

Parity in Practice:

THE DEFENDER'S DUTY
TO ETHICALLY USE AI

*A white paper on ethical and practical use of
generative AI by criminal defense lawyers*

Mason R. Clark





Copyright© 2026 National Association of Criminal Defense Lawyers

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

It may be reproduced, provided that no charge is imposed, and the National Association of Criminal Defense Lawyers (NACDL) is acknowledged as the original publishers and the copyright holders. For any other form of reproduction, please contact NACDL for permission.

For more information contact:

National Association of Criminal Defense Lawyers®

1660 L Street NW, 12th Floor, Washington, DC 20036

Phone 202-872-8600 [NACDL.org](https://www.nacdl.org)

This publication is available online at [NACDL.org/ParityInPractice](https://www.nacdl.org/ParityInPractice)

A version of this article is forthcoming in the inaugural edition of *The Journal of Advocacy and Litigation* (Vol. 1, 2026).

Parity in Practice:

THE DEFENDER'S DUTY
TO ETHICALLY USE AI

Andrew S. Birrell

President, NACDL
Minneapolis, MN

Lisa M. Wayne

Executive Director, NACDL & NFCJ
Washington, DC

Kyle O'Dowd

Deputy Director, NACDL
Washington, DC

REPORT AUTHOR

Mason R. Clark

Assistant Professor of Law
St. Mary's University School of Law
San Antonio, TX

Table of Contents

About NACDL	5
Acknowledgements	6
Executive Summary	7
Introduction	9
A. Purpose and Scope	9
B. Why This Matters Now	9
C. Timeline for Implementation	11
Part I:	
AI Overview	12
A. What Is Generative AI?	12
B. Use of AI in Defense	16
C. The Defense Lawyer's Duty to Use AI	20
Part II:	
AI Ethics and the Model Rules of Professional Conduct	23
A. Helpful Resources to Guide AI Use	23
B. Balancing Duties of Competence, Confidentiality, and Candor	25
Part III:	
Practical Guidance For Defense Lawyers	33
A. Practical Guidance for Defense Lawyers	33
B. AI Use Policies	41
C. Additional Policy Recommendations	45
Conclusion	47
Endnotes	49
Appendix A	56
Model AI Use Policy	56



About NACDL

The National Association of Criminal Defense Lawyers (NACDL) envisions a society where all individuals receive fair, rational, and humane treatment within the criminal legal system. NACDL's mission is to serve as a leader, alongside diverse coalitions, in identifying and reforming flaws and inequities in the criminal legal system, and redressing systemic racism, and ensuring that its members and others in the criminal defense bar are fully equipped to serve all accused persons at the highest level. NACDL is committed to enhancing the capacity of the criminal defense bar to safeguard fundamental constitutional rights. NACDL harnesses the unique perspectives of NACDL members to advocate for policy and practice improvements in the criminal legal system.

NACDL also actively engages with emerging technologies through its Fourth Amendment Center and related programming and task forces, educating defense attorneys on issues like probabilistic genotyping, forensic genetic genealogy, and the impacts of AI on policing, sentencing, and privacy rights. Additionally, NACDL has launched a Task Force on Artificial Intelligence to establish policy and the ethics of artificial intelligence and to monitor the impact that AI will have on criminal defense lawyers and their clients.

NACDL also believes "Justice is Green," and is committed to being a bar organization that leads in criminal legal reform as well as various going green initiatives. This commitment prioritizes NACDL's respect for the planet alongside its vision for a more just and fairer criminal legal system. NACDL believes we all need to modify our behavior to ensure the wellbeing of our planet and its inhabitants; for NACDL, this commitment will be reflected in its conduct as well as its messaging to other bar associations around the country.

Finally, NACDL has established the NACDL Foundation for Criminal Justice (NFCJ), a 501(c)(3) registered charity. The mission of the NFCJ is to preserve and promote the core values of the criminal legal system guaranteed by the Constitution — among them due process, fair sentencing, and effective assistance of counsel — by educating the public and the legal profession to the role of these rights and values in a free society.



