V.
Machine Learning and the Law

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

AI in the Judiciary
Practice and Policy

Helena Cecilie Lyng-Olsen & Daniel E. Ho

While numerous efforts have developed and codified AI use case inventories for government in the executive branch, no such inventory exists for the judicial branch. Much focus has been on attorney utilization of AI, but no work has examined the policies courts have instituted to regulate AI use by court employees and judges themselves. We fill this gap in two ways. First, we provide an overview of publicly known AI use cases in the courts, categorizing them by risk and potential impact on judicial discretion (Table 1).[1] *Second, we compile and analyze the internal AI use policies of the 17 U.S. state courts that have promulgated such policies (Figure 1b).*[2] *Our paper proceeds as follows. Section A discusses the methodology for systematically canvassing AI use cases in the courts and state policies. Section B presents an overview of these use cases. Section C analyzes state court policies.*

A. Methodology

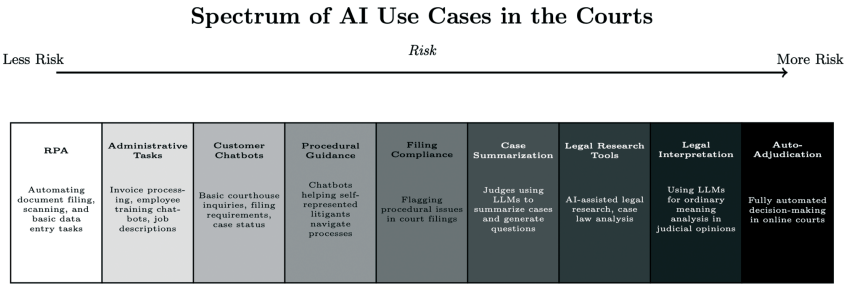
To identify AI use cases in the judiciary, we conducted an extensive manual search of all publicly available sources (performed from February to May 2025), detailed in Appendix B.[3] First, to identify use cases, we used keywords to search (1) judicial opinions on WestLaw, Lexis, and CourtListener, (2) when available, state judicial council websites, (3) publications by the six ABA Judicial Division Conferences and the Judicial Conference of the United States, and (4) general search engines. The most common sources providing information about use cases came from conferences attended by judges, state judicial council meetings, state reports, inter-

1 In Appendix C (https://bit.ly/3ZugNQ6), we include three in-depth case studies from these examples. A comprehensive review of related works can be found in Appendix A (https://bit.ly/3ZugNQ6)

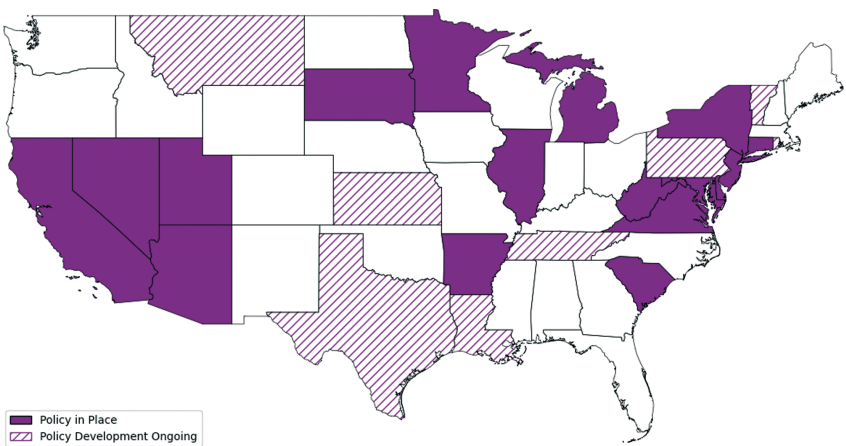
2 We focus on state courts as no federal courts have released formal, public policies governing the use of AI by their employees or judicial officers.

3 https://bit.ly/3ZugNQ6.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY



(a) Spectrum of AI Use Cases in the Courts



(b) Map of State Court Systems that Have Implemented Policies

Figure 1: AI Use Cases and Policy Implementation

views of judges, news articles, speeches, and judicial opinions. Second, to identify internal policies, we used keywords to search (1) each state's judicial branch or judicial council website, along with all federal appellate court websites and the federal Judicial Administration website, (2) general search engines for news articles and reports.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

B. Range of Use Cases

Table 1 provides an overview of AI use cases across countries. We include forms of Robotic Processing Automation (RPA), despite the fact that RPA may not clearly involve machine learning, as these are typically described interchangeably with AI by courts and judges. While fully documented examples of AI use are limited, numerous judges have expressed substantial interest. As Magistrate Judge Allison Goddard put it, "Anyone that comes up with a tool that can accurately summarize a social security record will be worth their weight in gold to every magistrate and district judge!"[1] Table 1 classifies use cases as "administrative" (i.e., aiding in the administration of judicial process) or "judicial" (i.e., aiding in the exercise of judicial discretion). In addition, we provide a rough typology of risk. As we discuss in Section 3, we caution that the same use can be designed and integrated in ways that heighten or mitigate risk, just as administration can bleed into judicial discretion. We also confine ourselves here to AI uses cases outside of administrative adjudication (within the executive branch), but we do include the use of AI for administration in an independent court.

