

THE EVIDENTIARY AUTHORITY OF AI AND THE RIGHT TO A FAIR TRIAL IN CRIMINAL PROCEEDINGS

Shahzad Khalid

Doctoral Researcher, Brunel University London

shahzad.khalid@brunel.ac.uk

DOI: <https://doi.org/10.5281/zenodo.21850234>

Keywords

Artificial Intelligence; Criminal Proceedings; AI-Generated Evidence; Fair Trial; Criminal Evidence Law; Algorithmic Accountability; Judicial Oversight.

Article History

Received: 07 January 2026

Accepted: 17 March 2026

Published: 31 March 2026

Copyright @Author

Corresponding Author: *

Shahzad Khalid

Abstract

The role of AI in the criminal justice system is growing, and so is its impact on the investigative process, forensic analysis, and evidentiary procedures. Facial recognition, probabilistic DNA interpretation, predictive algorithms, and digital forensic analytics are just a few examples of AI-driven technologies that provide valuable benefits in data processing and enhance investigative efficiency. But as their evidentiary function has become increasingly prominent, fundamental legal issues of reliability, transparency, explainability, and procedural fairness arise. In this article, the evidentiary status of AI-generated evidence in the criminal process is critically discussed, and its compatibility with the right to a fair trial is assessed. It examines the legal principles governing the admissibility and evidentiary value of AI-supported evidence and discusses issues arising from algorithmic opacity, confidential software, algorithmic bias, and the lack of adversarial challenge (Munir, Khalid, & Noreen, 2025). The discussion also reviews key judicial rulings, such as *State v Loomis* and *R (Bridges) v Chief Constable of South Wales Police*, as well as recent regulatory initiatives, including the EU AI Act and the Framework Convention on AI, Human Rights, CETS No. 225, adopted by the CoE. By comparing the legal frameworks in the U.S., the U.K., and the E.U., it becomes apparent that an international consensus is coalescing around the view that AI should be used to assist, rather than replace, human judges. The legitimacy of AI-generated evidence hinges on several factors, including the transparency of algorithms, independent validation, human oversight, disclosure requirements, and judicial oversight. It finds that rights-based regulatory systems are necessary to ensure that technological innovation improves the efficiency and accuracy of the criminal justice system without limiting the accused's rights to a fair trial, judicial independence, or due process.

INTRODUCTION

AI has rapidly become a new force in the administration of criminal justice, bringing powerful computational tools to bear on large amounts of data, patterns beyond human perception, and functions that support investigations and adjudication (Munir, 2025). In recent years, law enforcement and prosecution

agencies have increasingly turned to AI-based tools and systems such as facial recognition, predictive policing, automated digital forensic analysis, voice recognition software, and machine learning models for evidence evaluation (Ulenaers, 2020). The technologies hold great potential to enable faster, more effective, and more consistent criminal investigations and to help judicial systems

respond to increasingly sophisticated forms of crime, such as cybercrime, financial crime, and transnational crime (Raza, Khalid, Yasin, Kasuri, & Munir, 2023). With the growing prominence of AI in criminal investigations, however, its generated content is no longer considered a mere investigative tool, but rather evidence that can be used to sway decisions of guilt or innocence. In light of the growing reliance on machine-generated proof and the crucial need to uncover the truth during criminal hearings, it is clear that this technology poses serious legal challenges. Also, criminal proceedings have a larger procedural component that relies on individual guarantees, posing a serious barrier to deploying automated systems. The legitimacy of criminal adjudication relies on the reliability, transparency, and verifiability of the process, as well as the accused party's ability to challenge evidence (Nowotko, 2021).

Unlike traditional types of evidence, many AI systems rely on complex algorithms that make the reasoning behind the evidence inaccessible to judges, lawyers, and litigants. The "black-box" effect is a concerning issue that calls into question the effectiveness of courts in examining AI results and determining their probative weight under traditional principles of criminal justice (Kelly, 2021). Concerns regarding the precision of computational models and the quality of input information, as well as explainability, accountability, and reproducibility, are thus increasingly playing a prominent role in current discussions on the admissibility and evidentiary weight of AI-generated evidence. Such worries are not just technical; they also affect one of the most basic rights in any criminal system: the right to a fair trial. Fair trial rules ensure equality of arms – between the prosecution and the defense; the presumption of innocence; the impartiality of the judge; and the ability to challenge and rebut negative evidence before an independent adjudicator. In instances where AI-generated evidence affects an investigation, prosecution, or judicial decision but lacks adequate transparency or opportunities to review the evidence and its underlying determinations, there is a significant

risk to the integrity of the process (Kocagül & Kartal, 2026).

Defendants' lack of comprehension or ability to effectively challenge algorithmic reasoning may diminish traditional protections that have historically been a hallmark of criminal justice, relative to administrative and/or technological decision-making. As a result, it is essential to rethink many well-established rules of evidence concerning admissibility, the reliability of evidence, and the judge's evaluation (Trang et al., 2024). Judicial decisions, legislative efforts, and regulatory approaches aimed at balancing technological innovations with constitutional and human rights protections are now emerging across legal systems worldwide in response to these developments. However, substantial gaps remain regarding the legal nature of AI-generated evidence, the level of judicial involvement required, and the safeguards needed to ensure the integrity of the criminal process. The lack of harmonization has left the future of AI's evidentiary value and the question of when courts should accept algorithm-generated inferences in criminal cases uncertain (Molbæk-Steensig & Quemy, 2023).

