associated with the use of AI technologies.

Cybercrimes (Prohibition, Prevention, etc.) Act 2015 (as amended in 2024: The Cybercrimes Act plays a vital role in regulating digital conduct and safeguarding the integrity of computer systems in Nigeria. In the context of AI-generated evidence and synthetic media, several provisions of the Act directly engage with the risks posed by manipulated or fraudulent digital content in legal proceedings. Section 8 prohibits intentionally disrupting computer systems by altering, deleting, or suppressing data, which can compromise its intended operation. This applies to AI-generated content and deepfakes, particularly when manipulated to deceive. The law addresses data manipulation and system disruption used to mislead. This provides a framework to tackle challenges posed by AI-generated content and deepfakes. Section 13 further underscores the

[19] (2014) 4 NWLR (Part 1345) 534-594

legal consequences of inputting or altering data to produce inauthentic outputs that are passed off as genuine. This resonates with the concern that AI tools can fabricate seemingly reliable evidence, such as doctored confessions, manipulated videos, or fake digital footprints, potentially misleading courts and investigators. Relatedly, Section 14 on computer-related forgery and fraud addresses situations where data manipulation leads to the wrongful acquisition of property or economic benefit. These provisions establish AI-generated deepfakes as potential criminal acts with serious legal implications, rather than just technical issues. Another relevant provision is contained in Section 17, which legally recognizes electronic signatures and sets standards for their authenticity. As AI can now forge digital signatures or simulate user behaviour, the requirement of proving the genuineness of such signatures becomes crucial in maintaining the integrity of electronic records. Moreover, Section 45 empowers law enforcement to obtain search warrants for digital evidence, enabling them to access AI-manipulated content and background data during investigations. This is essential in cases where the origins or authenticity of digital evidence, such as AI-generated media, are in question. Further, Section 28 prohibits the production or distribution of software or devices intended to breach digital security, many of which may enable the creation or dissemination of deepfakes. Similarly, Sections 38 to 40 establish obligations for service providers to retain and disclose data when requested by law enforcement, thereby preserving logs or traffic data that could be useful in tracing the creation or spread of harmful AI-generated content. Together, these provisions demonstrate that while Nigeria's cybercrime legislation was not specifically drafted with AI in mind, it provides a flexible legal framework for addressing key risks posed by synthetic media and AI-based interference. Still, the sophistication of AI technologies continues to challenge the practical application of these laws, particularly in evidentiary settings where verifying authenticity, determining authorship, and meeting the burden of proof pose significant hurdles. The Act, though proactive, must be interpreted and perhaps further developed in line with evolving threats to ensure justice is not compromised in an increasingly digital age.

Recent Policy Developments on AI in Nigeria: In the absence of AI-specific legislation, Nigeria has initiated efforts to establish a dedicated policy and regulatory framework for AI. In 2022, the National Information Technology Development Agency (NITDA) initiated consultations with stakeholders to develop a National Artificial Intelligence Policy (NAIP), which led to the development of its first draft in March 2023.[20] Extending this progress, the Federal Ministry of Communications, Innovation and Digital Economy (FMCIDE) released a white paper in August 2023, setting the groundwork for a more expansive National Artificial Intelligence Strategy (NAIS).[21] By August 2024, the draft of the NAIS was officially published, outlining Nigeria's vision to harness AI ethically and responsibly while positioning the country as a global player in the AI space.[22] The Strategy identifies four core areas of risk: economic, ethical, societal, and AI model risks. To guide its response to these risks, it adopts the Identify, Assess, Mitigate, Monitor, and Review procedure from the U.S. National Institute of Standards and Technology (NIST) Framework for AI Risk Management. Implementation challenges such as inadequate funding, limited technical capacity, ethical concerns, and rapid technological evolution were also acknowledged.[23] Further reinforcing Nigeria's commitment to responsible AI adoption, the Nigerian Bar Association (NBA) released Guidelines on the Use of Artificial Intelligence in the Legal Profession in September 2024.[24] These Guidelines emphasize ethical compliance, human oversight, transparency, and data privacy in the application of AI by legal practitioners. Additionally, Nigeria's signing of the Bletchley Declaration on AI in 2023, alongside 27 other countries including the United Kingdom and France, signals its commitment to international cooperation in mitigating AI-related risks.[25] Collectively, these developments mark a progressive shift toward a comprehensive national AI governance framework, complementing existing legal instruments such as the NDPA and laying the foundation for future legislative action.

European Union General Data Protection Regulation (GDPR): The European Union (EU) General Data Protection Regulation (GDPR)[26] is a vital legal instrument which regulates how artificial intelligence (AI) interacts with personal data within the EU. Though it is primarily a data protection regime, the GDPR has significant implications for the integrity of digital evidence, investigation in judicial processes, and the governance of AI-generated content. Under Article 5, the GDPR establishes foundational principles of data processing, including lawfulness, fairness, transparency, accuracy, integrity and confidentiality. These principles apply directly to AI systems that process personal data for forensic, investigative, or evidentiary purposes. In the context of digital evidence, this means AI must not collect or manipulate data beyond what is necessary for its stated purpose, and it must process such data in ways that can be independently verified and audited, these are essential components for ensuring authenticity of digital evidence in court. Article 6 lays out the legal basis for data processing. Notably, it requires explicit, informed, and unambiguous consent where AI systems process sensitive data. Alternatively, processing may rely on legitimate interests, but only where such interests do not override the rights of the individual (the data

[20] Ema Obe, Chinelo Obiekwe, 'Artificial Intelligence And Nigerian Data Protection' (August, 2024) <https://www.mondaq.com/nigeria/privacy-protection/1505422/artificial-intelligence-and-nigerian-data-protection#authors> accessed 2nd July, 2025.

[21] White & Case, LLP, 'AI Watch: Global regulatory tracker – Nigeria' (January, 2025) <https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-nigeria> accessed 2nd July, 2025.

[22] Duale, Ovia & Alex, 'Pillars of Nigeria's Draft National Artificial Intelligence Strategy' (September 2024) <https://www.doa-law.com/wp-content/uploads/2024/09/Pillars-of-Nigerias-Draft-National-Artificial-Intelligence-Strategy.pdf> accessed 1st July, 2025

[23] U S Department of States, 'Risk Management Profile for Artificial Intelligence and Human Rights' (July 2024) <https://2021-2025.state.gov/risk-management-profile-for-ai-and-human-rights/> accessed 1st July, 2025

[24] Nigerian Bar Association, 'Guidelines for the use of artificial intelligence in the legal profession in Nigeria' (2024) <https://nbaslp.org/wp-content/uploads/2024/04/Guidelines-for-the-Use-of-Artificial-Intelligence-in-the-Nigerian-Legal-Profession.pdf> accessed 2nd July, 2025

[25] AI Safety Summit 2023: The Bletchley Declaration (February, 2025) <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> accessed 2nd July, 2025

