

## AI's Scope: Synthetics in Courtrooms

Sadhvi Mehandru

In the current age of vigorous artificial intelligence (AI) employment on a global scale, familiar inquiries may arise—questioning the integrity of services seemingly facilitated by humans. During recent years, one of AI's distinct applications has been exhibited in the legal field and the national judicial system. Latter-day issues concerning fabricated evidence and synthesized imagery/video have taken root in American courts, taking away from the crucial role of rectitude embedded in the legal system. Despite AI as a research tool being outwardly permissible, many federal judges have professed their disapproval of AI in the context of translating into “hallucinated” cases or citations. Judges have also expressed concerns regarding the lack of transparency within AI algorithms, and subsequent worries about potential biases (or perpetuation of such prejudices) against litigants that may arise as a result. The issue of AI's newfound prevalence in court emphasizes the necessity for increased accountability and fundamental American principles that should be carried forward into court.

Considering the inflow of AI-generated or fabricated evidence in court has increased only in recent years, a significant number of deepfake exhibits haven't been announced. However, relative to the amount of evidence that's recently entered court, the presence of AI “research” via attorneys and in case studies has proven its potential for a tangible negative impact on courts. One notable instance of this newfound potential was demonstrated in 2021, in *State of Washington v. Puloka*. The case overview included the state of Washington charging defendant Joshua Puloka with three counts of murder, originally stemming from a shooting. As a result of the judge's concerns about credibility, the court excluded an AI-enhanced piece of video evidence from the case. This video evidence had been derived by a defense expert, causing the

court to voice concerns about the process's reliability, suggesting that the video could have been altered in an unverified way.

The credibility of this AI-enhanced video was assessed on multiple levels, beginning with a Frye hearing (a legal proceeding that determines if a scientific method is permissible in court)—as the use of AI in this situation was deemed a “novel scientific technique.” Based on the standards of Frye, the court had to determine if the technique was accepted by the relevant scientific community on a general basis. Within the forensic video analysis community specifically, the court uncovered that the method utilized by the defense expert was not generally accepted. In addition, the court went on to raise concerns regarding the limited transparency (“opaqueness”) of the aforementioned AI-enhancing process, introducing the possibility of the original contents of the video being altered.

Similar suspicions about the unjust utilization of AI enhancement and other techniques in court were brought about in the Supreme Court case *Huang v. Tesla* in 2018. This case was instigated via a death lawsuit filed against Tesla, as Walter Huang (an engineer for Apple) was killed in a crash involving a Tesla Model X. According to the lawsuit, the Tesla vehicle's Autopilot system was defective, in failing to prevent the fatal accident—even testifying that that Tesla had been aware of these systematic shortcomings and refused to address them. To justify the crash, Tesla argued that Huang had mishandled the Autopilot system by inadequately supervising the vehicle and being “distracted,” employing evidence displaying that his hands were not on the wheel at the time of the crash. Beyond raising awareness and worries regarding Tesla's driver-assistance technology, the case highlighted the difficulty of determining liability in accidents or offenses concerning automated or AI-correlated systems.

While *Huang v Tesla* didn't outwardly exhibit the aforesaid dangers of AI in court (notably, evidence of deepfakes and fabrication), it readily conveyed the outputs of AI-related events and the complexities that arise when they become central to legal proceedings. Further beyond the scope of the legalities of AI concerning artificial evidence and fabricated defense in court, misleading AI-generated legal research has also occurred recently. Such research brings to light a more integral issue within the generalized legal process, with various accounts of lawyers utilizing AI for legal research. Essentially, this research produced "hallucinated" or untraceable case citations; this occurred in *Kohls v. Ellison* (2024) in Minnesota. In a turn of ironic events, an expert witness (on the uncertainty behind AI's employment in court) submitted a legal report consisting of nonexistent (AI-fabricated) citations. Expectedly, the court rejected the document, and explained the dangers and unreliability of AI output during legal proceedings; also deeming it a "waste of judicial resources."