In light of this, the current study examines the probative value of machine-assisted outputs in criminal proceedings with respect to ensuring fair judicial treatment. It starts by examining how automated systems have evolved across legal processes and how algorithm-derived materials have emerged as a new type of proof (Kumar et al., 2023). The discussion then reviews the principles underlying the admissibility, reliability, transparency, and judicial evaluation of AI-based evidence, and then considers the implications for procedural fairness, equality of arms, and the defense rights of algorithmic decision-making. Finally, the article reflects on comparative legal developments and suggests that a framework grounded in fundamental entitlements may help ensure that the integration of automated systems into penal judicial structures enhances operational effectiveness while preserving the core tenets underpinning lawful governance.

The Evolution of Artificial Intelligence in Criminal Justice Systems

Institutional Expansion of Artificial Intelligence in Criminal Justice

Artificial intelligence (AI) has become a reality in criminal justice systems, moving from a few experiments to a broad institutional practice. Over the past decade, law enforcement, prosecution services, forensic laboratories, and courts have turned to AI-powered tools to streamline their workflows, manage large volumes of digital evidence, and support their decision-making. This has been driven in large part by the exponential increase in digital data. Today, nothing is excluded from being included on a smartphone, stored in a cloud computer, recorded by a video camera, shared on social media, used in financial transactions, or captured by Internet-connected devices (Brayne & Christin, 2021). There is no data that isn't on a smartphone, stored in a cloud computer, captured by video cameras, recorded in social media, processed by financial transactions, or generated by Internet-connected devices, and all this data is routinely accessed during criminal investigations today. Examining this evidence manually is increasingly impractical, and criminal justice institutions are increasingly using machine-learning systems that can analyze millions of records in a fraction of the time it used to take to examine all the evidence.

Predictive analytics is one of the first and most impactful instances of AI (Abbasi, Munir, Jayaram, & Rivas, 2025). In the U.S., courts have applied the COMPAS algorithm to help estimate recidivism risk in decisions to grant bail, sentence offenders, or parole. It considers several factors such as criminal record, behavioral traits, and demographics to calculate a risk assessment score. While intended to increase consistency in the judicial process, the use of COMPAS sparked legal controversy in 2016 when the Wisconsin Supreme Court recognized that computational risk assessments could be helpful in the sentencing process but could not supplant the judge's discretion (Anastasopoulou, 2025). The court also acknowledged concerns about the algorithm's proprietary nature and the defendants' inability to review its methodology. Similarly, the police have

begun using facial recognition to identify criminals in surveillance footage and public image databases. It's alleged that one of the most talked-about facial recognition platforms, Clearview AI, built a database of over 30 billion public faces gleaned from social media and websites.

The software has been employed by thousands of North American and other crime-fighting agencies in the investigation of child exploitation, organized crime, homicide, and terrorism. But its use has sparked a range of legal issues, including concerns about privacy rights, data protection, and the lack of regulation (Raza, Munir, Ali, Othi, & Hussain, 2024). The UK has also seen a growth in the use of LFR by the police, including the MPS. Within the recent judgment of *Bridges v South Wales Police* [2020] EWCA Civ 1058, England's appellate tribunal determined that automated face-scanning systems, by their very nature, are not unlawful, but that they have not been deployed in a way that provides enough clarity around their legal protections regarding data protection, equality obligations, and operational discretion (Garrett & Rudin, 2023). The decision set a precedent for the proper use of AI in criminal investigations, highlighting the need for legal accountability for AI tools. The judgment highlighted legal accountability issues surrounding the use of AI tools, making an important point about the proper use of technology in criminal investigations. AI is now much more than a tool for policing; it is a key part of digital forensic investigations. Modern forensic software can be used to retrieve lost files, decrypt messages, flag tampered images, establish timelines, categorize e-documents, and analyze large volumes of data recovered from electronic devices. These have drastically shortened investigative time and enhanced forensic capabilities to identify potentially relevant evidence in intricate cybercrime, terrorism financing, money laundering, and organized criminal network investigations (Gyau et al., 2026).

Artificial Intelligence as an Emerging Source of Criminal Evidence

AI has revolutionized the type of evidence brought before a criminal court. The traditional methods

of criminal proceedings were witness testimony, documents, physical evidence, and expert testimony. AI has added a new category of machine-made evidence that goes beyond information processing to the creation of conclusions based on analysis that can influence decisions about criminal liability. The area of forensic science has seen the greatest strides. Forensic genetics has been revolutionized by DNA mixture interpretation software, such as STRmix and TrueAllele, which laboratories use to analyze complex biological samples containing genetic material from multiple contributors (Cascone, 2025). Probabilistic genotyping software uses advanced mathematical algorithms to assign a probability to the likelihood of a match between a suspect's DNA profile and a DNA profile from a crime scene, whereas traditional statistical methods often fall short in analyzing such mixtures. Such systems have been accepted as evidence in many criminal prosecutions throughout the US, Australia, New Zealand and the UK. However, defense attorneys have, in some cases, objected that algorithmic DNA interpretations were admissible only if they were made available without proprietary source code, thereby making it difficult for the defendant to effectively challenge the evidence.

AI has also revolutionized biometric recognition. Using machine-learning methods, Automated Fingerprint Identification Systems (AFIS) search a large national fingerprint database for latent fingerprints found at a crime scene within minutes. An increasing role of voice recognition software in organized crime and terrorist investigations has been to compare the acoustic qualities of intercepted communications (Borgesano et al., 2025). Similarly, AI-powered ballistic comparison tools analyze the micro-markings on spent cartridge cases and bullets to determine whether they came from a specific gun, aiding in standard ballistic analysis. Digital investigations have been a stark example of the rapid growth of AI-generated evidence in the era of mass data volumes. The proliferation of AI-generated evidence has been especially noticeable in digital investigations. Investigators can use advanced analytical platforms such as Cellebrite

and Magnet AXIOM to recover, classify, and analyze large volumes of electronic evidence from smartphones, computers, and cloud storage accounts. Such platforms use AI to find relevant conversations, build communication graphs, retrieve deleted content, flag suspicious activity and behavior, and conduct relevance-based investigations of digital evidence. These have greatly boosted the efficiency of criminal investigations but have also placed greater evidentiary burdens on algorithmic systems whose reasoning might not always be obvious (Sani, 2025).