[26] Regulation (EU) 2016/679

subject) whose personal data is being processed. In judicial contexts, this becomes vital: if digital content, such as surveillance data, online profiles, or location metadata, is used as evidence, it must be proven that it was lawfully obtained and that its use does not infringe the privacy or procedural rights of the parties involved. The requirement for informed consent imposes a significant duty on those processing personal data, which can help prevent manipulation of data, including through deepfakes, as they must demonstrate compliance with these requirements. Importantly, Article 22 directly addresses automated individual decision-making, including data-driven evaluation. It gives individuals the right not to be subject to decisions based solely on automated processing where such decisions produce legal or similarly significant effects. This provision is particularly relevant when AI tools are used to flag suspects, generate risk levels, or assess guilt based on behavioural patterns or online data. The provision demands that such processes involve human supervision and that affected persons are given the opportunity to contest or understand the basis of the decision. This is reinforced by Recital 71, which warns that automated decision-making, especially where it includes profiling, should not lead to discriminatory or unjust outcomes. It further stresses that such systems must include safeguards such as the right to human intervention, the ability to express one's point of view, and contest the decision. These safeguards are essential for preserving the truth and fairness of AI-generated or AI-influenced evidence. Additionally, Article 15(1)(h) enhances the transparency requirement by giving data subjects the right to obtain meaningful information about the logic and functioning of automated systems. In legal proceedings, there is a need for courts, parties, and counsel to understand how an AI system processed information and reached its conclusions, particularly where such outputs are relied upon in proving or disproving facts. Without this clarity, digital evidence generated or analysed by AI risks being incomprehensible or unverifiable, thereby undermining its probative value. The GDPR also enshrines the principle of 'data protection by design and by default' under Article 25, which mandates that systems, such as those generating or handling evidence, must be built with privacy and security as their fundamental features. This provision supports the development of forensic AI tools that preserve data authenticity and limit tampering or unauthorized manipulation, including the risks posed by deepfakes or AI-edited content. To ensure accountability, Article 24 requires controllers of personal data (including AI developers and institutions using AI systems) to demonstrate compliance with all GDPR principles. This includes maintaining detailed records, conducting internal audits, and showing that appropriate technical and organisational measures are in place to prevent misuse. When dealing with high-risk processing, such as surveillance or predictive policing, Article 35 mandates a Data Protection Impact Assessment (DPIA) to assess risks to individual rights and identify measures to minimize those risks.

These provisions, taken together could offer a strong legal foundation for assessing the reliability, fairness, and admissibility of AI-generated evidence in judicial processes. The GDPR underscores the importance of transparency, human accountability, and regulatory safeguards, all of which are essential for maintaining the truth and integrity of justice in the digital age. While not directly enforceable in Nigeria, the GDPR sets a global benchmark, offering critical lessons for jurisdictions grappling with the evidentiary and ethical challenges posed by artificial intelligence.

UNESCO Recommendation on the Ethics of Artificial Intelligence (2021): The United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation on the Ethics of Artificial Intelligence, adopted by 193 member states in November 2021,[27] represents the first global standards specifically aimed at guiding the ethical development and deployment of AI. Although not legally binding, the Recommendation carries significant persuasive authority and has been instrumental in shaping national AI regulations worldwide. Its relevance to digital justice systems, especially in the age of deepfakes and AI-generated evidence, lies in its commitment to protecting human rights, preserving human dignity, and ensuring transparency and accountability in AI systems. These principles speak directly to the foundational elements of fair trial rights, evidentiary integrity, and procedural justice. One of the Recommendation's core concerns is the potential of AI systems to exacerbate discrimination or distort truth.[28] It explicitly calls for the mitigation of algorithmic bias and the implementation of effective mechanisms to ensure that automated systems do not replace human judgment in critical decision-making areas such as the judiciary. This emphasis on human oversight and algorithmic transparency provides a global ethical baseline for ensuring that AI-generated evidence, such as manipulated audio-visual material or predictive analytics, does not unjustly influence judicial outcomes or undermine public trust in the legal system. Furthermore, the Recommendation warns against the use of AI for mass surveillance or rating system, this reinforces the need to protect privacy rights and prevent unlawful or invisible forms of digital monitoring. It also aligns with the growing challenge of AI-generated content being used to profile or mislead individuals, particularly in legal contexts where electronic records can be altered or misrepresented without clear traceability. As Nigeria and other jurisdictions grapple with integrating AI into evidentiary frameworks, UNESCO's guidance helps protect the integrity of justice systems as technology advances.

EU Artificial Intelligence Act 2024: The EU Artificial Intelligence Act which came into force on August 1, 2024, is the world's first comprehensive legal framework on AI. The Act establishes a regulatory system for the use of AI systems in EU Member States, prioritizing safety, transparency and accountability. It is the first legal regime globally to regulate AI in a risk-based, proportionate manner. It classifies AI systems into four categories: unacceptable risk, high risk, limited risk, and minimal risk, each with corresponding legal obligations designed to protect fundamental rights, democracy, and the rule of law. This Regulation aims to balance innovation with protection, fostering trustworthy AI that safeguards health, safety, fundamental rights, and the environment while promoting the internal market.[29] AI systems deemed to pose an unacceptable risk are

[27] Evelyne Para, 'The UNESCO Recommendation on the Ethics of Artificial Intelligence' (July 2024) <https://www.soroptimistinternational.org/2024/07/18/the-unesco-recommendation-on-the-ethics-of-artificial-intelligence/> accessed 2nd July 2025

[28] Ahuerma, Fabio Morandín. 'UNESCO Proposal for the use of Generative AI in Education: Eight Challenges and Seven Actions.'

[29] EU AI Act, art 1

prohibited under Article 5, including those that manipulate human behaviour through subliminal or deceptive techniques, exploit vulnerabilities of specific groups or individuals, engage in social scoring that leads to detrimental treatment, predict criminal risk based solely on profiling or personality traits, collect facial images data without specific purpose or consent, or infer emotions in workplaces and education institutions without medical or safety reasons. These prohibitions protect citizens against automated systems that could undermine autonomy, privacy, and dignity, which are principles that are essential in maintaining integrity in justice systems. For high-risk AI systems, such as those deployed in law enforcement, judicial decision-making, biometric identification, or evidence assessment, the Act imposes stringent regulatory controls to safeguard fundamental rights and uphold institutional integrity. These obligations include risk management requirements under Article 9, ensuring that potential harms are systematically identified and mitigated. Article 10 establishes standards for data quality and governance, mandating the use of relevant, representative, and error-free datasets. To enhance accountability, Articles 11 and 12 require comprehensive technical documentation and record-keeping throughout the AI system's lifecycle. Article 13 mandates transparency obligations, requiring that high-risk AI systems be understandable to users and subject to clear instructions. Article 14 introduces mandatory human oversight, ensuring that automated outputs do not operate unchecked and that meaningful human intervention is possible. Finally, Article 15 enforces standards for accuracy, robustness, and cybersecurity, obliging developers to prevent errors, manipulation, and system vulnerabilities. Together, these provisions aim to ensure that high-risk AI systems operate in a manner that is lawful, traceable, and aligned with the principles of accountability and fairness.

Further, high-risk AI systems must undergo conformity assessments before entering the market,[30] and remain under continuous post-market monitoring,[31] ensuring that their operation remains lawful and accountable even after deployment. A key innovation in the AI Act is its direct regulation of synthetic media, especially deepfakes. As outlined in Recital 133 of the EU AI Act, developers of AI systems that generate synthetic content should integrate technical solutions to mark and detect AI-generated output in order to show that the output has been generated or manipulated by an AI system and not a human. Recital 133 further emphasizes the need for reliable techniques, such as watermarks, metadata, or cryptographic methods, to prove the origin and authenticity of the content. This aims to mitigate risks of misinformation, manipulation, and deception by ensuring transparency and accountability in the information ecosystem.

The definition of deepfakes, as provided in the Act, encompasses 'AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful.'[32] This broad definition ensures that the legal framework addresses not only fake representations of individuals but also of objects, places, and entities such as businesses, governments, or institutions. This expanded scope is crucial, as it acknowledges that deepfakes can misrepresent not just people but organizations and inanimate things, which can potentially cause reputational, economic, or legal harm. However, experts worry this definition is too vague. Consider a scenario where AI sharpens a photo of a tree to make it more realistic, does that qualify as a deepfake? The uncertainty surrounding such cases raises concerns about the Act's ability to effectively identify deep fakes and regulate AI-generated content. Interestingly, Article 52(3) states that if a provider claims its AI model does not pose significant risk, the European Commission can reject this claim if it is inadequately substantiated, and designate the model as posing significant risk. This mechanism ensures that models capable of generating deep fakes or mass-manipulated content are scrutinized at the highest level. This is particularly relevant to the justice system: high risk models are now controlled strictly, this increases the likelihood that their outputs are traceable, labelled, and auditable. In practice, this means courts would be better equipped to verify whether such AI-generated content is reliable evidence. Thus, the EU model offers a foundation for enhancing judicial integrity by requiring that advanced AI systems be subjected to higher regulatory scrutiny, thereby ensuring that the truth and integrity of justice are not compromised by complex and opaque AI tools.