I. Administrative Use Cases

1. Lower Risk Examples

- OCR and Robotic Process Automation for Automatic Case Filing and Categorization:
  Multiple jurisdictions have implemented automated document processing systems that handle court filings. Palm Beach County Courts, along with other jurisdictions, operates one of the most extensive systems, processing approximately 12,000 filings weekly through systems that automatically scan, classify, and assign to judges incoming documents.[2] This system has been operational since 2020 and represents a significant automation of traditionally labor-intensive clerical work.
- HR Tasks and Invoice Processing:
  The District of Columbia Courts have implemented an AI-automated financial processing system that handles invoice receipts and vendor verification.[3] Orange County Courts uses two AI chatbots to provide administrative and legal training to new employees and to assist HR

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

employees with managing personnel records.[4] A municipal court in Texas has begun using ChatGPT to write job descriptions and performance improvement plans, claiming to reduce time spent on such tasks by 60 percent.[2]

- Chatbots for Employee and Basic Customer Support:
  Multiple courts, including Maricopa County Courts, Miami Courts, New Jersey State Courts, and Los Angeles Superior Court, have developed chatbot systems for customer support with varying levels of efficacy and sophistication. These chatbots allow for litigants to ask questions similar to what they might ask at the front desk of a physical courthouse.[2,4–6] While many of these courts do not specify what aspects of their system integrate AI and how, they do claim AI is used in the tool. Independent audits of state chatbots for unemployment assistance have found them to be limited, but no similar audit has been done on these judicial branch chatbots.[7]
- Social Media and Keyword Tracking:
  Maricopa County Courts uses a system to monitor social media platforms for concerning language patterns, conduct keyword analysis to analyze sentiment and threat levels, and identify potential threats against judicial officers.[4]
- Docket Scheduling:
  The New Jersey state court system has developed an AI system that schedules its docket, although detailed information on how that system operates is not currently publicly available.[2]
- AI Court Reporters:
  The Arizona Supreme Court created two lifelike generative AI avatar "reporters" who summarize the court's decisions in YouTube videos posted by the court intended to make court opinions more accessible to the public.[8] Critics worry that AI summarization of opinions, rather than independent journalists analyzing the case, could lead to bias.
- Triaging and Prioritizing Cases:
  Several courts use AI systems to triage incoming filings and prioritize the most urgent cases. The Massachusetts Appeals Court developed a tool summarizing emergency eviction appeals for easier review. Colombia's Supreme Court identifies for expedited review health related cases requiring urgent attention due to potential harm, and Brazil's Supreme Court also uses an AI system that prioritizes pending appeals cases based on potential constitutional significance.[9]