Regulatory Developments and the Changing Legal Landscape

AI's swift integration into criminal justice has prompted regulations from both governments and international bodies addressing legal issues surrounding the deployment of automated systems in policing. The EU has passed the AI Act, which classifies automated platforms used in policing, judicial processes, population movement, and frontier oversight as elevated-threat deployments. The systems must also comply with strict requirements related to transparency, human involvement, accuracy, cybersecurity, record-keeping, and risk management, among other things, before they can be put into operation. Some of the above AI applications are banned or subject to very strict regulation because of their impact on fundamental rights, such as indiscriminate biometric surveillance and social scoring by public authorities. The Convention on AI, Rights, Democratic Governance and Legal Standards, approved by Europe's intergovernmental body, represents an unprecedented global accord focused squarely on guaranteeing that automated technologies remain consistent with individual liberties and shared societal principles (Shandilya et al., 2024).

States are obligated under the Convention to put in place safeguards to ensure fairness, accountability, transparency, effective remedies, and meaningful human oversight in any cases where AI systems are involved in rights-protected activities, such as those occurring in criminal proceedings. In the United States, although there

is no comprehensive federal law regulating AI in the criminal justice field, NIST has issued the AI Risk Management Framework, which outlines principles for fostering trustworthy AI systems, including validity, reliability, explainability, privacy, security, and accountability. A number of federal agencies have also provided guidance stating that algorithmic systems cannot be used in place of human judgment in matters involving constitutional rights (Zaidan & Ibrahim, 2024).

The Evidentiary Authority of Artificial Intelligence in Criminal Proceedings Admissibility and Reliability of AI-Generated Evidence

AI is becoming an integral part of criminal investigations, thereby changing the nature of evidence. The use of AI in criminal investigations has, in many ways, revolutionized the nature of evidence by adding machine-generated items. AI-generated evidence is created using algorithms and large datasets, unlike traditional evidence from real-world observations, expert examinations, or documentary sources. Thus, the reliability, transparency, and reproducibility of the underlying computational process are as important as the traditional evidentiary requirements of relevance, authenticity, and probative value. The judicial gatekeeping approach is still used in most common-law jurisdictions to assess scientific and technical evidence. In *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), the highest court in the US set standards for assessing the reliability of expert testimony. These standards included experimental confirmation, external examination, established error margins, procedural guidelines, and broad recognition among relevant specialists (Grossman & Grimm, 2024).

While *Daubert* was not specifically about AI in the context of forensic evidence, the rule's tenets have become increasingly relevant to AI-generated forensic evidence, especially when AI conclusions are offered as scientific findings. Likewise, the precedent of *Frye v. United States* (1923) holds that novel scientific methods must be generally accepted before they become admissible as evidence, a standard still applied by some

jurisdictions in assessing new forensic technologies. The practical application of these standards has gained particular importance in connection with probabilistic genotyping software. Complex DNA mixtures that can't be analyzed using conventional statistical techniques have been introduced in criminal prosecutions by systems such as STRmix and TrueAllele. They are equipped with complex mathematical algorithms that calculate the likelihood that a suspect had made a genetic contribution to a crime-scene sample. Many courts have admitted evidence of this nature because the algorithms have been extensively validated by scientists, but defense counsel have repeatedly argued that the algorithms could not be independently verified through proprietary source code. The highlighted conflicts show that the reliability of the evidence increasingly depends on scientific practice and transparency, which allow for effective scrutiny by the judiciary (Stone & Vaidyan, 2026; Zhang, 2025).

Judicial Scrutiny of Algorithmic Decision-Making

The ability of algorithms to reason and the judiciary's capacity to comprehend and assess that reasoning determine AI's authority. The authority of AI is not only computational but also the judiciary's ability to understand and evaluate algorithmic reasoning. Traditionally, criminal courts have granted considerable discretion to scientists and experts in assessing the significance of scientific evidence and expert testimony. This judicial role is undermined by the use of AI, as many machine-learning systems are "black-box" models whose decision-making processes are difficult for judges, lawyers, or even software developers to decipher. Algorithmic opacity was clearly important in a case that came to light in 2016, *State v. Loomis*. The Wisconsin Supreme Court addressed the use of the COMPAS algorithm in the sentencing process and accepted that algorithmic assessments could be useful for assessing the risk of recidivism. But the court warned that due to the proprietary nature of the software, it could not examine in depth the methodology behind the creation of the

COMPAS scores, which the court prohibited judges from using as a basis for determining sentencing (Vidaki & Papakonstantinou, 2025). Rather, the court expressed that AI-generated evaluations were to be used in tandem with – not in place of independent judicial thought. The ruling brought to light a number of issues with automated accountability, proper balance, and unclear governance frameworks governing machine deployment, despite the hearing's primary focus on observation. In accordance with the principle that sophisticated equipment is nonetheless subject to fundamental legal concepts requiring permission and accountability, the appellate board also concluded that the use of observation rosters and biometric identification techniques lacked sufficient legislative protections. The scrutiny of proprietary forensic software has also increased in the Courts. In several criminal prosecutions, defense counsel has contended that denying access to source code undermines the adversarial process by depriving the defense of the opportunity to examine the algorithm's methodology (Dancy & Zalnieriute, 2026).