Institutional Framework

National Information Technology Development Agency: The National Information Technology Development Agency (NITDA) plays a central role in Nigeria's digital era and serves as a key institution in the governance of emerging technologies, including AI.[33] NITDA promotes research, development, capacity building, and innovation in AI and related technologies through its dedicated unit, the National Centre for Artificial Intelligence and Robotics (NCAIR). As part of NITDA's mandate to foster digital transformation, NCAIR spearheads the development of locally-driven AI solutions tailored to address Nigeria's unique socio-economic challenges. The Centre also contributes to policy development by providing technical support and coordinating stakeholder engagement in the drafting of the National Artificial Intelligence Policy (NAIP) and the National Artificial Intelligence Strategy (NAIS). In the context of AI-generated content and digital evidence, institutions like NCAIR are instrumental in setting technical standards, promoting ethical practices, and supporting initiatives that ensure the reliability, security, and traceability of AI systems.[34] Their involvement strengthens Nigeria's capacity to manage the risks associated with AI while enhancing the country's ability to enforce accountability and transparency in AI deployment, particularly in sensitive sectors like justice and public safety.

[30] EU AI Act, art. 16–23

[31] *Ibid* art. 61

[32] *Ibid* art. 3(60)

[33] National Information Technology Development Agency, 'Background' <https://nitda.gov.ng/background/> accessed 2nd July, 2025

[34] *Ibid*

Nigeria Data Protection Commission: The Nigeria Data Protection Commission (NDPC) is the principal regulatory body established under the Nigeria Data Protection Act, 2023 (NDPA) to ensure the protection of personal data and enforce compliance across both public and private sectors.[35] The Commission plays an important role in regulating how personal data are collected, processed, and stored, particularly as it relates to emerging technologies such as artificial intelligence. With AI systems often relying on large datasets that may include personal or sensitive information, the NDPC's role ensures that such data is handled lawfully and ethically. The NDPC plays a crucial role in enforcing the data protection principles set out in the Nigeria Data Protection Act. These include restrictions on fully automated decision-making and mandatory safeguards for securing personal data. In exercising its regulatory authority such as, ensuring compliance with the NDPA, licensing Data Protection Compliance Organisations (DPCOs), ensuring the appointment of Data Protection Officers in public and private entities, and investigating data breaches. These initiatives underscore the Commission's commitment to accountability, security, and transparency in personal data processing, which is essential for managing AI-driven systems that rely on personal or biometric data. Ultimately, through its regulatory authority, the NDPC strengthens governance mechanisms that are critical to ensuring the integrity and reliability of AI-generated evidence, and to uphold the integrity of justice in the digital era.

Federal Competition and Consumer Protection Commission: The Federal Competition and Consumer Protection Commission (FCCPC) is Nigeria's primary agency responsible for enforcing consumer rights, promoting fair market practices, and regulating competition under the Federal Competition and Consumer Protection Act. While the FCCPC does not directly regulate AI or digital evidence, its mandate intersects with the broader implications of AI use in consumer-facing applications and digital markets. As AI technologies are increasingly deployed in sectors such as finance, e-commerce, advertising, and telecommunications, there is a growing risk of algorithmic manipulation, automated misinformation, and deceptive content, including deepfakes, being used to mislead consumers or unfairly influence decision-making. The FCCPC's role in protecting consumers from false or misleading representations, unfair trade practices, and ensuring digital transparency is crucial in this context. For instance, where AI-generated content is used to impersonate individuals or businesses, spread misinformation, or manipulate online consumer behaviour, such practices may fall within the scope of prohibited conduct under the Federal Competition and Consumer Protection Act. The Commission's regulatory powers can prevent AI misuse that undermines trust in digital systems or manipulates judicial evidence in consumer disputes. By upholding ethical standards and accountability, the FCCPC complements efforts to ensure the integrity of data, protect consumers, and verify the authenticity of digital content in commercial and legal contexts.

Federal Ministry of Communications, Innovation and Digital Economy: The Federal Ministry of Communications, Innovation and Digital Economy (FMCIDE), established in 2011 (originally as the Federal Ministry of Communications Technology), plays a central role in driving Nigeria's digital transformation. Its core mandate is to harness the power of Information and Communication Technology (ICT) to foster inclusive access, improve economic development, and promote transparency in governance. The Ministry is strategically tasked with four primary responsibilities: enabling universal access to digital infrastructure; promoting the integration of ICT across sectors; fostering growth in Nigeria's ICT industry; and leveraging digital tools to improve transparency and efficiency in public service delivery. In the context of AI and the regulation of digital evidence, particularly AI-generated content such as deepfakes, the FMCIDE's role is highly relevant. The Ministry is not only a major unit for digital innovation but also a coordinator of national strategies for safe and ethical AI adoption. This was evident in its role in releasing the white paper in August 2023 that set the foundation for the National Artificial Intelligence Strategy (NAIS). The NAIS, released in draft form in August 2024, provides a framework for responsible AI development, by outlining procedures to identify, assess, and mitigate risks associated with AI, including the manipulation of data and digital content. Given the Ministry's mandate to drive transparency in governance and ensure cost-effective and reliable public services, its role directly intersects with efforts to safeguard the integrity of digital systems, particularly those that may be affected by AI misuse in critical sectors like law enforcement, the judiciary, and data regulation. For instance, in contexts where deepfakes technology could distort digital evidence or compromise legal processes, the Ministry's role ensures that Nigeria's digital system is guided by principles promoting ethical AI development, risk management, and accountability. Ultimately, the FMCIDE regulatory framework supports other regulatory institutions, such as the Nigeria Data Protection Commission (NDPC) and NITDA, in creating a digital environment that is secure, transparent, and prioritizes human rights, which are essential for addressing evolving challenges posed by AI-generated evidence in Nigeria's legal structure.

Nigerian Communications Commission: The **Nigerian Communications Commission (NCC)** is the primary regulatory authority for the telecommunications industry in Nigeria. NCC, established by the Nigerian Communications Act 2003, promotes fair competition and secure quality communications services. As AI technologies increasingly rely on telecommunications infrastructure to collect, transmit, and process vast amounts of data, particularly in areas like surveillance, biometric identification, and automated decision-making, the NCC's role becomes critically relevant in the governance of digital systems. In relation to AI and digital evidence, the NCC plays an indirect but pivotal role in safeguarding the infrastructure upon which AI systems operate. For example, secure networks and data transmission channels are essential to prevent tampering with digital information, including the type of data that may be presented as digital evidence in legal proceedings. A compromised communication channel can serve as a point of vulnerability where **deepfakes** or other forms of manipulated content are introduced or disseminated. The NCC's regulatory oversight helps to minimize such risks by enforcing cybersecurity standards and promoting data integrity, in collaboration with national security and law enforcement agencies. Moreover, the NCC collaborates with other institutions such as the Nigeria Data Protection Commission (NDPC) and National Information Technology Development Agency (NITDA) in shaping policies around data protection, digital rights, and AI

[35] Nigeria Data Protection Commission, 'About Us' <https://ndpc.gov.ng/> accessed 2nd July, 2025

governance**.** Through its Strategic Vision Plan (SVP 2021–2025)**,** the Commission has highlighted emerging technologies, including Artificial Intelligence**,** as focal areas for regulatory attention, particularly regarding ethical data use**,** network security**,** and consumer protection**.**[36] Despite the absence of AI-specific regulations, the NCC's responsibilities in digital infrastructure management and data security highlight its importance in shaping Nigeria's AI regulatory framework. Its contribution helps ensure that AI applications, especially those that could influence judicial outcomes or infringe on individual privacy, are deployed within a secure, monitored, and privacy-preserving digital ecosystem.