Acknowledgements

The author would like to thank Jonathan Brayman (Breen & Pugh, Chicago), Melanie Foote (Kentucky Department of Public Advocacy), Katherine Tang Newberger (Office of the Federal Public Defender, Maryland), Giselle Pomerleau (Office of the Federal Public Defender, New Jersey), and the National Association of Criminal Defense Lawyers for spearheading this project and providing thoughtful review and suggestions. Additional thanks to the following NACDL staff for their contributions to this report: Lisa Wayne (Executive Director), Kyle O'Dowd (Deputy Director), Jumana Musa (Director, Fourth Amendment Center), Jonathan Hutson (Communications Director), and Cathy Zlomek (Art Director).

As nearly two dozen jurisdictions now require attorneys to provide disclosure and/or verification statements for documents prepared with assistance from or containing information generated by generative artificial intelligence, it seems fitting to include a specific and conspicuous disclosure and verification statement here: This white paper was prepared with the assistance of generative artificial intelligence. The author used various large language models to organize some parts and subsections; to rephrase some content for brevity; to format citations; to summarize lengthy reports or articles; and to compile research material into accessible folders. The author certifies that all citations have been manually verified to ensure their existence and confirm their propositions as accurate representations.



Executive Summary

Generative artificial intelligence (AI) presents both opportunities and risks for criminal defense lawyers.¹ With practical, strategic guidance and dedicated organizational governance, these tools can improve efficiency, support legal research, and help defense lawyers manage overwhelming caseloads and provide effective counsel. However, generative AI raises serious concerns about confidentiality, accuracy, fairness, and ethical compliance, and these concerns impact all participants in the American criminal legal system — from prosecution and defense to judges and juries. Some prosecutors' offices are increasingly experimenting with enterprise-level AI tools, though resource disparities may hinder comparable adoption among some defender organizations. Moreover, AI has already impacted workflows and organizational structures at firms and defenders' offices, and its ubiquity encourages skill flattening.² What is clear, though, is that defenders have an ethical obligation to use AI in their practice and in the courtroom, and a failure to do so could produce incompetent representation of and negative outcomes for criminal defendants.

This white paper provides a framework for ethical and practical AI use across all types of criminal defense practice, grounded in the ABA Model Rules of Professional Conduct and best practices as described by practitioners, academics, and other experts. While AI tools can meaningfully improve efficiency and expand analytical capacity, those benefits are highly contingent on tool selection, implementation, and oversight. The white paper also recognizes that cost, access to training, and institutional resources remain significant barriers — particularly for solo practitioners, small firms, and under-resourced public defense offices — and discusses AI throughout as a set of tools whose value depends on informed, responsible, and ethical use. It explains what generative AI is, identifies its risks and limitations (including emerging issues like deepfakes and climate impacts), surveys its current use by prosecutors and its reception in courts, and provides practical guidance for defenders on important day-to-day decision-making, such as which tools to use and how to use them ethically and efficiently. It uses job-specific examples (e.g., sentencing memos and character statements) to explain where and how AI could be used and highlight potential ethical concerns in its use. Most importantly, it sets out a roadmap for drafting AI policies, including considerations for data privacy and security, verification, training, and compliance with judicial standing orders.



Finally, this white paper acknowledges the realities of technology and the law: The technology is rapidly evolving, and the guidance will likely change, too. With that in mind, this white paper recommends that defenders adopt formal, written AI policies, secure feasible AI tools, implement staff training protocols, and establish flexible yet clear guidelines for permissible and prohibited uses. Flexible guidelines are essential because the goal of any model AI Use Policy is to provide a consistent analytical framework to evaluate and guide use of tools in a dynamic environment. A model AI Use Policy is provided in Appendix A. It is meant to serve as a model only, and practitioners and firms are expected to tailor its provisions in ways that best reflect the workload and culture of each individual attorney or office. The paper also highlights the environmental footprint of AI and recommends strategies for minimizing energy consumption when selecting and deploying AI tools. Defenders must ultimately balance innovation with responsibility, ensuring that generative AI enhances zealous advocacy and competent representation without sacrificing ethical obligations.