Transparency, Explainability and Evidentiary Weight

The legitimacy of algorithmic evidence ultimately rests on its transparency and explainability in relation to the basic principles of the legal system. The level of explainability and transparency of algorithmic evidence with respect to the bottom-line principles of the legal system is the test for the legitimacy of the legal authority of algorithmic conclusions. Evidence required for criminal proceedings must contain meaning to it for both parties and be subject to meaningful examination by the court. Many AI systems, unlike human experts, can provide probabilistic answers and do not offer intelligible explanations of how certain conclusions were reached through testimony and cross-examination. It also makes it hard for judges to establish the weight of evidence to be given to algorithmic conclusions (Morison & McInerney, 2025).

The providers of these systems must have an extensive risk management process, extensive technical documentation, ensure data quality,

maintain operational records, and provide meaningful human oversight throughout the system's life cycle. These requirements are based on the recognition in legislation that algorithmic evidence does not have legal authority merely because of the degree of sophistication in the algorithm or technology, but because of evidence of reliability, traceability, and accountability. International progress is on the same theme. Several requirements are established by the Council of Europe's treaties on automated systems, basic freedoms, popular governance, and legal supremacy, requiring all participating States to ensure that systems that affect the enjoyment of fundamental rights are subject to frameworks that guarantee transparency, effective remedies, and human supervision. The safeguards are especially important in criminal proceedings, where AI-generated evidence could be a pivotal factor in convictions, sentencing, or limitations on personal freedom (Bialek, 2026).

In an evidentiary sense, then, AI must not be considered a decision-maker in its own right, but rather a high-tech tool for evidence that is subject to the independent assessment of the judge. The courts have the power to evaluate the quality of the training data, algorithmic validation, error rates, the extent to which the results can be reproduced, the extent to which they meet recognized forensic practices and standards, and whether there are meaningful challenges to be had by those who are averse to the algorithm. No matter how complex the computation, evidence obtained from invalid systems or systems with limited or weak validation ought to be accorded less weight. As AI evolves, the fact that a technology can prove its authority does not necessarily mean it actually does. (Chemerinsky, 2021).

Artificial Intelligence and the Right to a Fair Trial

As AI becomes an increasingly integral part of the criminal process, the notion of a fair trial has undergone unprecedented change, especially when algorithmic systems exert significant influence on decisions in the investigation, arrest, prosecution, bail, sentencing, and assessment of evidence in a case. AI can be incredibly useful in

terms of efficiency, uniformity, and analytical power, but it also raises deep constitutional and human rights questions when algorithmic decisions affect an individual's liberty without proper safeguards in the procedures for its development. Article 10 of the UD (1948), Article 14 of the ICCPR, and Article 6 of the ECHR all provide equitable adjudication as a fundamental cornerstone of criminal justice. When taken as a whole, these clauses ensure that the accused has adequate resources to present their case, that the bench is impartial and independent, that both parties have equal opportunity, that the accused is presumed innocent, and that they can challenge opposing evidence (Sroka, 2024).

The guarantees were made on the basis that evidence comes from recognizable human actors who are subject to cross-examination for their observations, expertise, and credibility. The traditional model is undermined by the emergence of this new type of evidence, which is AI-generated, as it is often the case for AI-generated evidence that the conclusions are drawn by a computational process that the judge, prosecutor, defense counsel, and defendant are not able to understand. Therefore, the core question is not whether AI can enhance the criminal justice system, but whether it can be done in a fair, due-process manner. *State v. Loomis* (2016) is one of the most important recent cases involving this dilemma in which the Wisconsin Supreme Court weighed the constitutionality of employing the COMPAS risk assessment algorithm in sentencing. While judges were allowed to take into account COMPAS scores, the Court explicitly noted that the algorithmic assessment should not be binding on the court, as the defendants were not permitted to review the proprietary algorithm used to generate the risk score. The Court acknowledged the issues this lack of access to algorithmic methodology posed for the defendant in contesting negative evidence, thereby giving rise to fundamental due process issues (Selçuk et al., 2025).

While the case was not about criminal court proceedings, its reasoning helped to affirm that AI systems used in law enforcement are bound by the rules of law, proportionality, accountability, and

judicial review. The discussion over fair trial rights has been renewed and bolstered by concerns about algorithmic discrimination. Independent research testing the COMPAS algorithm indicated that the algorithm was more likely to predict recidivism for Black defendants than for White defendants, thereby posing questions about the algorithm's racial bias in training datasets (Munir, 2025; 2026). Despite the disagreements the developers raised about some of these conclusions, the debate only highlighted how algorithmic results are not necessarily neutral and can reinforce and perpetuate past patterns of discrimination within the data from which they are derived. The same type of issues have been raised about facial recognition systems. A series of tests by the United States NIST revealed that many of the facial recognition algorithms had much higher false identification rates for female, older, and racial and ethnic minority subjects than they did for White male subjects. The differences have even more pronounced importance in criminal investigations, where a false "match" could lead to arrest, surveillance, or prosecution when there is no corroborating evidence. These risks directly undermine the presumption of innocence and create an impression of scientific certainty that could unduly influence investigators, prosecutors, judges, or juries (Hloviuk & Hryniuk, 2025).