Nigerian Bar Association: The Nigerian Bar Association (NBA) plays a fundamental role in Nigeria's legal system, it promotes the rule of law, protects human rights, and safeguards the integrity of the legal profession and the judiciary. As the primary regulatory and advocacy body for legal practitioners, the NBA also contributes to law reform, access to justice, and professional development. In the context of rapid technological advancement, especially the integration of AI in legal processes, the NBA has begun to address the challenges that AI presents to legal ethics, procedural fairness, and evidentiary integrity. Recognizing the disruptive potential of AI tools, particularly in the context of deepfakes and manipulated digital evidence, the NBA released its Guidelines on the Use of Artificial Intelligence in the Legal Profession in 2024. These Guidelines provide practical and ethical guidance for lawyers navigating the use of AI in practice. Among other provisions, the Guidelines stress the importance of human supervision, data privacy, and transparency in applying AI tools within the legal practice. They also warn against over-reliance on automated systems, which may compromise professional judgment or due process. Crucially, the Guidelines recommend that lawyers conduct an AI Impact Assessment prior to adopting AI tools. This includes identifying and evaluating potential risks to client confidentiality, evidentiary integrity, and fairness. To support this, the Guidelines offer a sample checklist that helps legal practitioners examine the implications of AI use in their work. This is particularly relevant as courts are increasingly confronted with the challenge of authenticating digital evidence, including detecting deepfakes or other manipulated materials. By setting ethical standards and promoting responsible AI adoption, the NBA contributes meaningfully to upholding the truth and integrity of justice in the digital age. Its guidance ensures that legal practitioners are better equipped to contest unreliable AI-generated content, protect client data, and maintain public trust in the legal system. In doing so, the NBA strengthens Nigeria's institutional readiness to confront the risks posed by AI in legal and evidentiary contexts.

Organs of Government: Under the Nigerian constitutional structure, the Legislature, Executive, and Judiciary, as the three arms of government, play a vital role in shaping how the country confronts the challenges and opportunities presented by AI, particularly in relation to digital evidence and deepfakes. The Legislative arm, established under Section 4 of the Nigerian Constitution, is vested with the power to make laws for the peace, order, and good governance of Nigeria. This includes the authority to enact future-specific legislation on AI governance, digital evidence, and cyber integrity. In the face of increasingly sophisticated AI-generated content. The National Assembly has the constitutional mandate to update or introduce laws that ensure the admissibility and credibility of digital evidence aligns with current technological realities. The Executive arm, provided for under Section 5 of the Nigerian Constitution, is responsible for implementing laws and driving national policies. This role involves overseeing regulatory bodies like the FMCIDE. NDPC and NITDA, which collectively develop and enforce AI and digital policies. Through the execution of national AI policies and data protection legislation, the Executive is crucial in preventing the misuse of personal data in the creation of deepfakes, thereby helping to preserve the integrity of digital evidence. The Judiciary, under Section 6 of the Nigerian Constitution, is charged with interpreting the law and ensuring justice is done in all legal matters. As AI-generated content increasingly enters courtrooms, the Judiciary bears the critical burden of assessing the reliability and admissibility of such evidence. In particular, judges must contend with the technical complexities of identifying manipulated evidence and balancing the right to a fair trial with the need to exclude falsified digital content. The courts must also ensure that due process is upheld where automated decisions or profiling may have influenced a case, issues already contemplated under Section 37 of the NDPA, which prohibits significant decisions based solely on automated processing without human review. Together, these three organs of government form a constitutional tripod essential for navigating the legal, policy, and ethical implications of AI. Their collaboration is necessary to ensure that Nigeria not only leverages AI for national development but also safeguards the truth, fairness, and integrity of its justice system in the digital era.

4. An Analysis of Deepfakes, AI, and the Integrity of Justice

The Evidentiary Dilemma: Deepfakes and Digital Truth: The concept of truth in judicial proceedings is deeply tied to the integrity and authenticity of evidence presented before the court. In both civil and criminal proceedings, evidence must satisfy legal standards of admissibility, relevance, and must be pleaded, before it can influence a court's decision. However, the emergence of deepfakes poses a fundamental challenge to these principles. Deepfakes, such as hyper-realistic, AI-generated content that can falsify audio, video, or images, are especially problematic when indistinguishable from authentic digital evidence. Deepfakes are concerning because they take advantage of the long-standing reliance on visual and audio evidence as highly persuasive. Historically, the courts have accepted digital recordings, CCTV footage, and voice recordings as probative tools for establishing facts. But the growing sophistication of generative AI tools, such as those that replicate facial expressions, vocal tones, and even speech patterns, means that deepfakes content can convincingly imitate real people saying or doing things they never said or did. This creates a new category of manipulated evidence that is difficult to detect using the standard test for computer generated evidence, and it raises critical questions about chain of custody, authenticity verification, and ultimately, whether the burden of proof lies with the proponent to prove its authenticity or the opposing party to disprove its authenticity.

[36] Samson Akintaro, 'NDPC to launch regulatory AI sandboxes for data protection in Nigeria' (May, 2025) <https://nairametrics.com/2025/05/12/ndpc-to-launch-regulatory-ai-sandboxes-for-data-protection-in-nigeria/> accessed 2nd July, 2025

In jurisdictions like Nigeria, with limited digital forensics capabilities, the introduction of AI-generated false evidence in legal proceedings could mislead judges, which undermines fair trial rights, and weakens trust in the justice system. The situation is worsened by the reality that many legal professionals, including judges, lawyers, and prosecutors, may lack the technical competence to identify or interrogate deepfakes content effectively. The gap between rapid technological progress and the legal system's capacity to respond effectively creates a pressing issue in handling digital evidence, which is described as a digital evidentiary dilemma. Although the Evidence Act provides some procedural and substantive rules for the admission of electronic evidence, it does not address the specific threat posed by AI-manipulated content. Section 84(2) requires authentication of computer-generated evidence, but it assumes a level of traceability and human control that deepfakes often lack. Consequently, the assumption that digital evidence is reliable is now on shaky ground in the era of AI. Courts must now transition from passive receivers of digital evidence to vigilant guardians of digital integrity, while subjecting computer-generated evidence to stricter scrutiny and requiring expert forensic analysis as a prerequisite for admissibility. Legal practitioners must also take responsibility for scrutinizing the sources, metadata, and verifiability of digital content they present in court. To uphold the law's commitment to truth, justice, and fairness, it is essential to confront the unique challenges that deepfakes pose to the integrity of justice.