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Use Case Type	Examples
ADMINISTRATIVE	
OCR and Robotic Process Automation *Lower Risk*	• Palm Beach County Courts: 12,000 weekly filings processed automatically • Tarrant County Courts: Real-time e-filing monitoring and document classification
HR Tasks and Invoice Processing *Lower Risk*	• DC Courts: AI-automated financial processing system • Orange County Courts: EVA chatbot for employee training, EMI for personnel records • Texas municipal court: ChatGPT for job descriptions and performance plans
Chatbots for Employee and Basic Customer Support *Lower Risk*	• Miami Courts: SANDI for courthouse inquiries in English/Spanish • Maricopa County: Cleo (public-facing), YODA (employee policies), Alfred (IT support) • New Jersey: JIA for public questions and court information • LA Superior Court: GINA for ticket payments and jury duty scheduling • Hawai'i State Judiciary: Kolokolo Chat for court information
Social Media and Keyword Tracking *Lower Risk*	• Maricopa County Courts: AI monitoring for threats against judicial officers
Docket Scheduling *Lower Risk*	• New Jersey state courts: AI-powered docket scheduling system
AI Court Reporters *Lower Risk*	• Arizona Supreme Court: AI avatars summarizing court decisions for YouTube
Triaging and Prioritizing Cases *Lower Risk*	• Massachusetts Court of Appeal: LLMs to summarize emergency eviction appeals for review • Colombia Supreme Court: NLP for urgent health-related claims • Brazil Supreme Court: ML algorithms screening appeals for review
Automated Transcription and Translation *Higher Risk*	• India Supreme/High Courts: Real-time transcription and multi-language translation • Singapore: AI transcription across all state courts • New Mexico and Bahamas: FTR RealTime implementation
Chatbots for Guidance *Higher Risk*	• Alaska State Courts: AVA for probate estate cases • New Mexico Courts: Clara for self-represented litigants • LA Superior Court: Triage and referral tool development
Flagging Issues in Filings and Procedural Compliance *Higher Risk*	• LA Superior Court with Stanford: AI analysis of court filings for procedural issues • Florida Small Claims: AI review of Civil Procedure rules compliance
JUDICIAL	
Summarization of Filings and Question Generation *Lower Risk*	• New Jersey: NJ Courts AI Access for 1,000 users including all judges • Judge Xavier Rodriguez: Case summarization and attorney question generation • Judge Allison Goddard: Technical issue questions and document organization
Creation of Timeline of Events *Lower Risk*	• Judge Xavier Rodriguez: LLMs for key event timelines
Asking LLMs Questions Related to Case Facts *Lower Risk*	• Dutch judge: Scientific questions related to case facts • Mexican judge: Contextual analysis of phrases/concepts • Indian judge: Historical legal concepts like bail jurisprudence
Cross-media Search *Lower Risk*	• Nebraska courts: Search across text, audio, video case exhibits • Judge Anthony Capizzi: IBM Watson for juvenile case multi-media search
AI Tools for Legal Research *Higher Risk*	• Ninth Circuit: WestLaw Quickcheck Judicial tools for all judges • Selected judges testing additional LLM applications
Using LLMs for Legal Analysis and Interpretation *Higher Risk*	• Judge Kevin Newsom: ChatGPT for "ordinary meaning" in Snell v. United Specialty Insurance. Continued use in United States v. DeLeon • D.C. Circuit: ChatGPT for "reasonable" facts determination in Ross v. United States
Anticipating Outcome of Cases *Higher Risk*	• Australia: Predictive algorithms for divorce property division
Assistance with Opinion Writing *Higher Risk*	• Shenzhen Intermediate People's Court: Intelligent Adjudication System for opinion drafting • Use of LLMs to draft court orders in U.S. District Courts (released in error)
Auto-adjudication and Assisted Adjudication *Higher Risk*	• China: Fully automated online courts for e-commerce/IP disputes • Argentina: AI-generated draft rulings based on previous decisions • Mexico: Rule-based expert systems for pension eligibility

Table 1: AI Use Cases in Courts by Type, Examples, and Risk Level

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

2. Higher Risk Examples

– Automated Transcription and Translation:
Some courts use automated transcription and translation in their proceedings. India's Supreme and High Courts have systems that simultaneously convert spoken proceedings into written text while providing translation services across multiple regional languages.[9] Singapore provides AI-powered court transcription services across all state courts, and courts in New Mexico and the Bahamas have started using FTR RealTime, an AI tool that instantly creates text from speech.[4] Court-automated transcription and translation run the risk of violating due process if independent evaluation is not in place to ensure that the text has not been tampered with or changed to serve a certain interest.[10]
– Chatbots for Litigant Guidance:
Some courts have developed AI systems that provide procedural guidance for litigants. Alaska State Courts operates AVA (Automated Virtual Assistant), developed through a third-party vendor partnership, which helps users identify appropriate court processes and forms for probate estate cases.[11] The system asks diagnostic questions to determine user needs and provides tailored guidance. New Mexico Courts and Los Angeles Superior Court have deployed chatbots specifically designed to communicate with and assist self-represented litigants,[12] and the Los Angeles Superior Court is also developing a triage and referral tool meant to communicate with and assist self-represented litigants.[13]
– Flagging Issues in Filings and Procedural Compliance:
The Los Angeles Superior Court, in partnership with researchers from Stanford Law School, has prototyped AI to review submitted filings and automatically inform litigants of potential formatting, informational, jurisdictional, or statute of limitations issues.[13] Similarly, the Florida Small Claims Court previously reviewed each of its filings to ensure the Rules of Civil Procedure were invoked, but replaced this task with an AI tool, reducing review time from 15 to 2–3 hours a week.[2] In some instances, use of AI in this domain can be a short term "band-aid" solution for an underlying structural deficiency of the court system; for example, using AI to perform OCR on a paper document and analyze its contents to find missing information could be avoided by simply requiring litigants to fill out an online form with separate fields.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

