

UNITED STATES OF AMERICA

Tools	Axon Draft One Casetext CoCounsel ChatGPT Claude Conferbot Correctional Offender Management Profiling for Alternative Sanctions Cybercheck Gemini Geolitica Google Vertex AI JusticeText LegalServer Level of Service Case Management Inventory Level of Service Inventory - Revised Lexis+AI MateyAI NICE Justice PATTERN Pretrial Risk Assessment PROSECUTORbyKarpel Reduct.Video Relativity SentencingStates ShotSpotter SPARC-13 Specialised systems Westlaw AI
Tasks	Administrative support Case management Data review and analysis Evidence review and analysis Legal research, analysis and drafting support Operational support Predictive analytics Risk-assessment
Users	Law enforcement Prosecutors Courts Defence
Scope	Nationwide
Training	No systematic or widespread training
Regulation	No dedicated federal AI regulation. The American Bar Association (ABA) Standing Committee on Ethics and Professional Responsibility has issued opinions on the use of AI, courts have issued standing orders, and there are state laws, and state court rules and guidelines regulating the use of AI in court
Cases	There are many cases in which attorneys are sanctioned for reliance on unverified authorities hallucinated by generative AI tools. Common sanctions include fines, costs orders, the striking of submissions, removal from the case, and in some instances, suspension from practice
Insights	In an audit by the Los Angeles Police Department of its predictive policing system (Operation LASER) for identifying crime hot spots and “chronic offenders”, it was revealed that almost 50% of the “chronic offenders” had zero or one arrest for violent crime, and almost 10% had no “quality interactions” with police

AT A GLANCE

The United States has adopted AI in a piecemeal manner, with more extensive use in law enforcement and in courts on a state-by-state basis. Law enforcement uses tools like Geolítica for predictive policing (with a <1% success rate in Plainfield, NJ), Cybercheck for digital forensics, Axon's Draft One to auto-draft police reports, and ShotSpotter gunshot detection (decommissioned in Chicago after false positives). Prosecutors rely on PROSECUTORbyKarpel for case management, NICE Justice for evidence review, and in one Arizona case (2025), an AI avatar delivered a deceased victim's impact statement. Courts employ risk-assessment tools such as COMPAS, VPRAI, and PSA at pre-trial and sentencing, while Miami-Dade courts use the chatbot "SANDI" to guide litigants. Defence attorneys adopt SentencingStats to argue for reduced sentences, JusticeText to review bodycam footage, and Casetext CoCounsel for legal research. There is no systematic or widespread training available to judges, prosecutors, or defence counsel.

There is **no federal AI regulation and no federal framework to regulate the use of AI in court**. The ABA has issued opinions on the use of AI tools by counsel. Courts have issued standing orders and sanctioned lawyers for reliance on unverified authorities generated by ChatGPT and similar tools. There is a growing body of state rules and guidelines regulating the use of AI in court.

USE

There are 50 states in the US, each with its own laws, judicial systems, and approaches to criminal justice. The AI Justice Atlas does not attempt to document developments in every jurisdiction; rather, it offers a high-level overview of how AI is being integrated into criminal proceedings, highlighting emerging trends, significant initiatives, and patterns shaping the broader legal landscape.

LAW ENFORCEMENT

Operational support

AI-powered **dispatch systems** and **multilingual phone translations** help law enforcement prioritise calls and improve response times.

Palantir and Anthropic announced in November 2024 that they will be partnering with Amazon Web Services to make Anthropic's **Claude** models available to US intelligence and defence agencies. The companies said that access to the Claude models from within Palantir's data analytics platform will help agencies with tasks such as: processing vast amounts of complex data rapidly, elevating data-driven insights, identifying patterns and trends more effectively, streamlining document review and preparation, and helping US officials to make more informed decisions in time-sensitive situations while preserving their decision-making authorities.

Predictive analytics

In the US, approximately **38.2% of major US police departments are using or piloting predictive policing systems**, which ingest large volumes of historical crime data to **forecast high-crime 'hot spots'** and guide patrol allocation. These tools do not authorise arrests: they serve as human-reviewed leads. For example, **PredPol** (now Geolitica) uses historical crime incident data (such as location, time, type) to forecast 'hotspot boxes': small geographic grid cells where crime is likely to occur during a given temporal window (e.g. that day or shift), allowing police to allocate more patrol presence or preventative resources to those boxes. In one publicly released dataset, Geolitica generated 23,631 predictions for Plainfield, New Jersey between 25 February and 18 December 2018. Out of those, fewer than 100 predictions matched an actual crime of the predicted type in the predicted box during the time window: the tool had a success rate of less than 0.5%.