Introduction

A. Purpose and Scope

This white paper provides guidance on the ethical, responsible, and practical use of generative AI by criminal defense lawyers, from solo practitioners to firm members to public defenders (collectively, “defenders”). The guidance draws directly from the ABA Model Rules of Professional Conduct and reflects the realities of day-to-day representation and practice. Defenders at every level will find practical frameworks for integrating AI tools into their work while maintaining their professional and ethical obligations.

This white paper also serves as a call to action. Prosecutors across the country have already begun experimenting with enterprise-level AI tools to manage evidence, streamline discovery, and automate review of wrongful convictions and resentencing.³ As the Vanderbilt Project on Prosecution Policy has documented, prosecutors use AI even when it is “often unclear how accurately and fairly these tools complete the task at hand.”⁴ Defenders cannot afford to remain on the sidelines. The adversarial nature of the criminal legal system demands that defenders achieve parity or risk providing representation that falls short of constitutional and ethical standards.

B. Why This Matters Now

Generative AI has entered legal practice at a remarkable pace. When OpenAI released ChatGPT in late 2022, it reached approximately 100 million monthly active users within two months, making it one of the fastest-growing consumer applications in history.⁵ The technology’s ability to predict, analyze, and produce writing strikes at the core of much legal work. As one commentator observed, AI now targets the “bread and butter” of what lawyers do.⁶

Yet adoption remains uneven between prosecutors and defenders, and “[r]ecent years have seen an increase in the use of AI tools to support crime prevention, prosecution, sentencing, and other facets of the criminal [legal] system.”⁷ Judges have largely permitted police and prosecutors to use “black box AI,” which refers to the types of AI tools and systems whose internal decision-making processes are hard to understand or explain, leading to transparency concerns.⁸ Meanwhile, many public defender



offices and solo practitioners lack not only access to enterprise-level AI systems, but also the resources and expertise to challenge software outputs (including AI-generated outputs) as unreliable in court.⁹ This disparity raises urgent questions about competence and fairness, but also requires defenders to proactively engage with the AI tools that are available as a potential equalizer.

The proliferation of generative AI among the legal profession, generally, has been controversial and landed many lawyers in trouble with the court. There are countless recent examples of prosecutors and defenders misusing AI in briefing and in the courtroom, but two recent examples highlight the ethical concerns at the heart of this white paper. The *New York Times* recently published a story about a District Attorney in California who allegedly used AI in an 11-page brief asking the court to deny bail to the accused and keep him in jail, a brief allegedly “rife with errors that bear the hallmarks of generative artificial intelligence” and including “wholesale misrepresentations of the law, as well as quotations that do not actually appear in the cited texts.”¹⁰ And on the other side of the bench, a defender included several misrepresentations of law in a brief before the Illinois Supreme Court, something that was a source of tense questioning during the defender’s oral argument.¹¹ When one of the justices challenged the defender regarding the inaccurate case propositions, the defender said, “I can only tell you that we rely on the Lexis, the flags and so forth, and my associate was assisting me, so I will have to have a long chat.”¹² Just minutes before, the defender admitted to using X’s chatbot Grok to learn more about certain expert testimony requirements, one of the key issues to be reviewed by the Illinois Supreme Court on appeal.¹³