II. Judicial Use Cases

1. Lower Risk Examples

- Summarization of Filings and Question Generation:
  Some judges in interviews have described using LLMs for summarization and information consolidation. U.S. District Judge for the Western District of Texas Xavier Rodriguez uses LLMs to summarize cases, identify key players involved, and generate questions for attorneys based on their materials.[14] Magistrate Judge for the Southern District of California Allison Goddard uses LLMs to understand what questions to ask attorneys about a technical issue, to summarize and ask questions about orders from other judges, and to organize information from documents.[14]
- Creation of Timeline of Events:
  Judges, such as Judge Xavier Rodriguez, have used LLMs to analyze case documents to create a chronological timeline of relevant events in the case.[14]
- Asking LLMs Questions Related to Case Facts:
  Several judges have been documented consulting LLMs about facts related to their cases, phrases in their cases, and relevant legal history, including judges in the Netherlands, Mexico, and India.[15–17]
- Cross-media Search Across Text, Audio, and Video Case Exhibits:
  Nebraska courts have implemented search capabilities that can analyze and cross-reference information across all forms of case exhibits, including written documents, audio recordings, and video evidence,[18] while former Ohio Judge Anthony Capizzi used IBM Watson to search across his filings and spend more time engaging with the juveniles in his court.[19]

2. Higher Risk Examples

- AI Tools for Legal Research:
  The U.S. Court of Appeals for the Ninth Circuit has implemented a formal program authorizing all judges to use Westlaw Quick Check Judicial tools, and some to use additional LLMs under controlled conditions.[1] Concerns still exist about the accuracy of these tools' outputs.[20,21]

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

- Risk Assessments:
  COMPAS, a software that uses data about a defendant to predict their likelihood of recidivism, has been used by courts since as early as 2016. Because of the extensive scholarship around risk assessment tools, we limit our discussion of these here.
- Anticipating Outcome of Cases:
  Australia employs predictive algorithms to anticipate property division outcomes in divorce proceedings, providing judges with statistical analysis of similar cases and suggested property distributions based on relevant factors.[9] This system analyzes historical case outcomes, financial circumstances, custody arrangements, and other relevant variables to suggest likely resolution ranges.
- Using LLMs for Legal Analysis and Interpretation:
  Judge Kevin Newsom of the U.S. Court of Appeals for the Eleventh Circuit was the first U.S. judge to openly rely on and discuss the use LLMs for legal interpretation in a judicial opinion. In *Snell v. United Specialty Insurance Co.*, Judge Newsom consulted ChatGPT to determine the ordinary meaning of "landscaping" in an insurance coverage dispute involving a below-ground trampoline installation. *Snell v. United Specialty Insurance Co.,* 102 F.4th 1208 (11th Cir. 2024). He continued this practice in *United States v. DeLeon*, using AI assistance from multiple sources for ordinary meaning analysis in his concurrence. *United States v. DeLeon*, 116 F.4th 1260 (5th Cir. 2024). The D.C. Circuit subsequently cited Judge Newsom's precedents while using ChatGPT to assess whether certain facts were "reasonable" to know in *Ross v. United States*. 331 A.3d 220 (D.C. Cir. 2025).
  In both cases, judges acknowledged the experimental and potentially controversial nature of their approach. Critics have argued that such use constitutes inappropriate independent fact-finding outside the adversarial process, and that commercial LLMs are subject to corporate manipulation and training methods that may not reflect genuine ordinary usage patterns.[22]
- Assistance with Opinion Writing:
  The Shenzhen Intermediate People's Court in China deployed a system that analyzes case materials to summarize factual information, identify the central legal disputes, generate targeted questions for judges to ask, and write the factual sections of court opinions.[23] Once the judge indicates the decision they wish to make in the case, the system helps craft other sections of the opinion as well.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

In interviews, judges using the system expressed concerns about potential over-reliance leading to constrained judicial thinking. Researchers also found evidence of slight bias in judicial rulings made with LLM assistance, not from the AI system itself but from the reinforcement of existing judicial biases.

– Auto-Adjudication and Assisted Adjudication:
China operates some fully automated online courts handling certain e-commerce and intellectual property disputes.[24] These systems manage all aspects of case processing from analyzing filings, conducting discovery, facilitating hearings through chatbots, and applying legal rules, to rendering binding decisions without human judicial involvement.
Argentina has implemented systems in some courts that generate complete draft rulings for new cases from previous similar cases. While judges review these drafts, the AI systems provide initial analysis and suggested outcomes.[9] Mexico's pension courts have a system that generates preliminary decisions for review as well.[25]

C. Internal Policies

U.S. state judicial entities have been at the forefront of developing policies and guidelines addressing the use of AI within the judiciary (Table 3 and Figure 1b). We hence focus on these policies here.