Comparative Jurisprudence on AI Evidence in Criminal Proceedings

United States

AI has become a frontrunner in the use of technology in the criminal justice arena, specifically in predictive policing, forensic science, risk assessment, and digital investigations (Munir, 2024). Even with this technological progress, the admissibility of AI-generated evidence is largely governed by the current rules of evidence rather than specific legislation. Federal Evidence Standard 702, alongside the benchmarks established in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), governs the evaluation of specialized and technical evidence in federal courts; those same criteria determine whether expert testimony may be introduced in federal proceedings. The standards call for a court to

consider whether an AI-based technique has been validated by scientific research, is peer-reviewed, has a measurable error rate, and has been accepted by the scientific community in the field. In *Wisconsin v. Loomis* (2016), the Wisconsin Supreme Court examined potential constitutional issues with the use of the COMPAS algorithm at sentencing. The Court allowed the court to consider algorithmic risk profiles, but also determined that sentencing could not be based solely on COMPAS, as defendants had no access to the algorithm used by the computer program or the ability to critique the methodology used to create the assessment (Benjamin, 2025).

The judgment confirmed that while AI can serve as a tool for guidance, it is not a replacement for the human court's decision-making power, and that any decisions made using AI should be made with caution (Munir, 2024). The use of AI in criminal prosecutions has also become commonplace. Sophisticated DNA mixture analysis software, such as probabilistic genotyping software like STRmix and TrueAllele, has been accepted in many cases involving complex DNA mixtures, and AI-powered analysis software like Cellebrite, Magnet AXIOM, and FTK can automatically recover, organize, and analyze digital evidence from electronic devices. However, the issues of disclosure of proprietary source code, algorithmic explainability, and defense access to technical documentation remain ongoing in American courtrooms. While the US has yet to have federal legislation governing AI in the criminal justice sector in its entirety, the NIST Presidential Directive for Reliable, Protected, and Dependable Creation and Implementation of Intelligent Tools and the Automated Systems Threat Mitigation Blueprint have reinforced calls for transparency, accountability, and human oversight (Meghwal, 2025).

United Kingdom

The UK has taken a more measured stance on the deployment of automated tools within penal legal frameworks, prioritizing legality, proportionality, and judicial accountability. Police forces are increasingly using AI technologies in their activities, including in facial recognition, digital

forensic analysis, intelligence activities, and evidence management, but these are also subject to extensive judicial scrutiny under domestic human rights legislation and data protection law (Khan, Ahmad, & Munir, 2025; Munir, Irfan, & Bashir, 2024). The main laws controlling machine-assisted investigations into misconduct are the 1998 statute on individual freedoms, the 2010 fairness legislation, the 2018 measure managing personal records, and the national adoption of European information standards. The most significant ruling at that level is the appeal tribunal's decision in the Bridges case involving the Welsh police. That court examined the use of instantaneous biometric scanning by that police force and found that, while it did not constitute illegal activity, the regulatory framework was not sufficiently precise, requiring officials to exercise caution when exercising their power, in addition to their obligations to protect information rights and conduct fairness evaluations. The judgment called for greater transparency in the creation of watch lists and noted that emerging technologies should be grounded in fundamental rights and the rule of law (Ibrahim et al., 2026).

Other AI-led capabilities for British policing have included digital evidence analysis from mobile devices, encrypted communications, CCTV footage, and online crimes. The National Police Chiefs' Council and the College of Policing have provided guidance on promoting the ethical use of AI in its deployment, and the Crown Prosecution Service is increasingly using AI-assisted digital review technology to manage huge volumes of evidence. However, courts are unwilling to accord algorithmic results definitive credibility or evidentiary value without expert evidence and customary corroboration. Institutional safeguards have also been established in the United Kingdom: the Information Commissioner's Office (ICO) and the Center for Data Ethics and Innovation (CDEI) have both called for explainability, fairness and accountability in public-sector AI systems. The UK has taken a principles-based approach to AI regulation, building on existing principles and rules and judicial oversight, rather than implementing

broad legislation similar to the EU's GDPR (Irfan et al., 2025).

European Union

The EU has codified its obligations regarding high-risk AI systems in the context of law enforcement and the administration of justice, whereas other jurisdictions mainly rely on judicial interpretation. At the heart of this is the Artificial Intelligence Act, which identifies the following AI systems as high risk and imposes mandatory legal protections: AI systems used for criminal investigation, biometric identification, migration control, and judicial decision-making. In terms of risk mitigation, input reliability, supporting documentation, activity logging, network security, transparency, human oversight, and subsequent monitoring, the legislation on intelligent systems imposes significant obligations on developers and operators. Some automated processes, such as citizen rating by government agencies and undefined instantaneous body scanning in public spaces, are either forbidden or allowed only in very specific situations. The provisions are in line with the EU's efforts to ensure that AI won't compromise privacy, equality, and fair process (Moroşteş, 2025).

Art. 6 ECHR, Arts. 47 & 48 CFREU, GDPR, and the CoE AI Convention (CETS No. 225) reinforce the EU legal order. Together, these tools demand that AI systems involved in criminal proceedings be consistent, have effective judicial remedies, be nondiscriminatory, and have meaningful human oversight. The European judiciary has also shown greater sensitivity towards automated decision-making. The European approach, therefore, does not focus so much on the efficiency of AI but rather on its legitimacy, which is based on the principles of respect for the dignity of human beings, democracy, and fundamental rights, and a legal framework (Mustafa & Ahmed, 2025).

Emerging Challenges and the Future Regulatory Framework

Emerging Challenges in the Evidentiary Use of Artificial Intelligence

Opacity of algorithmic decision-making is the foremost challenge. Forensic science, facial

recognition, predictive systems, and digital investigations are among the many areas that use AI systems with proprietary machine-learning models whose inner workings are not fully understood even by the AI systems' creators. This "black-box" aspect directly contradicts one of the fundamental rules of criminal evidence law: that evidence must be subject to meaningful judicial review and/or challenge. It is difficult to gauge the reliability and probative value of an evidentiary conclusion when the judge, prosecutor, or defense counsel cannot understand the reasoning that led to it. The second challenge is algorithmic bias stemming from training sets that do not reflect diverse populations (Organization, 2026).