These two examples underscore two serious concerns with overreliance on AI to be discussed in this white paper. First, AI often “hallucinates,” or invents, fabricates, exaggerates, or misstates cases, rulings, and even facts.¹⁴ According to researcher Damien Charlotin, there are at least 500 cases in the United States, identified as of January 6, 2026, where generative AI produced hallucinated content through fabricated, falsely quoted, or misrepresented case law, and a court or tribunal specifically levied warnings or other punishments.¹⁵ Second, in the *Smith* case mentioned above and in countless other cases, lawyers who have been caught using hallucinated case law often blame the mistake on a lack of supervision of younger or more junior associates. As discussed later in the white paper, both prosecutors and defenders are on notice that AI hallucinates, courts can and will sanction the misuse of AI in briefings and in the courtroom, and blaming a junior associate will not relieve an attorney of their ethical obligations or any punishment for violating them.

Beyond the misuse of AI in the courtroom, there are access-to-justice concerns about AI, too. As Professor Drew Simshaw has explained, although legal AI has “the potential to increase the efficiency and cost effectiveness of work done by lawyers” and “help people solve their own legal problems or connect them with licensed legal professionals who can,” overreliance on and legitimization of technology-driven legal services, “will lead to one or more inequitable two-tiered systems.”¹⁶ AI will either be superior, cost more, and only be available to large firms and wealthy clients, or the inverse



will be true.¹⁷ Either way, the benefits of AI depend on who can access and effectively use the technology. Without deliberate efforts to ensure equitable access, “the realization of any combination of these two-tiered systems” risks widening rather than narrowing the justice gap.¹⁸ These concerns matter to NACDL, which has long engaged with emerging technologies through its Fourth Amendment Center and related programming, but has a strong commitment to increasing access to justice.

Finally, NACDL believes “Justice is Green” and remains committed to environmental sustainability. AI systems — particularly large language models — consume substantial electricity and water, with research suggesting that “by 2030, the current rate of AI growth would annually put 24 to 44 million metric tons of carbon dioxide into the atmosphere, the emissions equivalent of adding 5 to 10 million cars to U.S. roadways” and “equal to the annual household water usage of 6 to 10 million Americans.”¹⁹ Defenders should consider these environmental costs when selecting and deploying AI tools, consistent with NACDL’s commitment to reducing its carbon footprint while pursuing a more just criminal legal system.

The tools and platforms discussed in this whitepaper are cited as illustrative examples only. Their inclusion does not constitute an endorsement of any particular product, vendor, or service provider, nor does it imply that the tools described are superior to alternatives not discussed.

C. Timeline for Implementation

Defense lawyers should aim to implement formal AI use policies within one year. However, the rapid pace of technological change may require interim safeguards sooner. To meet the goal of developing and implementing AI policies within a year, defenders must start the process now by:

1. Understanding the basic capabilities and limitations of both generative and non-generative AI tools, and identifying tools that may enhance client representation within their practice.
2. Inventorying current AI usage (if any) in their offices and how AI is being used.
3. Drafting an enterprise-level AI use policy.
4. Training all necessary staff on internal policies related to the use of AI tools.
5. Reviewing any standing orders or local rules in their jurisdictions that govern AI use, and learning how prosecutors in a given jurisdiction are currently using AI tools in practice.

The model AI Use Policy provided in Appendix A offers a starting framework that practitioners can adapt to their specific circumstances.



PART I:

AI Overview

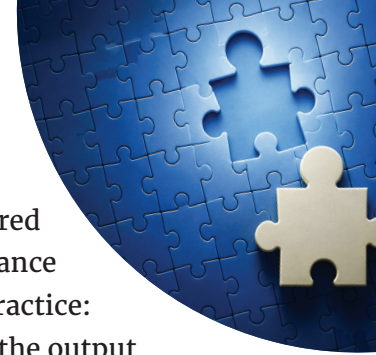
Part I provides an overview of generative artificial intelligence and its emerging role in criminal defense practice. It begins by explaining, at a high level, what generative AI is, how certain systems (e.g., LLMs) work, and why their design and limitations matter in practice. The discussion then turns to the ways defenders are already using generative AI tools in practice, including legal research, drafting, case organization, discovery review, and client communication, while also noting the risks of inaccuracy, bias, confidentiality breaches, and overreliance. Finally, Part I affirms that defenders have an ethical duty to understand and use AI. Together, these sections establish the technological, practical, and ethical foundation necessary to assess how AI should be responsibly integrated into criminal defense practice.