State	Approved Tools									
	Copilot	Claude	ChatGPT	Gemini/Bard	Perplexity	Adobe AI	WestLaw Precision	LexisAI+	Zoom AI Companion	FTR Justice Cloud
Arizona	✓	✓	✓	✓	✓	✓	✓		✓	✓
Maryland	✓	✓	✓	✓						
New York	✓		✓		✓					
Utah		✓	✓	✓						
Virginia							✓	✓		
Connecticut										
Delaware										

Table 2: AI Tools Approved in Different State Court Jurisdictions. States with grey rows require supervisory approval of tools but do not publish approved tools.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

State	Policy Characteristics							Appr. Req.		Flexibility		Ind.
	Bindingness	Number of Pages	Mentions GenAI	AI Use Case Inventory	Disclosure	Monitoring and Evaluation	Includes Judges	For Tool Use	For Use Cases	Local Policy	Review Regularly	Restrictiveness Index
Connecticut	✓	21	✓	✓	✓	✓		✓	✓		✓	0.481
Utah	✓	2				✓	✓	✓	✓			0.296
New York	✓	5	✓				✓	✓				0.148
Maryland		3	✓				✓	✓	✓			0.148
Minnesota	✓	1				✓	✓					0.148
New Jersey		2				✓	✓		✓			0.148
South Carolina	✓	4	✓			✓	✓					0.148
Arizona	✓	4	✓	✓		✓	✓	✓		✓	✓	0.111
Arkansas	✓	3	✓				✓					0.111
California	✓	2	✓		✓		✓			✓	✓	0.111
Michigan		24	✓			✓	✓					0.111
South Dakota		3	✓		✓							0.111
Delaware	✓	4	✓				✓				✓	0.074
Virginia	✓	4	✓					✓		✓	✓	0.037
West Virginia		5					✓					0.000
Illinois		1	✓				✓				✓	0.000
Nevada		2	✓				✓					0.000

"Policy Characteristics" indicates what provisions the policy contains. "Bindingness" indicates whether the policy is intended to be binding to the judicial officer or employee, or conversely whether it is meant to be a non-binding recommendation. "Number of Pages" is the length of the policy, thereby a proxy for the number and specificity of its stipulations. "Mentions GenAI" indicates whether the policy specifically includes generative AI. "Use Case Inventory" indicates whether the state's courts are required to create an inventory of all AI or GenAI uses. "Disclosure" indicates whether there are any provisions requiring public disclosure of use of AI or generative AI for textual, image, or other outputs from the courts. "Monitoring and Evaluation" indicates whether there exists a provision requiring some sort of monitoring of procured or in-house AI or GenAI tools. "Includes judges" indicates whether the policy includes judges or not in its scope. "Appr. Req." indicates that leadership approval is required either "for tool use" (see Table 6)—any AI or GenAI tool being used must be pre-approved first—or "for use cases," meaning approval is required before implementing any new applications. "Flexibility" indicates that the policy contains provisions that have some degree of permissiveness; under that, "Local Policy" indicates that local courts have discretion to develop different policies for their own jurisdictions and "Review Regularly" indicates the policy contains a provision to review and update it at some time. "Ind." indicates index, and the restrictiveness index is calculated to indicate the degree to which the given policy restricts potential usage of AI by judicial officers and court employees. The index was calculated as follows: +3 points for characteristics we consider "strongly restrictive" (bindingness of the policy, whether disclosure of AI use is required, and whether leadership approval is required for use cases); +2 points for characteristics we considered "somewhat restrictive" (if the policy is more than 10 pages long); +1 point for characteristics we consider "loosely restrictive" (AI use case inventory is required; monitoring and evaluation of tools is required; leadership approval is required for specific tool usage); -1 point for characteristics that offer loose flexibility (policy is reviewed regularly and potentially revised); -2 points for characteristics that offer moderate flexibility (differing policies are allowed for local courts). The overall score is then divided by the maximum possible score (27) to create the index.