Deepfakes and Denial: How the Liar's Dividend Undermines Justice: The rise of deepfakes technology has not only introduced a new category of synthetic misinformation, but also empowered a more insidious threat to truth and accountability– the liar's dividend.[37] This term is used to describe how genuine evidence may be discredited merely by raising the possibility that it could be fake. This means that the mere existence of deepfakes provides the opportunity for individuals to deny the authenticity of legitimate information by falsely claiming it is AI-generated. This phenomenon enables wrongdoers to cast doubt on legitimate evidence, creating uncertainty that can benefit guilty parties and compromise the truth and integrity of justice. This tactic is particularly effective in legal and political contexts, where digital evidence like audio or video recordings can be crucial. In court, a party facing damaging evidence may claim that the authentic content is fake. The problem is that once doubt is raised, the burden shifts to proving not just relevance, but the authenticity of digital evidence, which becomes a far more complex and time-consuming task. One illustrative example, often cited in discussions of the liar's dividend, involves the 2016 Access Hollywood video in which then-candidate Donald Trump was recorded making lewd comments about women.[38] Although he publicly acknowledged and apologized for the video upon its release, reports later emerged that he suggested the following year that the clip was not authentic. In a world where deepfakes were more mainstream at the time, it is plausible that a candidate might immediately dismiss such a video as AI-generated and find support from political allies or social media influencers in doing so. This illustrates how the liar's dividend could be operationalized not only by the accused themselves, but also by supporters seeking to cloud the truth. High-profile figures may strategically choose not to outrightly deny a content, but instead use proxies or media outlets to raise suspicions about its authenticity, suggesting it might be AI-generated. What makes the liar's dividend particularly damaging to justice is that it erodes public and judicial confidence in all digital content, not just fakes. Legal systems, already burdened by evidentiary standards and technical limitations, may struggle to conclusively authenticate or debunk AI-generated media, especially without specialized forensic resources.[39] As a result, genuine evidence may be wrongfully discredited, and fabricated content may slip through unnoticed. In Nigeria, where the judiciary's forensic capabilities are still evolving, the liar's dividend poses a significant risk. It erodes trust in both the rule of law and digital justice systems, which increasingly rely on audio-visual evidence. Thus, there is an urgent need to address this issue, where the widespread of false AI content threatens to undermine the credibility of genuine evidence and compromise the integrity of the justice system.

AI and the Risk to Judicial Integrity and Fair Trial: As AI becomes ever-more integrated into legal systems, whether through tools used by lawyers, platforms for dispute resolution, or even AI-generated evidence, the risks it poses to the integrity of the judiciary and the right to a fair trial have become more pronounced. These risks are not speculative; they stem from real-world developments in how AI intersects with digital evidence, information processing, and legal proceedings. As mentioned earlier, one core concern is that AI-generated or manipulated evidence, could be introduced into courtrooms where neither judges nor legal practitioners are adequately trained to identify or challenge their authenticity. For instance, a litigant in a civil case could present an audio deep fake clip purporting to be a business agreement or defamatory statement, thereby influencing judicial perception, especially if the opposing party lacks the resources or technical support to contest it. Even more unsettling is the impact on judicial reasoning. Judges may rely on AI-driven legal research tools, predictive analytics, or even AI-generated summaries of case law. While these tools can enhance efficiency, they also raise the possibility of biased or erroneous outputs influencing judicial decisions, particularly in jurisdictions where AI regulation and legal tech literacy are still developing.[40] If these tools rely on flawed or skewed data, they could inadvertently reinforce existing biases or produce misleading interpretations of legal principles. Another critical issue is procedural fairness. The increasing reliance on AI tools in the justice system raises concerns about transparency and accountability, particularly when algorithms play a significant role in decision-

[37] Carpenter, Perry. Faik: A Practical Guide to Living in a World of Deepfakes, Disinformation, and AI-generated Deceptions. John Wiley & Sons, 2024. <https://books.google.com/books?hl=en&lr=&id=bnEbEQAAQBAJ&oi=fnd&pg=PT7&dq=The+rise+of+deepfakes+technology+has+not+only+introduced+a+new+category+of+synthetic+misinformation,+but+also+empowered+a+more+insidious+threat+to+truth+and+accountability%E2%80%93+the+liar%E2%80%99s+dividend.+&ots=epqj-PdBZM&sig=_DKPaXwfL-d3paoL6aiCdEsdxj0> accessed 2nd July 2025

[38] Josh A. Goldstein, Andrew Lohn 'Deepfakes, Elections, and Shrinking the Liar's Dividend' (January 2024)) <https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend> accessed 2nd July 2025.

[39] Apolo, Yvonne, and Katina Michael. 'Beyond a reasonable doubt? Audiovisual evidence, AI manipulation, deepfakes, and the law.' IEEE Transactions on Technology and Society 5.2 (2024): 156-168.

[40] Balakrishnan, Abhijith, *Ethical and Legal Implications of AI Judges: Balancing Efficiency and the Right to Fair Trial.* MS thesis. 2024.

making processes. In Nigeria, this threatens the constitutionally guaranteed right to a fair hearing under Section 36 of the 1999 Constitution. For instance, if AI-generated evidence is presented in court without clear explanations of its methodology, a party may struggle to adequately challenge its validity. They might be able to deny the truth of the evidence, but without understanding how it was generated, they cannot provide a sufficient explanation to discredit it. This opacity of AI tools and lack of traceability could undermine the adversarial process, leaving litigants unable to fully exercise their right to a fair hearing and challenge the case against them. Moreover, the public perception of justice could suffer in an age where people cannot easily distinguish between genuine and synthetic content. If a court's verdict relies on evidence later revealed to be AI-generated or tampered with, even unintentionally, it could lead to broader scepticism about the legitimacy of judicial decisions. In a country like Nigeria, where the justice system already struggles with issues of public confidence, this erosion of trust could have profound implications. Ultimately, while AI holds enormous potential to streamline judicial processes and expedite case disposition, it must be integrated with clear safeguards, technical expertise, and judicial oversight. The integrity of justice depends not just on the tools we use, but on our ability to ensure that those tools serve human fairness, not replace it.

Assessing Existing Safeguards: Are they Adequate?: As legal systems attempt to adapt to the rapid evolution of AI and synthetic media, questions naturally arise about the sufficiency of existing evidentiary safeguards. In Nigeria, one of the most relevant provisions in this regard is Section 84 of the Evidence Act, which was recently amended to reflect the growing importance of digital records in judicial proceedings. These amendments now recognize 'electronic records' alongside traditional documents, allowing electronically stored, sent, or received information to be admitted into evidence if certain criteria are met. These include demonstrating the reliability of the system used to create or store the record, and attesting to the regularity with which the system was used. Section 84A, newly introduced, further expands on this by validating electronic information even where a written or printed form is otherwise required, provided that such information is accessible and usable for future reference. The law also gives legal recognition to digital signatures and permits electronic deposition of affidavits, including the use of electronic oaths. These developments reflect an effort to align Nigeria's evidentiary framework with contemporary digital realities, where legal documents and communications are usually created and transmitted in electronic form. However, when assessed in light of threats posed by deepfakes and generative AI, the adequacy of these provisions becomes less certain. While the amendments to Section 84 create a pathway for the admissibility of electronic records, they remain primarily focused on procedural and technical criteria, such as the source and consistency of data, rather than addressing the integrity or authenticity of the content itself. In the context of AI-generated content, this distinction is critical. Deepfakes, by design, can produce highly realistic but entirely fabricated audio-visual material. Such content may satisfy procedural admissibility criteria under the Evidence Act, particularly if the party presenting the evidence can show it was stored in a reliable system. Yet the content could still be fundamentally false.[41] For example, if a litigant tenders a video purporting to show a defendant making incriminating statements, and the video was stored on a regularly used digital platform, there is currently no express requirement in Nigerian law for independent verification of whether the content itself is authentic and unaltered. This creates a concerning gap: AI-generated fabrications may be procedurally admissible even when substantively misleading. The law does not yet require courts to examine the content for signs of manipulation, nor does it mandate forensic analysis or expert input as a condition for admissibility in such cases.

Additionally, the current provisions are silent on the issue of human oversight in determining the veracity of digital evidence. This is significant given how easily AI-generated content can be presented in court as credible evidence. Without the requirement of meaningful review by experts or the application of content authentication tools, the system relies heavily on the opposing party's capacity to identify and challenge manipulated evidence. In practice, this can place an unfair burden on less-resourced litigants who may lack the technical expertise or financial means to contest convincingly falsified content. In this light, while the amendments to Section 84 and the introduction of Section 84A represent commendable efforts to modernize Nigeria's evidentiary regime, they appear to fall short in addressing the specific and growing risks posed by synthetic media. The law's focus on system reliability rather than content integrity may inadvertently allow AI-generated content with questionable accuracy to influence the outcome of cases. In the legal system where truth and credibility are paramount, this gap has profound implications, not just for individual cases but for public confidence in the legal process.

Balancing Innovation with Legal Safeguards: The Role of Human Oversight and Technical Expertise: The impact of AI in the legal system is promising– faster legal research, predictive analytics, enhanced access to evidence, and streamlined administrative processes. Yet, this promise exists alongside significant risks, especially where AI intersects with integrity of processes in court. Balancing innovation with legal safeguards requires more than simply updating statutes or allowing digital tools into the courtroom. It calls for deliberate, informed human involvement and technical fluency to ensure that these tools are used responsibly, particularly when they affect rights, liberties, or judicial credibility. AI does not operate in a vacuum; it is created, trained, and deployed by humans who make choices about data, models, and objectives. These choices inevitably reflect biases, gaps, or blind spots that can surface when AI tools are used in legal contexts. For instance, an AI-powered system used to analyse financial transactions for suspicious activity might be trained on data that doesn't account for the unique characteristics of informal financial systems in Nigeria, such as the prevalence of mobile money services or traditional savings groups. As a result, the system might flag legitimate transactions as suspicious, causing unnecessary delays or even false accusations. This highlights the need for human oversight and contextual understanding in AI-driven decision-making processes.

[41] Delfino, Rebecca A. 'Deepfakes on trial: a call to expand the trial judge's gatekeeping role to protect legal proceedings from technological fakery.' Hastings LJ 74 (2022): 293.

Moreover, the ability to question and challenge AI-generated content, whether a piece of digital evidence, a forensic reconstruction, or a data-driven legal prediction, depends on the presence of individuals with the necessary technical expertise.[42] Judges, lawyers, and investigators must be equipped not only with legal knowledge but also with a working understanding of how AI systems function, what their limitations are, and how errors or manipulations might occur. Without this foundational understanding, the courtroom risks becoming a stage where technological sophistication overshadows legal reasoning. This is particularly relevant in the context of deepfakes, where even trained, eyes can struggle to discern manipulation without the aid of forensic tools. If a court admits such content as evidence, the question becomes not only whether the content meets procedural thresholds as discussed earlier, but also whether anyone involved in the trial has the competence to interrogate its authenticity. The risk is in blindly trusting digital content as objective or trustworthy, when its very persuasiveness is what makes it potentially damaging. At the same time, outright rejection of AI technologies in the administration of justice would be equally short-sighted. When used transparently and responsibly, AI can support decision-making, reveal patterns of injustice, and enhance efficiency.[43] The challenge lies in embedding layers of accountability into the system: ensuring that AI tools complement human judgment, not replace it; that their outputs are explainable, not mysterious; and that users can interrogate their logic, not merely accept their conclusions. This is the essence of legal safeguards, not only rules on admissibility but also a wider framework of human responsibility and institutional vigilance.

Ultimately, innovation in legal practice must be approached with a mindset of cautious optimism. Thus, there should be a careful balance of enthusiasm for innovation with a critical and careful approach. Technical tools can enhance access to justice and promote fairness, but only when there is human involvement and technical literacy to keep pace with the speed of technological advancement.[44] In matters as serious as evidence, criminal liability, or judicial decision-making, one cannot afford to be dazzled by innovation at the expense of foundational legal values. Safeguarding justice in the AI age demands that legal professionals strike a balance between harnessing innovation and preserving core values.

5. Case Studies and Comparative Perspectives

Foreign Case Law on AI and Digital Evidence: The cases of *People v. Wakefield*[45] and *State v. Morrill*,[46] both decided in the United States, offer instructive insights into how courts in other jurisdictions have begun grappling with the admissibility and implications of AI-generated evidence. In *People v. Wakefield,* decided by a New York appellate court, the issue was based on the admissibility of DNA evidence processed by TrueAllele, a software system developed by Cybergenetics. The system integrates AI-driven analysis with probabilistic genotyping to interpret genetic data from complex DNA mixtures. At a hearing, the expert testimony from the software's developer, established that TrueAllele's methodology was generally accepted within the relevant scientific community. The court ruled that the system's analysis, which generates a likelihood ratio indicating how much more probable it is for a suspect to match a DNA profile than a random individual, met the required standard for admissibility as evidence. This case demonstrates that courts can subject AI-based forensic tools to rigorous scientific scrutiny before allowing them into legal proceedings. It also underscores the importance of expert involvement and transparency when introducing algorithmically generated results, ensuring that such evidence supports rather than undermines the investigation process. Further, *State v. Morrill*, decided by the New Mexico Court of Appeals, examined whether the outputs of two software programs, Roundup and Forensic Toolkit, should be classified as hearsay. The defendant argued that computer-generated conclusions should be treated as statements and thus inadmissible unless they fell within a recognised hearsay exception. However, the court disagreed, holding that these automated outputs did not qualify as 'statements' by a person under the rules of evidence. Instead, the court reasoned that the software merely logged and organised data without human intervention, making its outputs akin to neutral, mechanical recordings rather than subjective assertions. This ruling draws a critical distinction between software that passively processes or organizes data and more advanced AI systems that may generate conclusions or inferences. It highlights a growing challenge in evidentiary law: as AI systems become increasingly autonomous and generative, courts face the question of whether their outputs are objective data or testimonial evidence requiring validation and cross-examination.

Moreover, on the intersection of AI and justice, the case of *State v. Saylor*,[47] decided by an Ohio appellate court, provides a nuanced perspective on how algorithmic tools are affecting judicial discretion, particularly in sentencing. In this case, the court considered the Offender Risk Assessment System (ORAS), a tool designed to evaluate an offender's likelihood of recidivism. The concurring opinion is particularly relevant, as it critiques the increasing reliance on algorithmic assessments in criminal justice decision-making. While it acknowledged that such tools may represent a good-faith attempt to reduce arbitrary discretion and systemic bias, it also expressed significant concern about the opacity and perceived objectivity of these systems. Specifically, the opinion noted that although the defendant, Saylor, had a minimal criminal history, he received a 'moderate' risk score, which the trial court viewed as unexpectedly high. This discrepancy raised doubts about how ORAS scores are computed and interpreted, especially in the absence of transparency about the underlying logic. It emphasized that even

[42] Ahmed, Saquib, et al. 'Enhancing Crime Scene Analysis: The Impact of AI Technologies on Evidence Processing.' Forensic Intelligence and Deep Learning Solutions in Crime Investigation. IGI Global Scientific Publishing, 2025. 63-84.

[43] Leslie, David, and Antonella Maia Perini. 'Future Shock: Generative AI and the international AI policy and governance crisis.' Harvard Data Science Review Special Issue 5 (2024).

[44] Donoghue, Jane. 'The rise of digital justice: Courtroom technology, public participation and access to justice.' 80.6 (2017): 995-1025.

[45] 175 A.D.3d 158, 107 N.Y.S.3d 487 (3d Dept. 2019) *The Modern Law Review*

[46] No. A-1-CA-36490, 2019 WL 3765586 (N.M. App. July 24, 2019)

[47] *State v. Saylor,* 2019-Ohio-1025.

algorithmic systems marketed as objective are ultimately shaped by human decisions—about what variables to include, how to prioritize them, and what outcomes to emphasize. This case also reveals the difficulty of reconciling machine-generated assessments with human judgment. The trial court's scepticism toward the ORAS score suggests a natural friction: while AI can offer helpful insights, it should not override a judge's informed intuition, especially when rights and liberty are at stake. The concern deepens when courts either over-rely on or casually disregard algorithmic outputs without justification, especially when they do not fully understand how those outputs are generated. Ultimately, the case of *State v Saylor* serves as a reminder that while AI holds potential to support judicial efficiency and consistency, it cannot substitute for human discernment and procedural fairness, which remain the bedrock of justice.