A widely renowned incident displaying AI's gravity in legal issues was shown through the case of Pikesville High School in Baltimore County, Maryland, in January of 2024. Instigated when an antisemitic audio clip featuring the school principal (Eric Eiswert) was circulated widely, the case shed light on the strength of counterfeit media produced by AI. To elaborate, after the audio clip (which consisted of racist remarks) was circulated, Eiswert received various threats and was placed on administrative leave until authorities later determined that the video clip had been "deepfaked" or fabricated with the use of AI. With the implementation of forensic analysis, the falsely generated audio recording was tied back to the school's athletic director, Dazhon Darien, who was then arrested. The significance of this Pikesville school case was largely transparent, despite not occurring in a courtroom; the case thoroughly depicted the surfacing dangers and expanding range of legal controversies that corresponded to AI.

Exceeding recent struggles with AI both in courts nationwide and in public facilities, legal experts have begun to disclose increasingly plausible scenarios entailing AI-deepfakes in future legal contexts. For instance, manipulated surveillance footage in cases regarding civil disputes and criminal investigations could come into common use in order to either implicate (blame) or exculpate (“prove” innocent) certain individuals. Another major concern unveils itself as forged documents and financial records become more likely to appear in court, considering AI’s current abilities in forging realistic counterfeit documents and contracts/financial statements. In addition, in accordance with the recently resolved event at Pikesville High School, the amount of fabricated voicemails/voice messages may increase as well. The prevalence of these deep fake voice messages, based on research from legal experts and scholars, could concerningly come to play a role in divorce cases, creating nonexistent evidence of abuse in relationships. A major concern, which corresponds to the rudimentary elements of court hearings and legal proceedings, could also be disrupted in the long term if synthetic witness testimony comes to be.

In accounting for the upsurge of AI employment in court and inappropriately integrating it into the legal system, foundational attributes of the national legal process (and its utilization on local levels) are disrupted. By way of elucidation, the decomposition of justice is essentially catalyzed by the introduction of unreliable and unsafe employment of AI into legal issues. Most notably, the integrity of the dominant and structural legal process of the United States is undermined by the circulation of AI utilization in cases and proceedings, whether by way of research, conviction, evidence, etc. In full, AI heavily increases the risk of wrongful convictions in court (and outside of it, as seen in the case of Pikesville High School in 2024) when fabricated evidence is found and submitted by defense experts, lawyers, or witnesses.

Furthermore, an overbearing failure to not only preserve the probity of the nationwide legal system but also to urge accountability for guilty parties has been prevalent as a result of the irresponsible implementation of AI's skill sets. Major disputes regarding the authenticity and reliability/transparency of AI algorithms and their occurrence in court are plausible gateways for heightened complexity and extended legal arguments. Also significantly instrumental in the legal system's continuity and overall success is public trust and regard for the current measures and systems in place. The application of AI as a part of new legal processes and research has come to be known as susceptible to corruption and misuse by the public, and hence, beliefs of simple manipulation of certain cases have become widespread. Therefore, risks that have become inherent with the utilization of AI-generated evidence, deep fakes, and even cases stemming from AI in court demonstrate the vast scope and potential of dangers surrounding its continued embedment into legal issues. Ensuring that the integrity and fundamental goal of the pursuit of justice remain central values in our judicial system (and beyond the government) should be of top priority.

## Works Cited

Akerman LLP. "The Challenges of Integrating AI-Generated Evidence Into the Legal System."

Accessed May 15th, 2025. <https://www.akerman.com>

Minami Tamaki LLP. "Family of Tesla Car Crash Victim Walter Huang Files Lawsuit." Accessed

May 16th, 2025. <https://www.minamitamaki.com>

CNN News. "Tesla settles with Apple engineer's family who said Autopilot caused his fatal

crash." Accessed May 16th, 2025. <https://www.cnn.com>

Justia Law. "Kohls et al v. Ellison et al, No. 0:2024cv03754 - Document 46 (D. Minn. 2025)."

Accessed May 17th, 2025. <https://law.justia.com>

CBS News. "Former Pikesville High School principal sues Baltimore County Schools over racist

AI case." Accessed May 18th, 2025. <https://www.cbsnews.com>

Stanford University. "AI on Trial: Legal Models Hallucinate in 1 out of 6 (or More)

Benchmarking Queries." Accessed May 20th, 2025. <https://hai.stanford.edu>