A. What Is Generative AI?

1. *Definition and Core Functionality*

Artificial intelligence encompasses a rather large umbrella of technologies that in some way mimic human intelligence. The National Institute of Standards and Technology defines AI as systems that can “make predictions, recommendations or decisions influencing real or virtual environments” for a given set of human-defined objectives.²⁰ These systems work with machine learning²¹ and data analytics²² to form conclusions using real-time data. However, to make intelligent decisions, computer programmers must first build intelligent algorithms.²³

Generative AI refers to a specific subset of artificial intelligence. These systems use machine learning models trained on vast datasets to create new content (e.g., text, images, code) based on patterns identified in that training data. Large language models (LLMs), such as those powering ChatGPT, Claude, and similar tools, analyze enormous amounts of text and generate responses to user prompts based on the patterns they have learned. As OpenAI CEO Sam Altman explained to Congress in 2023, these models “generate responses by predicting the next likely word in response to the user’s request, and then continuing to predict each subsequent word after that.”²⁴ The capabilities of these LLMs have



advanced rapidly. When OpenAI released GPT-4 in March 2023, testing showed it scored in the top ten percent of July 2022 Uniform Bar Exam test takers.²⁵ This performance demonstrates why generative AI poses both opportunities and challenges for legal practice: The technology can now produce work product that, at least superficially, resembles the output of trained lawyers.

But generative AI is not limited to just LLMs and text generation for legal writing. Beyond large language models, lawyers are beginning to explore other forms of generative AI. Some of these generative AI tools include research and case analysis tools to summarize and compare cases, statutes, and regulations; image generators to create demonstrative exhibits and other visual aids; document review tools to classify documents and flag relevance or privilege; and video generators to “synthesize background sound effects and dialogue based on prompts.”²⁶ Perhaps more significantly, courts are grappling with deepfakes and manipulated audio and video evidence appearing in litigation, prompting the Advisory Committee on Evidence Rules to consider proposed Rule 901(c) governing potentially fabricated electronic evidence.²⁷

2. *Open or Closed Systems*

a. *Open/Closed Source v. Open/Closed System*

Defenders must understand the critical distinction between the meanings of “open” and “closed” in AI discourse. The terms “open” and “closed” carry two related but distinct meanings in AI discourse. One meaning concerns the AI model’s development and distribution. The Open Source Initiative defines “open-source” AI as any system, model, weights, and/or parameters that grant users the freedom to (1) use the system for any purpose without asking for permission; (2) study how the system works and inspect its components; (3) modify the system for any purpose, including to change its output; and (4) share the system for others to use with or without modifications, for any purpose. Open-source models make their underlying code and models publicly available, allowing anyone to inspect, modify, or deploy them. “Closed-source” or proprietary models — such as ChatGPT or Claude — keep their architecture and training data confidential, offering access only through controlled interfaces or application programming interfaces (API). Proponents of open-source models argue they “level[] the playing field by making technology accessible to all” and enable transparency into how the model was trained and what it was trained on, both key metrics for evaluating an AI tool. They can also be run on local hardware with significantly fewer environmental effects. Proponents of the closed-source models, however, contend they allow developers to maintain centralized control over deployment and rapidly patch vulnerabilities, reducing the risk that malicious actors will bypass safety guardrails or deploy models without use restrictions.²⁸



The second meaning concerns the environment in which users use the AI. This is perhaps more relevant to defenders reading this white paper and using or preparing to use AI. A closed environment (sometimes called an “enterprise” or “private” deployment) keeps user inputs (mostly) confidential and segregated, either on-premises (i.e., on a localized server on the firm’s premises) or in a cloud-based instance (i.e., the platform stores the data in the cloud). An open environment, by contrast, may incorporate user interactions into future training data or make them accessible to the provider, law enforcement, or other unintended third parties. This distinction matters particularly for attorneys handling confidential client information. Bar ethics opinions have emphasized that lawyers must understand how their AI tools operate with respect to preserving confidentiality under Model Rule 1.6.²⁹

b. What are the risks of open and closed systems?