The COMPAS risk assessment system sparked controversy over its ability to produce different outcomes for different demographic groups, and a set of performance evaluations conducted by the NIST found that several facial recognition algorithms were more precise at recognizing the age, gender, and ethnicity of certain groups than others. The same issues have arisen with the automated license plate recognition system, predictive policing programs such as PredPol, and biometric surveillance programs based on past policing data. Historical data may present patterns that have a discriminatory effect, which AI systems can amplify at an even greater scale, thereby threatening equality under the law and public trust in criminal justice institutions (Munir, Abbasi, Wilson, & Colombo, 2025). Another challenge to procedural fairness is that commercial AI software is often proprietary. The source code, training methods, and optimizations for forensic platforms such as TrueAllele, STRmix, Cellebrite, and Magnet AXIOM are often protected as trade secrets and developed by private companies. However, it is very hard for a defense lawyer to obtain copies of technical documents that would assess whether an algorithm was used correctly in a specific case (Maaitah et al., 2026).

This restriction may undermine the adversarial system by limiting cross-examination and the defendant's ability to challenge evidence that may have a significant impact. A new challenge is to ensure cybersecurity and evidentiary integrity. Criminal justice data is used in vast quantities by

AI systems, such as biometric information, DNA profiles, digital communications, and forensic databases, all of which contain sensitive information. The integrity of the algorithmic results may be compromised by one or more of the following: a cyberattack, unauthorized access, model manipulation, or data poisoning, which can lead to inaccurate or misleading evidentiary conclusions. The current chain-of-custody rules are mostly tailored for physical evidence and do not necessarily meet the needs of algorithmic changes that may occur during automated data processing (Siemens et al., 2025).

Toward a Future Regulatory Framework for AI Evidence

To meet these challenges, there needs to be a regulatory regime that treats evidence created with AI as a separate category of evidence, but also respects the rule of law and constitutional values. The first step in a framework like this is to set up requirements for compulsory transparency. AI systems used in criminal justice should have all technical documentation that explains the architecture of the systems, how they are trained, validated, how they are measured, how well they perform, what errors they can make, software upgrades, and any limitations on their use. The courts should be given the power to require such disclosure in criminal proceedings, with due regard to well-balanced procedures to safeguard legitimate commercial interests without abusing defence rights when algorithmic evidence is used in a criminal proceeding. Secondly, law should mandate independent scientific validation and certification before AI systems are used in law-enforcement, forensic laboratories or in the judicial system. AI systems must be subjected to regular evaluations by independent technical entities to check for accuracy, reproducibility, security in the face of cyber attacks, adversarial robustness and potential discriminatory effects, in a manner similar to ISO/IEC 17025 accreditation requirements for forensic laboratories (Ford et al., 2026).

Certification does not imply forever certification, but rather a re-certification should be done regularly when algorithms are retrained or

materially changed. The regulatory system to be developed should also implement procedural protections for defendants. There is a need to explicitly mandate the rights of the defence to challenge algorithmic evidence, disclosure of any technical information relevant to the algorithm, appoint independent computational experts, and to seek judicial review of the use of AI tools in investigations. The role of court is to ensure that the prosecutors prove that the evidence provided by AI is relevant, has been scientifically validated, has documented reliability, and adherence to known operational standards before it is used as evidence. But it's also vital to maintain relevant human control. AI should not act as an independent adjudicator but rather be a decision support tool. Any issues concerning whether a person has been arrested, charged, convicted, sentenced or assessed evidence should be the prerogative of independent judges making their decisions based on reason and law (Kumari, 2025). Any legal system, then, should explicitly prohibit fully-automated decisions with direct consequences for criminal liability without any substantive human review. International cooperation will also be required. States should take steps to build on these efforts to establish internationally agreed-upon principles and practices for explainability, algorithmic auditing, evidentiary disclosure, digital chain of custody, and certification of forensic AI systems. The training of judges, prosecutors, and defense counsel should also be expanded to enhance technical literacy to better prepare them to critically assess evidence generated by AI (Lenarczyk et al., 2025).

Conclusion

AI has profoundly reshaped the handling of evidence in criminal justice, influencing its collection, analysis, and presentation, with AI-generated evidence increasingly playing a significant role in the criminal justice process. Facial recognition, probabilistic DNA interpretation, digital forensic analytics, predictive algorithms, and biometric identification are notable technologies that have shown great promise in boosting investigative efficiency and

forensic accuracy. The study has indicated, however, that nothing can serve as an evidentiary source of power unless it is technologically sophisticated. Ultimately, the validity of AI-generated evidence will depend on its adherence to well-established principles of criminal evidence law, such as reliability, authenticity, transparency, explainability, and meaningful judicial review. The comparative analysis confirms that, while the EU, the UK, and the US have taken different regulatory approaches, they share a similar legal stance: AI is not a replacement for independent judicial thinking. The courts, in cases like *State v Loomis* and *R (Bridges) v Chief Constable of South Wales Police*, are increasingly hesitant in giving final credence to the output of algorithms when the algorithms are proprietary, cannot be challenged, or cannot be explained. Similarly, the CoE's treaty on automated systems, freedoms, governance, and legal standards and the EU AI Act reflect a global trend toward entitlement-focused oversight, acknowledging that machine platforms used in criminal proceedings must continue to be consistent with legitimate authority, balanced advocacy, judicial autonomy, and procedural fairness (Munir, Khan, & Ahmed, 2020). This study also confirms that the lack of harmonized legal principles regarding algorithm transparency, disclosure requirements, independent validation, cybersecurity, and accountability is the key issue, rather than the issue of AI admissibility. Even if technology is correct, the core principles of a fair trial are undermined when defendants are unable to access the computations used to examine the processes that produced the incriminating evidence. AI, therefore, is not a decision-making entity but rather a tool that can be presented to the judge as evidence, who will ultimately assess its reliability and probative value in the context of the overall evidence.

REFERENCES

- Abbasi, M. S., Munir, B., Jayaram, V., & Rivas, P. (2025, September 12). Leveraging autocorrelation in a dilated CNN-LSTM framework for predicting the U.S. Supreme Court decisions. *IEEE Access*. IEEE.
- Anastasopoulou, M. (2025). Exploring algorithmic governance: The AI Act and new realities for criminal justice, and fundamental rights. *New Journal of European Criminal Law*, 16(2), 176-196.
- Benjamin, M. S. (2025). THE RISE OF AI IN PROCEDURAL JURISPRUDENCE: GLOBAL INNOVATIONS, LEGAL FRAMEWORKS, AND FUTURE IMPLICATIONS. *Journal of Law and Sustainable Development*, 13(9), e4534.
- Bialek, L. (2026). The Corporate Transparency Act: Substantial Modifications are Desperately Needed (and They Might Just Help the Law Withstand Judicial Scrutiny). *Michigan Business & Entrepreneurial Law Review*, 15(2), 117-138.
- Borgesano, F., De Maio, A., Laghi, P., & Musmanno, R. (2025). Artificial intelligence and justice: a systematic literature review and future research perspectives on Justice 5.0. *European Journal of Innovation Management*, 28(11), 349-385.
- Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. *Social problems*, 68(3), 608-624.
- Cascone, G. (2025). The use of artificial intelligence in EU criminal justice systems: first insights and emerging trends in an evolving landscape. *Open Research Europe*, 5, 361.
- Chemerinsky, A. (2021). Tears of Scrutiny. *Tulsa L. Rev.*, 57, 341.
- Dancy, T., & Zalnieriute, M. (2026). AI and transparency in judicial decision making. *Oxford journal of legal studies*, 46(1), 1-34.
- Dragomir-Constantin, F.-L. (2025). Algorithmic transparency in information systems: A legal necessity for the protection of fundamental rights. *Acta Universitatis Danubius. Juridica*, 21(1), 126-136.

- Ford, J. E., Wing, D., & Hurt, M. (2026). Artificial V. Expert Intelligence: The Risk of Ethical, Constitutional, and Evidentiary Violations in Representing Artificial Intelligence as Linguistic Expertise in the Legal System. *International Journal for the Semiotics of Law- Revue internationale de Sémiotique juridique*, 1-31.
- Garrett, B. L., & Rudin, C. (2023). The right to a glass box: Rethinking the use of artificial intelligence in criminal justice. *Cornell L. Rev.*, 109, 561.
- Grossman, M. R., & Grimm, P. W. (2024). Judicial approaches to acknowledged and unacknowledged AI-generated evidence. *Colum. Sci. & Tech. L. Rev.*, 26, 110.
- Gyau, E. B., Musah, S., Blake, J. J., & Basigi, S. (2026). Promoting crime prevention and criminal justice: Examining the role of artificial intelligence, regulatory quality, and digital infrastructure. *International Journal of Law, Crime and Justice*, 86, 100878.
- Hloviuk, I., & Hryniuk, V. (2025). Some Issues concerning the Right to a Fair Trial under the Statute of the Special Tribunal for the Crime of Aggression against Ukraine. *Law Ukr.: Legal J.*, 18.
- Ibrahim, A., Gourari, F. M., AbdelAziz, G. M., & Hashish, A. (2026). Artificial intelligence in criminal justice governance: Opportunities and legal challenges (A comparative study). *Corporate Law & Governance Review*, 8(2), 77.
- Irfan, M., Bennett, O., & Müller, A. (2025). Synthetic Evidence and Procedural Fairness: Admissibility Standards for AI-Generated Outputs in Civil and Criminal Courts. *International Journal of Law, Policy and Scientific Research*, 1(1), 60-73.
- Kelly, B. (2021). Seeking compatibility in preventing crime with artificial intelligence and ensuring a fair trial. *Masaryk University Journal of Law and Technology*, 15(1), 25-52.
- Kocagül, H., & Kartal, M. (2026). AI, Evidentiary Authority, and the Right to a Fair Trial in Criminal Proceedings. *Laws*, 15(3), 58.
- Kumar, S., Joshi, K., Singh, S., Anandaram, H., Sindhu, K. S., Mahariya, S. K., & Kaushik, V. (2023). AI Technological Interference in Court Proceedings: Right to Fair Trial Decision. 2023 3rd International Conference on Advancement in Electronics & Communication Engineering (AECE),
- Kumari, P. (2025). Legal Frameworks for AI Regulation: A Comparative Study. *Advances in Consumer Research*, 2(2).
- Khan, A. N., Ahmad, I., & Munir, B. (2025, December 31). Data regulations and digital sovereignty: A new challenge to international investment law. *Policy Research Journal*, 3(12), 883-895.
- Lenarczyk, G., Minssen, T., Price, N., & Rai, A. (2025). The future of AI regulation in drug development: a comparative analysis. *Journal of Law and the Biosciences*, 12(2), Isaf028.
- Munir, B. (2025). *Artificial intelligence for lawyers: Navigating novel methods and practices for the future of law*. Punjab Law Book House.
- Maaitah, Q., Alghazali, M. I., & Jasim, A. A. (2026). The Legal Validity and Evidentiary Value of Artificial Intelligence-Generated Digital Evidence. 2026 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSYS),
- Munir, B., Abbasi, M. Z., Wilson, W. B., & Colombo, A., Jr. (2025). Hallucinations in legal practice: A comparative case law analysis. *International Journal of Law, Ethics, and Technology*, 5(2), 85-109.
- Munir, B. (2026). Islamophobic artificial intelligence in the USA: A critical analysis of religious bias in datasets. *The Law Library Journal*, 118(2), 217-237.
- Munir, B. (2025). Evaluating AI in legal operations: A comparative analysis of accuracy, completeness, and hallucinations in ChatGPT-4, Copilot, DeepSeek, Lexis+ AI, and Llama 3. *International Journal of Legal Information*, 53(2), 103-114.

- Munir, B. (2024, December 31). Artificial intelligence and legal decision-making in the USA and Pakistan: A critical appreciation of regulatory frameworks. *Global Foreign Policies Review*, 48-58.
- Meghwal, T. (2025). Admissibility of AI-Generated Forensic Evidence: Legal Standards, Ethical Challenges, and Comparative Jurisprudential Analysis. Available at SSRN 5998017.
- Munir, B., Irfan, M., & Bashir, N. (2024, September 30). Artificial intelligence, data protecting and transparency: A comparative study of GDPR and CCPA. SSRN.
- Munir, B., Khan, A. N., & Ahmed, N. (2020, December 30). An assessment of right to fair trial under the Constitution of Pakistan, 1973: A comparative study of the US and Pakistan. *Global Strategic & Security Studies Review*, 5(4), 31-39.
- Molbæk-Steensig, H., & Quemy, A. (2023). Artificial intelligence and fair trial rights. *Artificial Intelligence and Human Rights*, 265-280.
- Munir, B., Khalid, S., & Noreen, U. (2025). Exposing Islamophobia in machine learning: A critical analysis of the existing theories and biases. *Center for Management Science Research*, 3(3), 1-9.
- Morison, J., & McNerney, T. (2025). When should a computer decide?: Judicial decision-making in the age of automation, algorithms and generative artificial intelligence. In *Research handbook on judging and the judiciary* (pp. 54-87). Edward Elgar Publishing.
- Moroşteş, A. F. (2025). Artificial Intelligence in Forensic Investigations. Doctrinal Gaps, Fundamental Rights and the Rule of Law. *Journal of Legal Studies "Vasile Goldiș"*, 36(50), 164-175.
- Mustafa, H. K., & Ahmed, M. K. (2025). Balancing AI and Judicial Conviction in Criminal Evidence. *Twejer Journal*, 8(4), 23-48.
- Nowotko, P. M. (2021). AI in judicial application of law and the right to a court. *Procedia computer science*, 192, 2220-2228.
- Organization, W. H. (2026). Artificial intelligence and evidence-informed policy: emerging challenges and opportunities: discussion paper.
- Raza, A., Khalid, S., Yasin, A., Kasuri, S. M., & Munir, B. (2023, July 31). Automation in judicial administration: Evaluating the role of artificial intelligence. *International Journal of Contemporary Issues in Social Sciences*, 2(3), 1544-1552.
- Raza, A., Munir, B., Ali, G., Othi, M. A., & Hussain, R. A. (2024). Balancing privacy and technological advancement in AI: A comprehensive analysis of the US perspective. *International Journal of Contemporary Issues in Social Sciences*, 3(3), 3732-3738.
- Sani, A. J. (2025). Artificial Intelligence and the Criminal Justice System: An Analysis of Liabilities and Consequences. *Legal Studies in Digital Age*, 1-14.
- Selçuk, S., Konca, N. K., & Kaya, S. (2025). AI-driven civil litigation: Navigating the right to a fair trial. *Computer Law & Security Review*, 57, 106136.
- Shandilya, S. K., Datta, A., Kartik, Y., & Nagar, A. (2024). Navigating the regulatory landscape. In *Digital resilience: Navigating disruption and safeguarding data privacy* (pp. 127-240). Springer.
- Siemens, W., von Elm, E., Binder, H., Böhringer, D., Eisele-Metzger, A., Gartlehner, G., Hanegraaf, P., Metzendorf, M.-I., Mosselman, J.-J., & Nowak, A. (2025). Opportunities, challenges and risks of using artificial intelligence for evidence synthesis. *BMJ evidence-based medicine*, 30(6), 381-384.
- Sroka, T. (2024). Artificial intelligence and the right to a fair trial. In *Artificial Intelligence and International Human Rights Law* (pp. 250-277). Edward Elgar Publishing.

- Stone, M. R., & Vaidyan, V. (2026). Confidence levels of artificial intelligence generated evidence in cybercrime investigations—a survey: W. Zheng et al. *Security Journal*, 39(1), 35.
- Trang, N. T. T., Linh, N. H., Hoang, N. T. C., Kiet, P. V. T., Loan, L. T. N., & Phuc, N. T. H. (2024). Right to a fair-trial when applying artificial intelligence in criminal justice—lessons and experiences for Vietnam. *Journal of Law and Sustainable Development*, 12(3), 34.
- Ulenaers, J. (2020). The impact of artificial intelligence on the right to a fair trial: towards a robot judge? *Asian Journal of Law and Economics*, 11(2).
- Vidaki, A. N., & Papakonstantinou, V. (2025). Democratic legitimacy of AI in judicial decision-making. *AI & SOCIETY*, 40(8), 6025-6035.
- Zaidan, E., & Ibrahim, I. A. (2024). AI governance in a complex and rapidly changing regulatory landscape: A global perspective. *Humanities and Social Sciences Communications*, 11(1).
- Zhang, H. (2025). The evidentiary value of AI-generated data: A framework for reliability and admissibility. *Education and Social Work*, 2(2), 34-41.