Table 3: Comparative Analysis of Policies.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

I. Current Landscape of Internal AI Policies Across State Courts

A threshold question is whether AI specific policies are necessary. Existing Rules of Court and Codes of Judicial Ethics already govern judicial use of AI to some extent. Illinois's AI fact sheet for judges, for instance, highlights its Code of Judicial Conduct, which requires judges to "decide matters assigned to the judge" and "act at all times in a manner that promotes public confidence in the independence, integrity, and impartiality of the judiciary."[26] Prior efforts to address the rise of technology may also apply. Courts, for instance, have grappled with judicial notice of websites, when such websites can be altered or deleted.[27]

Seventeen states have promulgated AI-specific policies, with eight other states planning to do so. Table 3 characterizes some of the key differences across existing policies. One notable dimension is the force of these policies: six are non-binding recommendations, one is a non-binding advisory opinion, four are administrative orders, and six have amended Rules of Court or the Standard of Judicial Administration (or analogues). The last column of Table 3 also provides a rough measure of the facial restrictiveness of these policies.

II. Discussion

1. The Enforceability Question

These policies present a core question about enforceability. To begin, six states have issued merely advisory, non-binding policies. Second, for states purporting to issue binding policies, enforcement is clearer for employees of the court – with typical employment sanctions – but less clear for judges. California's Judicial Council maintains one policy for court employees and judicial officers in their non-adjudicative roles in the Rules of Court, and another policy for judicial officers in their adjudicative role is adopted into the Standard of Judicial Administration, a set of strongly encouraged "guidelines."[28] Yet prior work has demonstrated significant variability in mechanisms to enforce against judicial misconduct.[29] Third, enforcement against judges raises questions of judicial independence. Connecticut, for instance, appears to require preapproval for the use of AI tools. If a Connecticut judge followed Judge Newsom to explore the use of an LLM for statutory interpretation, should such behavior be sanctioned as violating the policy or be controlled through ordinary appellate review of judicial interpretive methodology?

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Outside of the judiciary, we have already seen the limits of these policies. New York State's audit of its initial state executive branch agency AI policy revealed widespread "inadvertent noncompliance" and significant reliance on individual discretion in determining appropriate use, despite formal policy requirements.[30] While some jurisdictions attempt to strengthen compliance through deterrent language, such as North Carolina's state executive branch policy's warning of "severe penalties" for violations,[31] the effectiveness of such approaches remains untested in the judicial context.

2. Policy Ambiguity

Heightening the enforcement challenge is significant ambiguity in these policies. West Virginia's advisory opinion declares "A judge should NEVER use AI to reach a conclusion on the outcome of a case."[32] But the boundary between AI for legal research and "reaching" a conclusion is muddy. Other states, such as California, do not explicitly prohibit its courts from using AI for adjudicative purposes, but instead provide high level principles – privacy, bias, accountability, and transparency – when much scholarship has focused on often subtle tensions between such principles. Some policies, such as Connecticut's, require citation of "use of AI" when "required by law," although Connecticut's does not specify what "use of AI" encompasses, nor which laws would be relevant. "Use of AI" could range from fixing grammar and spelling mistakes to complete rewording, and clearer language would make these policies easier for judicial officers and employees to achieve compliance. All these efforts obscure the difficulty of drawing lines given the spectrum of AI integration (Figure 1a).[33]

3. Regulatory Burden

What is also notable is the significant variation in the regulatory burden for judicial adoption of AI, as measured by the restrictiveness index. At one extreme, court systems in states such as Connecticut have frameworks that impose extensive compliance obligations on judicial personnel, creating detailed procedural requirements that, while thorough, may prove burdensome. In Connecticut, a judicial branch division looking to use an AI system must submit a request to the State Judicial Branch's AI Committee describing the purpose of the request and considerations relating to various privacy, IP, and procurement concerns, agree to monitoring and

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

evaluation while using the tool, acquire approval from a supervisor before each new use of the tool, and cite use of AI "where required by law."[34] At the other extreme, court systems in Illinois and Nevada have adopted minimalist approaches, offering a small amount of optional guidance for judicial officers and court employees alongside unrestricted AI use.[26,35]

California and Arizona's judicial policies present a more balanced regulatory posture, establishing guardrails alongside relatively light reporting requirements. California requires courts to take measures to ensure privacy, bias minimization, accountability, and transparency when using the tools, and also allows courts to individually choose to unilaterally ban use of generative AI if they so wish; Arizona allows AI and GenAI use with certain tools, requires courts to seek approval for tools and use cases outside that realm, and documents use cases publicly on its website.[28,36]

The core challenge is calibrating safeguards to risk, given the spectrum of AI integration. As illustrated in Section 3, some use cases pose minimal risks to core judicial functions, and too much process may lead to suboptimal innovation. While California's policy presents separate policies for AI for adjudicative uses and AI for non-adjudicative uses, it fails to account for the differing risks within those categories. While much focus has been on ex ante procedural controls, too little attention has been placed on ex post evaluation and post-deployment monitoring. One positive counterexample is that Michigan recommends the establishment of pilot programs, qualitative and quantitative evaluation, and iteration for AI tool use before scaling.[37]