Open AI systems use publicly available models, transparent training datasets, and are often freely accessible to consumers.³⁰ When users enter information into open systems, that data may become part of the broader training dataset or otherwise leave the user’s control. Open systems offer certain advantages — including broader training data and greater resources for innovation — but this openness and transparency comes at a cost for confidentiality. Searches performed and content generated through open systems may effectively enter the public domain. Moreover, stored data is likely used to train the system and otherwise shared with third parties.

Closed AI systems, by contrast, operate within self-contained and controlled environments. These enterprise-level deployments house data on private servers or secure cloud instances, limiting access to authorized users within an organization. Closed systems restrict training data to information the organization provides or approves, and they better maintain confidentiality (albeit not entirely) by preventing data from leaving the secure environment. However, closed systems typically require significant financial investment and technical infrastructure that many solo practitioners and under-resourced public defender offices cannot easily afford.

Finally, defenders may learn about “on-prem” (i.e., on premises) AI environments. These are often a subset of closed AI environments in which the data actually stays within the firm’s or office’s own infrastructure and is managed by its own internal information technology and security teams. On-prem AI systems can be beneficial for defenders (and lawyers, generally) because they are more customizable and directly managed by those using the system every day. However, on-prem AI requires upfront spending on the actual hardware and physical infrastructure (e.g., servers, space, and other physical infrastructure, like cooling) that many solo practitioners or under-resourced firms may not be able to afford.³¹



c. *Why does this matter?*

For defense lawyers, this distinction matters enormously. Entering privileged client information, case facts, discovery, or any other materials subject to protective orders into either system risks breaching confidentiality obligations under Model Rule 1.6, but closed systems remain a safer option for defenders who disclose this type of information in an AI tool. State bar opinions have likewise required lawyers to understand how different AI tools handle data.³² Even if the risk of a specific query resurfacing in another user's results seems remote, the possibility is not zero and defenders must weigh that risk against the potential benefits.

3. *Examples in Practice*

Several AI tools have emerged as relevant to legal practice. General-purpose tools include ChatGPT (developed by OpenAI), Claude (developed by Anthropic), and Gemini (developed by Google). These LLMs allow users to generate text (e.g., drafting an email to the client or preparing a section of a brief or motion), summarize documents, and receive answers to questions across virtually any subject matter.

Legal-specific AI tools have also proliferated. Thomson Reuters has integrated AI capabilities into Westlaw through its CoCounsel product, which was the first legal AI assistant powered by OpenAI's GPT-4 technology and now provides agentic functionality using Anthropic's Claude SDK (software development kit). LexisNexis offers Lexis+ AI for legal research and document analysis. Harvey, a startup backed by Sequoia and OpenAI, has partnered with major law firms to develop AI tools tailored specifically for legal work.

But legal-specific AI tools share the same problems as other AI tools and LLMs. Legal tech startups often promote "retrieval-augmented generation" (RAG), which act as LLMs but instead of drawing from the entire internet, RAG models "get their information only from a closed set of data, such as Westlaw, Bloomberg, or the lawyer's own collection of documents, allowing RAG systems to provide accurate citations when answering user prompts."³³ Although legal AI tools may reduce errors compared to the general-use LLMs, RAGs still hallucinate "at an alarming rate."³⁴ Researchers at the Stanford RegLab and Institute for Human-Centered Artificial Intelligence (HAI) found "the Lexis+ AI and Ask Practical Law AI systems produced incorrect information more than 17% of the time, while Westlaw's AI-Assisted Research hallucinated more than 34% of the time."³⁵

Defenders evaluating these tools should consider several factors, including but not limited to: (1) whether the system operates as open or closed (both source and environment); (2) what data security guarantees the vendor provides, such as data retention and deletion controls, encryption standards, role-based access controls, and breach notification procedures; (3) how the system was trained and on what data; and (4) what verification mechanisms exist to catch errors or hallucinations in the output.