In addition to questions surrounding the admissibility and weight of AI-generated evidence, courts in several jurisdictions have also begun addressing an equally pressing issue: the misuse of generative AI in legal filings, particularly the submission of fabricated case law produced by tools like ChatGPT. This trend was most infamously highlighted in *Mata v. Avianca, Inc*.,[48] where two attorneys submitted a motion citing multiple non-existent judicial authorities generated by an AI tool. The Judge imposed sanctions and emphasized that lawyers have a continuing duty to verify the accuracy of their legal submissions, regardless of the technological tools they employ. The court made it clear that while AI may assist in research, it cannot substitute for professional due diligence. This pattern of concern continued in *Hamad Al-Haroun v. Qatar National Bank QPSC & QNB Capital LLC* (unreported), where a UK judge uncovered that 18 of the 45 cited authorities were fictitious, generated by AI and unverified by counsel. The solicitor involved had relied on AI for legal research without independently confirming the cited precedents, this led the court to highlight the risk of blindly depending on generative technologies in serious litigation.
A similar outcome occurred in *R v. London Borough of Haringey*,[49] where the claimant's counsel cited five entirely fabricated cases. Although the counsel denied using AI, the court observed that the only plausible explanations were either deliberate fabrication or undisclosed AI usage, both potentially amounting to contempt of court. The judge referred the matter to the Bar Standards Board, reinforcing that misrepresentation, even if assisted by AI, remains a breach of legal ethics and judicial integrity. Outside the common law jurisdictions, the Grand Court of the Cayman Islands has also weighed in. In *Bradley & Chuang v. Frye-Chaikin*,[50] the court criticised submissions that had clearly been generated with AI tools and contained 'hallucinations,' including fake procedural rules and non-existent case law. It also used the opportunity to issue broader guidance: while the use of AI is not inherently improper, it carries with it a duty of verification and accuracy. The court warned that litigants and counsel alike may face personal consequences, including wasted cost orders, if they fail to ensure that AI-generated material is truthful and legally sound.

These recent cases reveal a new risk at the intersection of law and AI: the spread of AI-generated legal content that appears authentic but is actually fabricated. Courts are drawing a clear line: AI is a tool, not a shield. Its use must be transparent, supervised, and subject to professional standards. These cases demonstrate that while AI can enhance efficiency, it can also introduce profound vulnerabilities if used without rigor. For jurisdictions like Nigeria, these decisions underscore the urgency of developing ethical and procedural safeguards, not only around AI-generated evidence but also around the reliability of legal submissions crafted in the age of machine intelligence.

Comparative Legal Frameworks: India and China: As AI continues to transform legal systems worldwide, comparative insights from countries that are proactively addressing the intersection of AI and evidentiary law are increasingly instructive. While Nigeria has yet to develop clear legal frameworks or judicial precedents specifically addressing AI-generated evidence, examining the approaches adopted by jurisdictions like India and China offers valuable guidance. These countries showcase two distinct approaches: India, with its gradual integration of digital evidence into its traditional system, and China, with a more proactive regulatory stance on AI and synthetic media in courts.

India: A Traditional System Struggling to Modernize: India's legal framework for evidence remains rooted in the Indian Evidence Act of 1872, a colonial-era statute that was never designed to grapple with the complexities of modern technology, let alone AI. While the statute has undergone amendments to accommodate electronic evidence, its foundation lacks the specifics required to address the unique challenges posed by AI-generated content. One of the primary provisions governing digital evidence in India is Section 65B of the Evidence Act. This section permits the admissibility of electronic records, provided they are accompanied by a certificate affirming their authenticity and origin. Though effective for conventional digital formats such as emails, SMS messages, or electronic documents, Section 65B falls short when applied to AI-generated which often result from opaque processes rather than straightforward data storage or transmission. Consequently, the mere certification of a document's origin under Section 65B does not address the more pressing concern of how an AI system arrived at its conclusion, especially when the algorithm's logic is not transparent or interpretable. In an effort to modernize its legal framework, India introduced the Bharatiya Sakshya Adhiniyam, 2023,[51] which aims to replace the existing Evidence Act. While this new legislation does more to acknowledge digital evidence, it still lacks targeted provisions dealing with machine-generated content or deductive outputs from AI systems. This legislative ambiguity leaves a significant gap in addressing the admissibility and evaluation of AI-generated material, particularly where transparency and accountability are essential. Courts have focused on electronic evidence broadly, but haven't specifically addressed AI-generated content yet. Landmark decisions such as *Anvar*

[48] 678 F. Supp. 3d 443 (S.D.N.Y. 2023)
[49] [2025] EWHC 1167 (Admin)
[50] [2025] CIGC (Civ) 5
[51] Act No. 47 of 2023

P.V. v. P.K. Basheer[52] and *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*[53] have reinforced the requirement of strict compliance with Section 65B. However, these cases relate to the authenticity of digital evidence, not the reliability or transparency of outputs created by autonomous AI systems. In such contexts, the courts have not yet engaged with questions about algorithmic bias, the need for accountability, or the role of human oversight in interpreting machine-generated evidence. Thus, while India has laid some foundational legal structures for dealing with digital evidence, it remains largely unprepared to confront the complex challenges of AI in legal proceedings.

China: A Model of Proactive Regulation and Protections: In contrast, China has adopted a far more proactive and multifaceted regulatory approach to the challenges posed by AI and synthetic media. Rather than relying solely on judicial precedent or general electronic evidence provisions, China has established specific rules and policy guidelines aimed at both the regulation of AI technologies and their controlled integration into judicial systems. One of the cornerstone regulatory developments in China is the introduction of the Deep Synthesis Provisions, which came into effect in January 2023. These rules mandate that any platform or provider utilizing generative AI to produce synthetic media, including audio, video, text, or virtual scenes, must clearly label such content. Requirements include visible and embedded markers of synthetic origin, identity authentication of users, monitoring obligations, and regular algorithmic audits. This framework is designed to enhance transparency and accountability, ensuring that courts and the public can clearly distinguish between authentic and AI-manipulated content. In addition to this, the Draft Measures for the Administration of Generative AI Content, introduced in September 2024, propose stricter obligations for labelling AI-generated media. These measures mandate watermarking and metadata tagging, particularly for content with the potential to mislead. This regulatory emphasis on transparency of content stands in sharp contrast to jurisdictions like India, where legal mechanisms primarily focus on data authenticity rather than the synthetic nature or intent behind content creation.

Moreover, China’s Interim Measures for Generative AI Services (August 2023) extend the regulatory reach to AI service providers, compelling them to ensure compliance with national security standards, data protection requirements, and accuracy obligations. These measures also require platforms to monitor content and remove illegal or harmful material, creating a more controlled and accountable AI ecosystem. Perhaps most significantly, China has issued judicial policy directives to govern the use of AI within the court system itself. In December 2022, the Supreme People’s Court released formal guidelines stating that AI may be used to assist judicial functions but must not replace human judgment. The guidelines emphasize that any AI-assisted outcome must be traceable, explainable, and subject to human oversight. While ‘smart courts’ in cities such as Hangzhou and Beijing have implemented AI tools for tasks like e-filing, document review, and draft opinions, the final authority remains firmly with human judges. This deliberate stance highlights China’s commitment to preserving judicial integrity and accountability, even as it embraces technological innovation.

Together, the experiences of India and China reveal two distinct but complementary approaches to regulating AI-generated evidence. India's experience highlights the challenges of adapting outdated legal frameworks to rapidly evolving technologies. Its reliance on general electronic evidence provisions without specific guidance on AI-generated content highlights the risk of under-regulation. China, on the other hand, provides an example of forward-looking regulatory innovation with detailed rules on synthetic media, transparency obligations, and judicial oversight protocols. For Nigeria, where no dedicated legal framework yet exists to address the admissibility of AI-generated evidence, both models present valuable lessons. India's experience serves as a warning against complacency and overdependence on outdated frameworks, whereas China's approach demonstrates the advantages and of implementing a structured regulatory framework that prioritizes the integrity of content, transparency, and human accountability. As Nigeria explores legal frameworks for AI in its justice system, drawing from other jurisdictions can provide an informed balanced approach that integrates technology while protecting fundamental rights.