Prison Classification Systems apply risk-assessment tools to inform decisions about facility type, housing unit assignments, placement in general or special populations, and availability of programmes and services for incarcerated individuals. The Federal Bureau of Prisons uses the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN) to track dynamic changes in risk and uniform earned time eligibility, along with Standardised Prisoner Assessment

for Reduction in Criminality (13 domains) (SPARC-13) to identify programmatic and treatment needs. Most classification systems rely on risk-assessment tools originally designed to estimate post-release recidivism, though some have been modified for predicting prison misconduct. Correctional staff review risk assessments within established policies and procedures governing classification decisions. Deployment occurs as standard practice across federal and state prison systems, with tools integrated into regular classification and reclassification processes.

Chicago's Police Department previously used two predictive analytics systems: the '**Strategic Subject List**' (SSL) and the '**Crime and Victimization Risk Model**' (CVRM), which were designed to predict the likelihood that an individual would become a 'party to violence' (PTV)—that is, the victim or offender in a shooting. The attributes used by the models to generate risk scores and tiers were:

1. incidents as a victim of shooting;
2. age at latest arrest;
3. incidents as a victim of aggravated battery or assault;
4. trend in involvement in crime incidents;
5. arrests for unauthorised use of a weapon;
6. violent incidents as an arrestee;
7. narcotics arrests; and
8. gang affiliation.

The results of SSL were known as 'risk scores', while CVRM produced 'risk tiers', with higher scores or tiers indicating a greater risk of becoming PTV. Every individual arrested at least once within a four-year period prior to the model being built—regardless of whether they had a history of violence—received a risk score or tier. The Police Department decommissioned its PTV risk model programme on 1 November 2019 due to (among other reasons) unreliable risk scores and insufficient training.

The Los Angeles Police Department (LAPD) **Operation LASER** (Los Angeles Strategic Extraction and Restoration) was audited by the LAPD in early 2019. Operation LASER was a predictive

policing and 'chronic offender' targeting programme. It used, for example, historical crime data, field interviews, gang membership, and arrest records to identify 'hot spots' and 'chronic offenders' and to assign risk scores. The LAPD shut down LASER after the audit.

Closer look

The audit of Operation LASER by the LAPD in 2019 revealed significant inconsistencies in how individuals were selected and retained. Almost half of the 'chronic offenders' had zero or one arrest for violent crime, and almost 10% had no 'quality interactions' with police.

Data review and analysis

Cybercheck is an AI forensic tool, founded in 2016, that issues probabilistic reports to aid in suspect identification and crime scene analysis. The tool uses machine learning algorithms to analyse open-source intelligence data and link an individual's 'cyberDNA' or digital signature, to a crime scene. In an Ohio homicide trial, Cybercheck's founder testified that the tool's conclusions were 98.2% accurate, but provided no source for this value. The tool has been used in nearly 8,000 cases across 40 states and nearly 300 agencies, despite heavy criticism of its methodology.

Axon's Draft One is a new software product that drafts high-quality police report narratives in seconds based on auto-transcribed body-worn camera audio. Officers reportedly spend up to 40% of their time writing police reports and Draft One allegedly cuts that time in half. The system takes an officer-in-the-loop approach where Draft One generates report narratives based on body-worn camera audio, but these narratives cannot be submitted without officer review, editing and approval. Police have adopted Axon's Draft One in Lafayette, Indiana; Tampa, Florida; and Campbell, California.

Facial recognition technology (FRT) employs computer vision algorithms that detect faces in images, extract quantitative templates, and compare similar scores between facial features. Federal agencies such as the FBI use FRT systems to identify perpetrators, victims, and witnesses as part of authorised criminal investigations. The FBI's NGI Interstate Photo System incorporates data from 17 state agencies and two federal agencies, encompassing over 67 million arrest photos. Trained examiners independent of case teams must manually review FRT results, and agency policies prohibit using FRT results as sole proof of identity. Deployment varies significantly across jurisdictions, with some agencies prohibiting FRT use entirely while others permit broad application under different policy frameworks.

Automated fingerprint identification systems (AFIS) are statistical or algorithmic tools that compare friction ridge patterns and minutiae features. These systems perform automated matching for ten-prints with minimal human oversight except when evidence will be used in prosecution. Latent print analysis typically requires human examiner review. Deployment of this technology is national through federal systems, as well as state and local AFIS networks.

Agencies also use **automated licence plate recognition systems (ALPR)**, which are computer vision systems that capture and cross-reference licence plate data against law enforcement databases. Such systems have widespread use: almost all local law enforcement agencies have an LPR programme, as do many smaller agencies. Deployment occurs primarily at local levels with some agencies subscribing to commercial LPR services that operate networks of participating cameras.