4. Transparency

The challenges we faced in identifying judicial use of AI illustrate why others have attempted to codify such requirements for the executive branch. Some of the policies improve transparency around court and judicial use of AI. Connecticut and Arizona require judicial officers and court employees to report use cases to their respective AI committees to "identify emerging use cases and opportunities for knowledge sharing" and "maintain a list of AI tools" respectively.[34] Such transparency allow third parties – or other courts – to evaluate and potentially adopt the use case themselves. Arizona's online use case inventory provides insight into its operations, and court employees who have championed its AI use have become sources of expertise to other jurisdictions.[37]

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

That said, the implementation of an AI use case inventory is not straightforward. In the U.S. federal context, 76 % of agencies failed to publish an inventory within two years, including 48 % of large agencies actively known to have AI use cases in 2020.[38] Currently, Connecticut's Judicial Branch Artificial Intelligence Committee website announces that it "conducted an inventory of its internal systems and has determined that it does not employ artificial intelligence in its applications."[34]

Table 2, for instance, shows that while preapproval of AI tools is required, no information exists in Connecticut and Delaware on the preapproved tools. Tool-based approval may also obscure the range of ways a specific tool may be used. It is perplexing, for instance, that Maryland approved ChatGPT, but not LexisAI, when legal AI tools have reduced hallucination rates.[20] Consider the contrast of Judge Newsom. While one can dispute the appropriateness of utilizing Claude and ChatGPT for statutory interpretation, one virtue is the transparency that triggered extensive commentary on the approach.[22,39–41]

Acknowledgments

We thank the Daimler and Benz Foundation for supporting this work.

References

1. Generative AI in Action – Exploring Generative AI Tools for Legal Practice. Ninth Circuit Judicial Conference. Youtube. (2024). https://perma.cc/82DR-MVMR.
2. National Center for the State Courts. Artificial Intelligence in State Courts: Promise and Peril. (2025). https://perma.cc/J2LF-6K44.
3. Thomson Reuters Institute/National Center for State Courts AI Policy Consortium. AI in Action – Current Applications in State Courts. (2025). https://vimeo.com/1082965262.
4. National Association for Court Management. Courting AI: Understanding Artificial Intelligence in Courts. (2025). https://nacmnet.org/resources/store/.
5. New Jersey Courts. NJ Courts AI Access: How the Judiciary is Using Secure, Internal AI to Improve Court Services. (2024). https://njsbamaterials.com/MAM25107.pdf.
6. Sigler, E. SANDI: Improving Court Access and Service in Miami with an Advanced Artificial Intelligence Chatbot. Trends in State Courts. (NCSC, 2023).