6. Conclusion and Recommendations

The integration of AI into legal systems presents both significant opportunities and profound challenges, particularly in the realm of evidence law. As demonstrated throughout this analysis, AI-generated evidence introduces complex questions about admissibility, reliability, transparency, and accountability which are issues that traditional evidentiary frameworks were not designed to address. From the evolving jurisprudence in jurisdictions like the United States, to proactive regulatory models in China and transitional reforms in India, one central theme emerges: legal systems must adapt rapidly if they are to maintain procedural fairness and public confidence in the age of intelligent machines. While Nigerian law currently lacks specific provisions addressing the admissibility or regulation of AI-generated evidence, it is clear from comparative perspectives that a reactive or minimalist approach will be insufficient. AI-generated outputs, whether forensic conclusions, risk assessments, or even fabricated legal authorities, can significantly influence judicial outcomes. Therefore, the legal system must develop principled guidelines that ensure such technologies serve justice rather than undermine it. Foreign case law reveals how courts are beginning to confront questions about hearsay, algorithmic transparency, and the limits of judicial reliance on automated tools. At the same time, the rise in AI-assisted legal malpractice, such as the submission of fabricated citations, illustrates the dangers of uncritical dependence on generative technologies. Without rigorous verification mechanisms and ethical boundaries, AI can be weaponized to distort legal processes. Comparative insights from China and India further demonstrate that while the challenges are global, responses can vary dramatically. China’s emphasis on regulatory clarity, traceability, and human oversight offers a forward-thinking model, while India’s struggles highlight the risks of relying on outdated or overly general provisions. Together, these cases underscore that any credible approach to AI in law must prioritize transparency,

[52] (2014 10 SCC 473)

[53] Supreme court 4908, Aironline 2020 SC 641

accountability, and due process. Nigeria stands at a critical juncture. The rapid rise of AI in legal practice is inevitable, but whether it strengthens or destabilizes the justice system will depend on the frameworks adopted today. Courts, lawmakers, and legal professionals must collaborate to build an ecosystem where technology complements, rather than compromises, the rule of law.

As Nigeria confronts the growing use of AI technologies and digital evidence in its legal system, it is imperative to adopt a strategic framework that blends legal reform, technical safeguards, professional accountability, and global best practices. The following recommendations are offered to ensure that AI-generated evidence can be safely, reliably, and ethically incorporated into the justice system:

Enact AI-Specific Evidence Legislation: Nigeria should develop dedicated legal framework governing AI-generated evidence. This legislation should:

- Define what constitutes AI-generated versus human-generated digital evidence.
- Set out admissibility thresholds based on transparency, reliability, and expert validation.
- Provide standards for assessing contested AI media and algorithmic outputs.
- Distinguish between raw data and analysed or generated data. This legal clarity is essential for preventing ambiguity in admissibility rulings and protecting procedural fairness.

Amend the Evidence Act with Targeted Provisions: The Nigerian Evidence Act should be amended to:

- Introduce specific provisions addressing AI-derived evidence, including probabilistic tools, neural networks, and synthetic media.
- Require expert testimony in cases involving AI systems that perform non-transparent reasoning or inference.
- Codify judicial discretion to reject AI-generated content that fails transparency, or authenticity tests.

This will ensure the Act reflects modern evidentiary realities without relying on outdated assumptions about data integrity.

Strengthen Technical Infrastructure for Evidence Integrity: Nigeria must invest in secure digital infrastructure for the generation, transmission, and storage of digital evidence. This includes:

- Encrypted data channels (for instance, TLS- Transport Layer Security or HTTPS- Hypertext Transfer Protocol Secure) to prevent tampering during transmission.
- Digital signatures and hash values to verify the integrity of files and identify tampering.
- Access controls and audit trails for digital repositories to ensure chain of custody.

Establish Guidelines for Evaluating Contested or Synthetic Media: The courts must be equipped to assess whether digital content is genuine or AI-fabricated. This requires:

- Judicial guidance on evaluating deepfakes and synthetic media.
- Standards for admitting or rejecting contested videos, audios, or documents.
- Institutional capacity for forensic verification (e.g., media authenticity labs)

Given the liar's dividend, the risk that even authentic evidence may be discredited due to the existence of fakes, Nigerian courts need rules that go beyond general hearsay or relevance doctrines.

Mandate AI Impact Assessments for Legal Tech Use: Before AI tools are deployed in law firms, courts, or investigative agencies, an AI Impact Assessment (AIIA) should be conducted.[54] Key questions should include:

- Workflow Impact: How will this tool alter legal tasks or decision-making?
- Privacy and Confidentiality: Does it comply with Nigeria's Data Protection Act (NDPA)? How is client confidentiality maintained?
- Bias and Fairness: What safeguards are in place against algorithmic bias? How is fairness ensured in sentencing, evidence review, or risk scoring?

Such assessments ensure responsible integration and prevent blind dependence on unverified tools.

Ensure Compliance with NDPA and International Standards: AI tools used in the legal sector must align with Nigeria's Data Protection Act (NDPA), including principles such as: • Fairness and transparency.

- Purpose limitation and data minimization.
- Accuracy and accountability.
- Integrity, confidentiality, and lawful processing. International models like the EU's GDPR can offer additional guidance. Legal practitioners must adopt principles like:
- Privacy by design and by default.
- Algorithmic transparency—especially crucial in AI-generated evidence.
- Organizational accountability for AI use in sensitive matters.

[54] Nigerian Bar Association, 'Guidelines for the use of artificial intelligence in the legal profession in Nigeria' (2024) https://nbaslp.org/wp-content/uploads/2024/04/Guidelines-for-the-Use-of-Artificial-Intelligence-in-the-Nigerian-Legal-Profession.pdf accessed 2nd July, 2025

Develop Courtroom Guidelines and Judicial Training: Judges and magistrates must be equipped to understand and evaluate AI-generated content. To this end:

- The Nigerian Bar Association should provide tailored training on AI systems, digital evidence, and forensic evaluation.
- Standard judicial questions should include: Can this AI output be explained? Is it contestable? Has it been validated by an expert?

This capacity-building is foundational to fair adjudication.

Create Ethical Rules for Legal Professionals Using AI: The Rules of Professional Conduct should be updated to include duties related to AI use:

- Prohibiting submission of hallucinated or fabricated caselaw.
- Requiring verification of any AI-assisted research or filings.
- Penalizing misuse of generative AI that misleads the court.

Cases like *Mata v. Avianca* illustrate how generative AI can be misused by lawyers. Preventing this in Nigeria demands not just rules, but enforcement mechanisms.

Establish a National Oversight Body for Legal AI Tools: Nigeria should create a regulatory entity or unit, possibly under NITDA to:

- Certify and audit AI tools intended for use in legal and judicial settings.
- Monitor compliance with privacy, fairness, and security standards.
- Investigate abuse of AI in court processes and sanction responsible parties.

This institutional anchor will ensure long-term governance of legal AI tools.

Preserve Human Oversight and Judicial Discretion: No AI system should supplant human reasoning in justice delivery. Legal decision-making, legal research and legal reasoning must remain human-led. Courts should:

- Be empowered to override AI outputs where fairness or context demands.
- Demand clarity before acting on any automated recommendation.
- Reinforce that AI is an aid, not a replacement, for judicial wisdom.

By integrating these reforms, Nigeria can ensure that its legal system is not only technologically current but also ethically sound and well-equipped to handle the challenges of rapid digital transformation.