Drug Classification Systems employ machine learning models that analyse chemical composition data to classify the geographic origin of seized drug samples, particularly heroin and cocaine. The Drug Enforcement Administration (DEA) uses these systems to detect anomalies in analysis and low-confidence results, providing intelligence about drug trafficking patterns and trends. Analysts review machine learning outputs for intelligence and investigative purposes, though these results are not currently used as evidence in court proceedings. Deployment remains

limited to specialised federal forensic applications, primarily within DEA operations for understanding drug trafficking networks.

Gunshot-detection systems like **ShotSpotter** deploy acoustic sensors across urban areas, listening for impulsive sounds that may be gunfire. The system uses algorithms and human verification to filter sounds, triangulate the approximate location, and generate alerts (often within seconds) to law enforcement dispatch centres or ‘real-time crime cent[re]s’.

In New York, the NYPD is launching a ‘**Drone as First Responder**’ programme that links ShotSpotter alerts to drone deployment: when a shooting is detected, a drone (piloted from a centralised operations centre) is sent to fly over the site and stream live video and telemetry back to officers en route. By contrast, the City of Chicago announced in February 2024 that it decommissioned the technology. It has been reported that the system has high rates of false positives: for example, a 2024 audit claims that roughly 87% of ShotSpotter alerts did not correspond to confirmed shootings.

PROSECUTORS

Case management

PROSECUTORbyKarpel’ (PbK) is the most widely used prosecutor case-management system in the US. In August 2024, Karpel Solutions and NiCE announced a technology partnership to allow offices to pull digital evidence and AI features directly into case files. For example, users can manage documents, generate subpoenas, witness documents, or victim letters, and use eDiscovery or redaction tools. PbK claims to havestreamlined over 600 large and small prosecutors’ offices.

‘**Generative-AI avatars**’ (for scripted victim-impact videos) have been used to deliver emotional narratives in court. Some courts have issued rules, guidance or standing orders requiring disclosure of AI usage in documents or evidence, including when the AI-generated output is made available to the public (see below). **One Arizona judge in May 2025 allowed a dead man**

to deliver his own victim statement via an AI avatar. The man was shot and killed and, when it was time for the killer to be sentenced, the victim impact statement was rendered by an AI-generated avatar that used the deceased's face and voice. Though this marked the first time that a US court allowed an AI-generated victim to make this kind of statement, the judge in that case noted the AI nature of the impact statement, and the defendant's attorney noted the likelihood that the appeals court would weigh whether the judge improperly relied on the AI video when handing down the sentence. (As at September 2025, the judgment on appeal had not been delivered.) As at September 2025, this type of practice was still in limited pilot use.

Legal research, analysis and drafting support

Generative AI tools such as ChatGPT, Claude, Gemini, and Westlaw AI are used by US prosecutors for summarising precedents, drafting motions, and generating legal research outlines. All outputs must be verified, and all hallucinations corrected by prosecutors, and compliance with professional ethics must be upheld (see below). These tools are broadly available within legal tech firms and some law offices, with early adoption in the criminal setting as at September 2025.

Evidence review and analysis

Prosecutor offices in the US have used AI systems for evidence review and analysis. For example, **NiCE Justice** uses AI-automation for audio and video transcription and translation, optical character recognition software (for extracting text from images or scanned documents), object detection, analytics, and evidence 'connection-finding'. This system has been used by, among others, the Calcasieu Parish (LA) District Attorney (since December 2024), and the Monterey County (CA) District Attorney (since December 2024).

COURTS

Case management

Courts are using AI-enhanced reminder systems, such as Conferbot, which automate notifications to reduce failure-to-appear rates.

Courts are also exploring AI for public-facing tools and access to justice. For example, **chatbots** or **virtual assistants** are used to help litigants understand procedural steps, check deadlines, or route them to forms. In Miami-Dade courts, ‘**SANDI**’ is used to assist court users by helping them with case status, court forms, directions, and procedural questions.

Risk-assessment

US courts use AI-based risk-assessment tools in the pre-trial and sentencing stages of criminal litigation.

Pre-trial risk-assessment

Risk-assessment tools employ statistical models to estimate the likelihood that defendants will fail to appear in court, commit new offences before trial, or pose public safety risks during pre-trial release.

The federal court system uses the PreTrial Risk Assessment (PTRA), which is an algorithmic tool developed by the Administrative Office of the US courts. State and local jurisdictions deploy various tools including the Virginia Pretrial Risk Assessment Instrument (VPRAI), Public Safety Assessment (PSA), and Ohio Risk Assessment System Pretrial Assessment Tool (OPRAS-PAT).