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

7. Ho, D.E., Martin, O., Rodolfa, K., & Perez, A. Evaluation as Due Process: Civil Service in an Automated Age. Adm. Law Rev. (forthcoming). (2025).
8. Arizona Supreme Court Administrative Office of the Courts. Arizona Supreme Court Introduces AI-Generated Court News Reporters to Enhance Public Engagement. [Press release]. (2025).
9. UNESCO Global Judges' Initiative: Survey on the Use of AI Systems by Judicial Operators. (2024). https://perma.cc/PBD2-WFSN.
10. Jones, T., Kalbfeld, J. R., Hancock, R. & Clark, R. Testifying while black: An experimental study of court reporter accuracy in transcription of African American English. *Language* **95**, 216–252 (2019).
11. LawDroid. Alaska Courts Virtual Assistant: Preview and Demo. Youtube. (2025). https://www.youtube.com/watch?v=ZC0_Q0HnAQA.
12. New Mexico Courts. Language Access Services: Clara. New Mexico Courts. https://languageaccess.nmcourts.gov/clara/.
13. Lynch, S. Harnessing AI to Improve Access to Justice in Civil Courts. Stanford HAI. (2025). https://hai.stanford.edu/news/harnessing-ai-to-improve-access-to-justice-in-civil-courts.
14. O'Donnell, J. Meet the early-adopter judges using AI. MIT Technology Reviewhttps://www.technologyreview.com/2025/08/11/1121460/meet-the-early-adopter-judges-using-ai/.
15. Pascoe, R. Dutch judge uses ChatGPT to help reach a verdict. DutchNews. (2024). https://perma.cc/EX5S-3PAL.
16. Gutiérrez, J.D. Judges and magistrates in Peru and Mexico also caught the ChatGPT 'fever.' Agenda Estado de Derecho. (April 18, 2023). https://perma.cc/LQJ7-QL6V.
17. Sandhu, J. S. In a first, Punjab and Haryana HC turn to ChatGPT for view on bail in murder case. The Indian Express. (2023). https://perma.cc/C54E-KX4T.
18. Shetty, A. Nebraska Judicial Branch modernizes its Electronic Exhibits System using AWS. AWS Public Sector Blog (2023). https://perma.cc/SVJ5-JHMS.
19. Capizzi, A. Helping Juvenile Treatment Court Improve Efficiency and Outcomes with IBM's Watson Health Solution. Int'l Pers. of Crime Prevention. **12**, 67–74. (2023)
20. Magesh, V. *et al.* Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. J. Empir. Leg. Stud. 22, 216–242 (2025).
21. Dahl, M., Magesh, V., Suzgun, M. & Ho, D. E. Large legal fictions: Profiling legal hallucinations in large language models. J. Leg. Anal. **16**, 64–93 (2024).
22. Grimmelmann, J., Sobel, B. & Stein, D. Generative Misinterpretation. Harv. J. Legis. **63**, (2025).
23. Liu, J. Z. & Li, X. How do judges use large language models? Evidence from Shenzhen. J. Leg. Anal. **16**, 235–262 (2024).
24. Guo, M. Internet court's challenges and future in China. Comput. Law Secur. Rev. **40**, 105522 (2021).
25. Valverde, D. Supreme Court of Mexico Unveils AI Sor Juana for Judicial Clarity. Mexico Business News. (June 27, 2024). https://perma.cc/S65W-EWJQ.

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

26. Illinois Supreme Court Policy on Artificial Intelligence. Supreme Court of Illinois. (January 1, 2025). https://perma.cc/4HQR-N65G.
27. Barger, C. M. On the Internet, Nobody Knows You're a Judge: Appellate Courts' Use of Internet Materials. J App Pr. Process **4**, 417 (2002).
28. Rule and Standard for Use of Generative Artificial Intelligence in Court-Related Work. Judicial Council of California Artificial Intelligence Task Force. (2025). https://perma.cc/QV7E-SFSM.
29. Abel, J. Testing three commonsense intuitions about judicial conduct commissions. Stanford L. Rev. **64**, 1021 (2012).
30. DiNapoli: Improved Guidance Needed for State Agencies Using AI To Avoid Risks. Office of the New York State Comptroller [Press release]. (2025). https://perma.cc/LGB9-EB9U.
31. Use of Publicly Available Generative AI. North Carolina Department of Information Technology. https://perma.cc/5VZP-SS42
32. RE: Judicial Investigation Commission Advisory Opinion 2023–22. West Virginia Judicial Investigation Commission. (2023). https://perma.cc/H77X-8GH8.
33. Martin, O., Surani, F., Rodolfa, K., Perez, A. & Ho, D. E. The Spectrum of AI Integration: The Case of Benefits Adjudication. AI: Legal Issues, Policy, and Practical Strategies (American Bar Association, 2024).
34. State of Connecticut Judicial Branch. *JBAPPM Policy 1013 Artificial Intelligence Responsible Use Framework*. (2024).
35. Artificial Intelligence Guide. Nevada Administrative Office of the Courts. (2025). https://perma.cc/9LPU-RVW2
36. Section § 1–509: Use of Generative Artificial Intelligence Technology and Large Language Models. Arizona Court Rules. (2024). https://perma.cc/WL3H-84R4.
37. Generative AI and the Courts Workgroup: Report and Recommendations. Michigan Judicial Council (2024). https://perma.cc/GQE7-WADP.
38. Lawrence, C., Cui, I. & Ho, D. The Bureaucratic Challenge to AI Governance: An Empirical Assessment of Implementation at U.S. Federal Agencies. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. 606–652 (ACM, 2023).
39. Curl, J., Henderson, P., Kandula, K. & Surani, F. Judges Shouldn't Rely on AI for the Ordinary Meaning of Text. Lawfare (2025).
40. Engel, C. & McAdams, R. H. Asking GPT for the ordinary meaning of statutory terms. U Ill JL Tech Pol 235 (2024).
41. Waldon, B., Schneider, N., Wilcox, E., Zeldes, A. & Tobia, K. Large Language Models for Legal Interpretation? Don't Take Their Word for It. Georgetown L. J. **114** (forthcoming). (2025).

https://doi.org/10.5771/9783748947257-509 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY