

Practice Update

The Challenges of Integrating AI-Generated Evidence Into the Legal System

June 12, 2024

By Evelina Gentry

The integration of artificial intelligence (AI) into the legal system heralds a transformative era marked by both innovation and unprecedented challenges. After some lawyers made headlines for submitting legal briefs with fictitious case citations that were generated by AI, several judges have updated their standing orders to address the use of AI-generated content in court filings. However, as AI-generated evidence becomes more prevalent, and increasingly indistinguishable from its non-AI counterparts, the courts will also need to address complex issues concerning the authenticity, reliability, and admissibility of trial evidence generated by AI. The courts are especially concerned about the risks of AI being used to manipulate videos and images and create “deepfakes,” *i.e.*, artificial images, video clips, and audio recordings created by AI that are fake but appear to be the real, that could taint a trial. For example, in the realm of intellectual property, deepfakes have the potential to unlawfully exploit an individual’s image and likeness, as well as trademarks and labels. If deepfakes use copyrighted material, they could be subject to copyright infringement claims; they can also raise issues with authorship or inventorship. These issues will naturally lead to introduction of AI-generated evidence at trial. This article delves into the multifaceted challenges presented by AI-generated evidence, explores the principles governing its

Related People

Evelina Gentry

Related Work

Litigation

Related Offices

Los Angeles

admissibility, and discusses the broader implications for the legal system.

The Challenges of AI-Generated Evidence

AI-generated evidence encompasses a wide range of materials, from documents, photos, and video and communications synthesized by AI to data analysis and patterns identified through machine learning algorithms. One of the primary concerns is the *accuracy and reliability* of AI-generated evidence. Unlike traditional evidence, which can often be directly traced and verified through human sources, AI-generated content might not have a clear lineage or might be the result of complex processes that are difficult to audit. This raises questions about how to prove that such evidence has not been tampered with or generated based on biased data.

Moreover, the *authentication* and chain of custody of AI-generated evidence present significant hurdles. While the legal system demands rigorous standards for evidence admissibility, the opaque nature of AI processes complicates these requirements, challenging parties to prove the integrity of evidence that a machine, not a human, has generated. To alleviate these issues and ensure the integrity and admissibility of AI-generated evidence, courts may require parties to provide evidence demonstrating how the data was collected, processed, and analyzed, causing potential delays in the discovery process and any subsequent trial. The additional resources required of the parties and the court may ultimately impact the parties' success in pursuing their claims.

Further, the complex nature of many AI systems may require *expert testimony* to explain how the evidence was generated. Yet, this necessitates relying on highly specialized knowledge, which may not always be accessible to the litigants or comprehensible to judges and juries. However, the expense of retaining highly specialized experts will further increase the cost of litigation and make it even more cost-prohibitive. Moreover, it will require

the court to assess the qualifications of experts and the reliability of their opinions, which will further burden the already strained judicial system.

Additionally, disputes may arise over the *integrity* of the data inputted into AI systems, the algorithms' validity, and the potential for bias or error in their outputs. Indeed, AI systems are only as unbiased as the data on which they are trained. Historical data can contain biases, and these biases can be perpetuated or even amplified by AI. In the context of evidence, this raises concerns about fairness and the potential for AI to generate evidence that is prejudicial against certain groups.

Privacy and data protection concerns also emerge, as AI systems frequently rely on extensive datasets, including sensitive personal information. Parties may raise concerns about privacy and data protection laws when introducing AI-generated evidence, particularly if there are questions about the legality of how the data was obtained or processed.

Also, a *lack of transparency* and explainability in AI algorithms can undermine the admissibility of evidence in the legal context, where the reasoning behind evidence is as important as the evidence itself. Thus, courts may require evidence to be presented and explained, and the inability to explain how AI-generated evidence was derived could lead to its exclusion or to challenges in its credibility.

The overarching issue is that AI technology is rapidly evolving, making it difficult for legal systems to keep pace with it and for practitioners to remain knowledgeable about the best ways to handle, interpret, and challenge AI-generated evidence. These technological advancements present significant challenges to parties, their counsel, and the courts in determining whether evidence is authentic or fake. They also raise concerns about whether litigation costs will dramatically increase as parties are forced to hire forensic experts to address

AI-generated evidence, the ability of juries to discern authentic from fake evidence, and whether the courts will become overwhelmed with AI-generated evidence and lawsuits.

Evidentiary Principles for Addressing the Admissibility of AI-Generated Evidence

The integration of AI into the legal realm, particularly in the context of evidence admissibility, presents both novel opportunities and challenges. To date, few, if any, court decisions squarely address the admissibility of AI evidence, and the cases that have referenced AI evidence have often done so in a cursory manner. In general, the *admissibility* of evidence hinges on several critical factors, including relevance, reliability, authenticity, and fairness. However, an important factor in evaluating the admissibility of AI evidence is whether the functioning of the AI system that produced the evidence can be explained to the trier of fact so they can understand, at least at a general level, how the system operates and how it achieves its results, and thus evaluate the amount of weight they are willing to give to the evidence derived from it.

The Federal Rules of Evidence (FRE), specifically Rules 401, 402, and 403, entrust trial judges with the gatekeeping role of determining the admissibility of evidence. This responsibility extends to AI-generated evidence. To deem AI evidence relevant, the offering party must explain how the AI system operates (*i.e.*, how it produced its outcome) and how the evidence will aid, rather than confuse, the jury towards a just verdict. This involves disclosing sufficient information about the AI system's training data, development, and operational mechanisms to enable both the opposition and the judge to evaluate it. Several considerations affect the admissibility and relevance of AI evidence. The accuracy and reliability of AI systems are fundamental; however, the interpretability of complex AI algorithms and the potential for privacy infringements due to extensive data usage also weigh heavily on their admissibility.

Moreover, the timeliness of AI-generated evidence and the inherent biases within AI systems can significantly impact their relevance and, consequently, their admissibility in court.

Authentication, as set forth in FRE 901(a), involves demonstrating that the evidence is what it purports to be, and it is crucial for AI evidence to be considered by a jury. The rules most applicable to AI evidence authentication are 901(b)(1) and 901(b)(9), which concern witness testimony and evidence describing a process or system that produces accurate results, respectively. Challenges in authenticating AI evidence include the opacity of AI algorithms, potential biases in training data, the quality of data used, compliance with regulatory standards, and the general lack of legal expertise in AI technology. These factors can complicate the authentication process, raising questions about the evidence's reliability and accuracy.

In conclusion, the admissibility of AI-generated evidence is a complex issue that intersects technology and law. A comprehensive approach is essential to effectively address the challenges posed by the use of AI-generated evidence within the legal system. This includes the development of clear standards and regulations to ensure the reliability, fairness, and transparency of AI-generated evidence. Additionally, educating legal professionals on the capabilities and limitations of AI will enable them to use AI-generated evidence more effectively. Finally, promoting the development of AI in an ethical manner, with a focus on fairness, privacy, and accountability, is vital for mitigating potential issues. As AI continues to evolve and integrate into various aspects of society, addressing these challenges will be crucial for ensuring that the use of AI-generated evidence supports justice and fairness in the legal system.

This information is intended to inform firm clients and friends about legal developments, including

recent decisions of various courts and administrative bodies. Nothing in this Practice Update should be construed as legal advice or a legal opinion, and readers should not act upon the information contained in this Practice Update without seeking the advice of legal counsel. Prior results do not guarantee a similar outcome.