Human oversight on these systems ensures that risk assessments inform judicial decision-making about pre-trial release conditions but cannot determine or replace judicial discretion in these decisions.

Deployment spans all 50 states, with every state implementing some form of risk assessment for pre-trial decisions, though with local variation in specific tools and implementation approaches.

Sentencing risk-assessment

These tools use statistical models that predict recidivism likelihood to assist judges in making sentencing decisions within applicable statutory ranges and guidelines. Common instruments include:

- Level of Service Inventory - Revised (LSI-R): a structured actuarial risk/need assessment instrument designed to estimate an offender's likelihood of recidivism (reoffending) and to identify criminogenic needs (dynamic factors that can be changed through intervention).
- Level of Service Case Management Inventory (LS/CMI): a combined risk/need assessment and case management instrument for adult offenders. This system is not only designed to estimate recidivism risk, but also to guide supervision, intervention, and management plans.
- Correctional Offender Management Profiling for Alternative Sanctions (COMPAS): examines recidivism risk and failure to appear. The tool is widely used in states including California, Wisconsin, and Florida.

Deployment of risk-assessment tools for sentencing occurs widely across state and federal court systems, though with considerable variation in which specific tools jurisdictions choose to implement.

Judges may consult these scores in connection with decisions concerning bail, sentencing, or probation, although defendants often cannot examine the algorithmic logic. Studies indicate that judges using such AI tools gave shorter sentences to low-risk defendants, but racial disparities persisted even where the algorithm purported to be 'objective'. Risk-



Oxford Institute
of Technology and Justice

assessment tools support judicial decision-making but cannot replace or determine sentencing decisions, which remain within judicial discretion.

DEFENCE

Administrative support

LegalServer is a case management system widely used by public defenders, integrating AI for document automation, optical-character-recognition-based (OCR-based, used to turn documents, images or handwritten/typed evidence into machine-readable text), and legal analytics.

Legal research, analysis and drafting support

As with prosecutors (discussed above), **generative AI tools** such as ChatGPT, Claude, Google Vertex AI, Gemini, and Westlaw AI are used by defence attorneys in the US for summarising precedents, drafting motions, and generating legal research outlines. All outputs must be verified by defence attorneys, hallucinations corrected, and compliance with professional ethics must be upheld (see below). These tools are broadly available within legal tech firms and some law offices, with early adoption in the criminal setting as at September 2025.

'SentencingStats' is a machine-learning platform that analyses federal sentencing data and generates statistical reports on likely sentencing outcomes based on historical trends. Federal defenders have begun using SentencingStats to support sentencing advocacy and mitigation efforts. By analysing trends in judicial decisions, attorneys can present data-backed arguments for reduced sentences, demonstrating disparities or inconsistencies in sentencing patterns.

'Casetext CoCounsel' (Thomson Reuters) assists defence attorneys by conducting legal research, drafting motions, summarising discovery, and preparing deposition questions. The Miami-Dade County (Florida) Public Defender's Office was the first public defender office in the US to integrate CoCounsel, providing 100 attorneys access to the AI research assistant. The tool significantly reduced legal research time and improved efficiency in drafting motions and trial preparation. Attorneys noted that CoCounsel helped generate cross-examination questions and

identify overlooked case law. However, due to budget constraints, not all defenders received licenses.

'Westlaw Edge' and 'Lexis+AI' are used to allow attorneys to quickly identify relevant caselaw, analyse opposing briefs, and draft legal arguments. Larger criminal defence firms and well-funded attorneys have integrated these tools to accelerate legal research and ensure comprehensive case preparation. AI-powered brief analysis features have been particularly useful for identifying missing citations and legal precedents.

Evidence review and analysis

Public defenders increasingly use AI-driven tools such as **JusticeText** and **Reduct.Video** for AI-assisted speech-to-text and indexing evidence. Examples of tasks performed by these tools include transcribing police body-cam footage, or 911 (emergency) calls, and flagging key moments (such as a suspect's request for counsel or being read their Miranda rights). Prosecutors manually review flagged transcripts and internal protocols to ensure there is no over-reliance, limiting the use of these tools to supportive functions.

JusticeText

Kentucky Department of Public Advocacy (DPA) implemented JusticeText to handle the surge of bodycam footage in cases. Kentucky DPA defenders reported that JusticeText reduced time spent reviewing evidence by hours per case, allowing them to find key contradictions in police statements more efficiently. Estimates suggest that a single police officer's body camera will record around 32 files, seven hours and 20GB of video per month. JusticeText is reported to generate transcripts that deliver 50% in time savings. However, funding limitations meant that not all defenders statewide had access to the tool.

Santa Cruz County Public Defender (CA) adopted JusticeText to streamline discovery review, ensuring that attorneys could quickly locate and analyse critical evidence.

The **Harris County** defender's office in Houston also uses JusticeText.

The **Virginia Indigent Defense Commission** signed a contract with JusticeText after a 2021 pilot project involving more than 100 attorneys, investigators, and support staff.

In Nebraska, the **Sarpy County Public Defender's Office** adopted JusticeText.

Reduct.Video

The **Colorado State Public Defender** deployed Reduct to transcribe bodycam footage and police interrogations, making video evidence review more efficient. Attorneys could highlight key clips, generate captions, and create court exhibits. Reduct reduced the time required for evidence processing and improved courtroom presentation of video evidence.

'**Relativity**' offers a broad, end-to-end e-discovery and investigation platform designed to manage diverse data types (such as text, email, chat, multimedia) across the entire litigation, from data collection and processing through review, analysis, and production. AI capabilities are integrated to enhance these core workflow stages.

'**MateyAI**' is an AI tool that organises and analyses criminal eDiscovery, built specifically for criminal defence.

TRAINING

Systematic and widespread training is not available to judges, prosecutors, or defence counsel to help them understand how AI tools work and their limitations. There are, however, ad hoc training sessions and seminars.

REGULATION

At the federal level, **neither the Department of Justice nor the Administrative Office of the Courts has established comprehensive policies governing the use of AI in the courts.** As at August 2025, there were also no state statutes that explicitly regulate the use of AI in criminal proceedings or courtroom decision-making. What exists instead are ABA opinions on the use of AI tools, standing orders issued by courts, and the enforcement of sanctions, and state-specific judicial policies, administrative rules, and guidance that govern AI use by judges, court staff, and attorneys.

GUIDELINES FOR PRACTITIONERS

ABA Model Rules and Formal Opinions 512

In July 2024, the American Bar Association (ABA) Standing Committee on Ethics and Professional Responsibility published *Formal Opinion 512 - Generative Artificial Intelligence Tools*, which establishes the federal ethical rulebook for lawyers in the use of generative AI, construing new obligations under the existing *ABA Model Rules of Professional Conduct*.

ABA Model Rules:	ABA Opinion 512:
Competency:	To competently use a generative AI tool in client representation, 'lawyers must have a reasonable understanding of the capabilities and limitations' of the specific technology used. Lawyers must



Oxford Institute
of Technology and Justice

Model Rule 1.1 requires lawyers to provide competent representation to clients.

Confidentiality:

Model Rule 1.6 requires lawyers to keep all client information confidential subject to limited exceptions. Model Rules 1.9 and 1.18 apply similar protections to former and prospective clients.

Communication:

Model Rule 1.4 addresses lawyers' duty to communicate with their clients.

Meritorious claims and candour towards the tribunal:

Model Rules 3.1, 3.3 and 8.4(c) prohibit frivolous claims, false statements, and conduct involving dishonesty, fraud, deceit or misrepresentation

Supervisory responsibilities:

independently verify and review the generative AI tool's output.

Before inputting information into generative AI tools, lawyers must 'evaluate the risks that the information will be disclosed or accessed by others outside the firm'. Informed client consent is required when client information is inputted into a generative AI tool.

The 'facts of each case will determine' whether a lawyer is required to disclose the use of generative AI or obtain a client's informed consent. Client consultation about the use of a generative AI tool is 'necessary when its output will influence a significant decision', such as when a lawyer relies on it 'to evaluate potential litigation outcomes or jury selection'.

Lawyers must carefully review outputs from generative AI tools 'to ensure that the assertions made to the court are not false' and 'to correct errors', including 'citations to nonexistent opinions' and 'misleading arguments'.

Managerial lawyers 'must establish clear policies' on the law firm's permissible use of generative AI,



Oxford Institute
of Technology and Justice

Model Rules 5.1 and 5.3 address the ethical duties of lawyers charged with managerial responsibility concerning their firm and subordinate lawyers and non-lawyers.

Fees:

Model Rule 1.5 governs lawyers' fees and expenses.

ensure that subordinate lawyers and non-lawyers receive relevant training, 'make reasonable efforts' to ensure that nonlawyers conform with the lawyers' professional obligations, and ensure that AI tools are 'configured to preserve confidentiality and security of information'.

Lawyers may bill only for 'their time actually worked' even if a generative AI enables them to complete tasks more quickly. To the extent that a particular AI tool 'functions similarly to equipping and maintaining a legal practice', a lawyer 'should consider its cost to be overhead' and not charge the client for it absent contrary advance disclosure. A lawyer 'may not charge a client to learn about how to use' generative AI that the lawyer will use regularly 'because lawyers must maintain competence in the tools they use', including generative AI technology.

ABA Formal Opinion 517

In July 2025, the ABA made a further reference to generative AI in *Formal Opinion 517 - Discrimination in the Jury Selection Process* concerning the prohibition against discrimination in Model Rule 8.4 (g). Because AI-assisted juror selection programmes can unknowingly apply 'rankings in a manner that would constitute unlawful discrimination (e.g. based on the prospective jurors' race or gender)', lawyers should 'conduct sufficient due diligence to acquire a general understanding of the methodology employed by the juror selection program'.

Standing orders in federal courts

A number of judges throughout the United States have issued standing orders governing the use of AI by attorneys who appear before them. Examples include the US District Court for the Northern District of California, Order Nos. 23-0903 (Judge Araceli Martinez-Olguin), which requires certification that lead trial counsel has verified the accuracy of AI-generated content, and 23-0933 (Magistrate Judge Peter H. Kang), which requires disclosure of AI usage in documents, identification of AI-generated evidence, and adherence to confidentiality. Other federal courts with orders concerning the use of artificial intelligence include: the District of Hawaii, the Northern District of Illinois, the Eastern District of Missouri, the District of New Jersey, the Southern District of New York, the Southern and Northern Districts of Ohio, the Western District of Oklahoma Bankruptcy Court, the Northern District of Texas Bankruptcy Court, the Northern District of California, the District of Colorado, and the Western District of North Carolina.

Sanctions:

The courts have **sanctioned lawyers relying on unverified authorities created by ChatGPT or similar tools**. Federal courts can look to *Rule 11 of the Federal Rules of Civil Procedure*, which permits the imposition of ‘an appropriate sanction on any attorney, law firm, or party’ that violates the rule requiring them to certify that their contentions are warranted by existing law and the factual contentions will have evidentiary support.

Fines are a common sanction as are costs orders requiring payment of the opposing party’s legal fees incurred when responding to court filings containing hallucinated case law. Courts have also struck out filings, removed lawyers from the case or suspended them from practice. See ‘Cases’ below for some examples of the sanctions imposed in federal and state courts.

State rules and guidance

A number of states have adopted rules governing the use of AI in the courts. Examples include Delaware's interim policy, Illinois' progressive framework, and California's comprehensive set of rules.

Delaware: In October 2024, the Delaware Supreme Court adopted an *Interim Policy on the Use of GenAI by Judicial Officers and Court Personnel*, which allowed limited AI use but required administrative approval of AI tools and prohibited delegating decision-making responsibilities to AI.

Illinois: In December 2024, the Illinois Supreme Court announced its *Policy on Artificial Intelligence*, which states 'The use of AI by litigants, attorneys, judges, judicial clerks, research attorneys, and court staff providing similar support may be expected, should not be discouraged, and is authorized provided it complies with legal and ethical standards. Disclosure of AI use should not be required in a pleading'.

California: In July 2025, the California Judicial Council, the policy-making body of the Californian courts, adopted a regulatory framework requiring courts that permit AI use to adopt policies by 15 December 2025, which must prohibit entry of confidential information into public AI systems and must require disclosure when AI-generated content is provided to the public.

Other state courts and state bar associations have also issued rules or guidance concerning the use of AI in courts. They include Florida, Michigan, New Jersey, and Pennsylvania. Additional states are likely to follow suit. For example, in July 2025, Georgia's Judicial Council AI Committee released recommendations, which included 'establishing interim and eventually permanent policies governing the use of AI in Georgia's judicial system'.

CRIMINAL PROCEDURE RULES

Federal and state rules of procedure and evidence may also apply to the use of AI in court even if AI is not mentioned. For example, *Federal Rules of Evidence* 901(b)(1) and 901(b)(9) concern witness testimony and evidence describing a process or system that produces accurate results.

They govern authentication of AI evidence, requiring witness testimony and evidence describing processes that produce accurate results. In November 2024, the US Courts Advisory Committee proposed expanding Rule 901(b)(9) to require proponents of AI-generated outputs to produce evidence that outputs are 'reliable' and describe training data and software.

DATA PROTECTION LEGISLATION

Privacy laws, such as the *California Consumer Privacy Act* at the state level, may also be relevant to the use of AI in criminal proceedings and the processing of any personal data by AI tools.

HUMAN RIGHTS

Constitutional due process standards could also be construed to regulate AI use. For instance, mass surveillance and predictive policing powered by AI test the protection against unreasonable searches enshrined in the Fourth Amendment to the *US Constitution*, while opaque, AI-driven risk assessments threaten the guarantees of equal protection and due process provided in the Fourteenth Amendment. International human rights instruments ratified by the United States may also be relevant, including fair trial and privacy guarantees in articles 14 and 17 of the *International Covenant on Civil and Political Rights*.

OUTLOOK

In January 2025, President Trump issued *Executive Order 14179* entitled '*Removing Barriers to American Leadership in Artificial Intelligence*' to establish the Trump Administration's approach to AI policy with the stated goal of maintaining US leadership in AI by developing systems 'free from ideological bias or engineered social agendas' and removing barriers to American AI innovation. The order sets a policy to sustain and enhance America's global AI dominance to 'promote human flourishing, economic competitiveness, and national security'. President Trump also revoked President Biden's 2023 '*Executive Order on the Safe, Secure and Trustworthy Development and Use of Artificial Intelligence*' (Executive Order 14110), which had mandated transparency and agency bias audits, and required a review of policies, directives, and regulations issued under the revoked Executive Order 14110.

The action plan issued in July 2025 pursuant to the new *Executive Order 14179* also emphasised deregulation, stating that AI-related federal funding should not go to states ‘with burdensome AI regulations that waste these funds’ while acknowledging that the federal government ‘should also not interfere with states’ rights to pass prudent laws that are not unduly restrictive to innovation’. Insofar as the legal system was concerned, the action plan referred to AI-generated media, such as deepfakes, which could be used as ‘fake evidence’ to deny justice to the parties to litigation. The plan suggested that the federal administration give ‘the courts and law enforcement the tools they need to overcome these new challenges’, including by exploring ‘deepfake-related additions’ to the Rules of Evidence.

In the absence of comprehensive federal legislation, some states have enacted general AI laws with many taking effect in 2026. These include, for example, the *Colorado AI Act*, which was enacted in May 2024 and regulates developers and deployers of ‘high-risk’ AI systems involved in ‘consequential decisions’ including in legal services, with a particular focus on preventing bias and discrimination. Other examples include additional AI legislation passed in California in September 2024, including its *Generative AI Training Data Transparency Act*, which requires developers to publish summaries of datasets used in training.

CASES

There is **growing jurisprudence on the use of AI tools in court**. Set out below are a selection of examples focussing on the risk of bias and discrimination posed by AI tools used by law enforcement and on the sanctioning of lawyers for relying on unchecked hallucinations generated by ChatGPT and similar tools.

CASES CONCERNING AI TOOLS USED BY THE COURTS AND IN LAW ENFORCEMENT

State v. Loomis (Wisconsin S. C., 2016): In this case, Eric Loomis pleaded guilty to attempting to flee a traffic officer and operating a motor vehicle without the owner's consent. His pre-sentence investigation report included a **COMPAS risk assessment** (see above) that indicated that he presented a high risk of recidivism. The circuit court referenced the COMPAS risk score along with other sentencing factors in ruling out probation. Loomis argued that this violated his due process rights because: first, the proprietary nature of COMPAS prevented him from challenging its scientific validity; second, it violated his right to an individualised sentence by relying on group data; and finally, it improperly considered gender in sentencing. The Wisconsin Supreme Court held that if used properly with specific limitations and cautions, consideration of a COMPAS risk assessment at sentencing does not violate due process. But the court established strict limitations: risk scores may not be used 'to determine whether an offender is incarcerated', 'to determine the severity of the sentence', or as the determinative factor in deciding whether an offender can be supervised safely in the community. The court also required that any pre-sentence investigation report containing COMPAS include written advisement about the tool's limitations.

Williams v. City of Detroit (E.D. Mich. filed 2021): Robert Williams, a Black man, was wrongfully arrested by Detroit police in January 2020 after **facial recognition technology** incorrectly identified him as a shoplifter, making this the first publicly reported instance of a false face-recognition match leading to wrongful arrest in the United States. The case resulted in a groundbreaking settlement in June 2024 requiring the Detroit Police Department to implement strong policies constraining law enforcement's use of facial recognition technology, including prohibiting arrests based solely on facial recognition results.

State of Ohio v. Black (C.A. 9th J.D. 2024): One high-profile case includes Adarus Black, who was sentenced to life imprisonment, based predominantly on **Cybercheck data** (see above). Defence attorneys argued that jurors would not have convicted Adarus Black without this AI

evidence. While Adarus Black's conviction was affirmed, this sparked investigations into the tool and its founder, and led to exclusion or withdrawal of AI-based evidence in several cases.

CASES CONCERNING HALLUCINATIONS BY GENERATIVE AI TOOLS

Lawyers relying on generative AI tools

Lawyers are being increasingly sanctioned for their reliance on unverified and false outputs from generative AI tools, often by way of **payment of a fine**.

An early civil case was *Mata v. Avianca* (S.D.N.Y. 2023), with judgment imposing sanctions delivered in June 2023. Lawyers filed submissions containing 'non-existent judicial opinions with fake quotes and citations created by the artificial intelligence tool ChatGPT'. This error was exacerbated by their failure to verify submissions, and their continued defence of the fake material even under judicial scrutiny. During the court hearings, one of the lawyers admitted misunderstanding ChatGPT's capabilities, stating: 'I just never thought it could be made up.' Other issues in this case included subjective bad faith on the part of the lawyers, 'acts of conscious avoidance and false and misleading statements to the Court', and violations of procedural rules. These breaches led to a joint sanction of a fine of USD 5,000 against the lawyers and their firm.

Fines continue to be imposed, sometimes in combination with other sanctions. More recently, in *Lacey v. State Farm General Insurance Company* (C.D. California 2025), two law firms representing the plaintiff were held jointly responsible for 'bogus AI-generated research' contained in a brief, which they failed to properly correct before re-submission despite explicit notice of the issues. The court regarded their conduct as 'tantamount to bad faith', imposing litigation sanctions including striking the plaintiff's supplemental briefs, and financial payments totalling USD 31,100 to compensate the defence. And in *Gauthier v. Goodyear Tire & Rubber Co* (E.D. Tex. 2024), a lawyer used Anthropic's Claude AI to produce a filing that used several non-existent quotations from real cases, and cited two cases that did not exist at all. He tried but

failed to verify them through another legal AI tool, later compounding the error by not correcting his brief even after opposing counsel flagged the issues. The Western District of Texas imposed a USD 2,000 fine and also ordered the lawyer to complete a training course on generative AI.

Further sanctions include **dismissal of the filing, striking lawyers from the case, or suspending them from practice**. In *Bevins v Colgate-Palmolive Co* (E.D. Penn. 2025), a lawyer filed briefs containing fake cases. When questioned, he offered no explanation beyond asserting that they may have been the ‘result and consequence of a tired, rather than fresh eyed, last proof reading of the filing’. The court found this unpersuasive, notified the relevant state and federal bars, struck the attorney’s appearance, required the attorney to inform his client about the sanction, and required the client to find new counsel if she wanted to refile after dismissal. In the case of *In re Newsom* (M.D. Florida 2024), a lawyer was suspended for one year on the recommendation of the Grievance Committee of the Middle District of Florida, after admitting that he ‘may have used artificial intelligence to draft the filing(s) but was not able to check the excerpts and citations’.

Reliance on generative AI tools has also featured in **criminal cases**. In *United States v. Michel* (D.C. 2024), rapper Pras Michel’s lawyers asked an AI tool to write ‘a powerful, emotionally compelling closing argument’ for his trial. Using the AI tool, Michel’s attorney’s trial closing erroneously attributed another artist’s lyrics to Pras Michel’s group. The court observed that Michel had not explained how the mistaken attribution of a song in the closing argument ‘resulted in prejudice’, and for this and other reasons denied his claim for ineffective assistance of counsel.

In *J.G. v. New York City Department of Education* (S.D.N.Y. 2024), the Cuddy Law Firm sought to justify its fees in multiple cases by relying on ChatGPT’s suggestions about lawyers’ rates. The U.S. District Court judge dismissed the arguments as ‘utterly and unusually unpersuasive’ because ‘treating ChatGPT’s conclusions as a useful gauge of the reasonable billing rate for the work of a lawyer with a particular background carrying out a bespoke assignment for a client in a niche practice area was misbegotten at the jump’.

Self-represented litigants relying on generative AI tools

Courts are also taking a stricter stance with self-represented litigants. In *Ferris v. Amazon.com Services LLC* (N.D. Mississippi 2025), William Ferris, a self-represented litigant, used GenAI to prepare filings containing fake and misleading case citations, and also for an oral statement to the court at a show-cause hearing. The court issued the following rebuke:

Courts exist to decide controversies fairly, in accordance with the law. This function is undermined when litigants using AI persistently misrepresent the law to the courts. AI is a powerful tool, that when used prudently, provides immense benefits. When used carelessly, it produces frustratingly realistic legal fiction that takes inordinately longer to respond to than to create. While one party can create a fake legal brief at the click of a button, the opposing party and court must parse through the case names, citations, and points of law to determine which parts, if any, are true. As AI continues to proliferate, this creation-response imbalance places significant strain on the judicial system.

William Ferris was ordered to pay the costs Amazon incurred ‘as a reasonable result of Plaintiff’s false citations’.