B. Use of AI in Defense

Defenders can use AI for defender-specific work, from internal office operations to evidence review to courtroom advocacy. Understanding these use cases helps defenders identify where AI might improve efficiency without compromising ethical obligations. Although this section discusses multiple forms of artificial intelligence (i.e., GenAI tools v. traditional, non-generative tools), the analysis is intentionally broad to provide guidance for using different types of tools. The ethical concerns that arise from AI use often turn less on the specific model or system than on how, where, and why the model or system is deployed. Accordingly, the discussion that follows treats AI as a general category of support and automation tools and emphasizes recurring risks and safeguards that cut across different tools and use cases.

1. *Operations*

AI can streamline many administrative and preparatory tasks within a defense practice. Chatbots and virtual assistants can handle initial client intake, gathering basic information before an attorney becomes involved. AI-powered transcription services can convert audio recordings — including client interviews, witness statements, jail recordings, and court proceedings — into searchable text. Document automation tools can generate routine correspondence, fee agreements, and standard motions from templates.

Legal aid organizations across the country are adopting AI-powered case management software and other tools at almost twice the rate reported across the wider legal profession, and these tools can now conduct initial client interviews, gathering basic information about legal problems, applying eligibility criteria, and producing structured summaries for staff review.³⁶ These systems operate around the clock, meaning a client served with an eviction notice or facing charges can begin the intake process at midnight rather than waiting days for a callback during business hours. For offices serving large geographic areas, this also extends reach to rural clients who may be unable to travel for in-person interviews.³⁷

For example, in 2023 the Los Angeles County Public Defender's Office partnered with Amazon Web Services to develop a custom pipeline that is to reduce manual data entry from documents by up to 85%.³⁸ The new system makes all documents searchable and proactively alerts defenders to filings from law enforcement, courts and the district attorney's office. It ingests scanned PDFs, classifies useful pages, extracts agency-specific information, and allows staff to verify results through a user interface. The office's Chief Information Officer described: "When we did the first presentation to a group of attorneys, you should have seen their faces. It identified the summary sheet, got the relevant charges



and maximum sentences. It went and found witness information and created witness records. Everyone was leaning forward to the computer with their mouth open looking at this.”³⁹

In another example, the Maryland Office of the Public Defender implemented eDefender, a cloud-based case management system from Journal Technologies, Inc., that centralizes all case files into searchable folders accessible securely to defenders and paralegals working on the same case. The tool integrates with other third-party “justice partner systems,” APIs, and platforms (like Evidence.com) and centralizes critical tasks, reports, and communications.⁴⁰ Journal Technologies has similar platforms for prosecutors (eProsecutor) and courts (eCourts), and if all three are adopted in a given jurisdiction, this may facilitate equal access among all players in the jurisdiction.

NACDL spoke with a representative from the New Jersey Office of the Public Defender (OPD) for this white paper, and the conversation offered an instructive example of how a statewide public defender office can approach AI adoption incrementally using (1) internally developed and customized tools, (2) emerging tools with existing vendors like Westlaw, and (3) experimental tools with institutions researching AI use in certain industries. For example, the New Jersey OPD first started using an internal AI tool developed by New Jersey’s Office of Information Technology (NJOIT) specifically for state employees. This tool allows users to choose from models such as Claude and other large language models, that enable attorneys, investigators, or staff to quickly generate first drafts of informational memos, letters and policies, yet it is not used for legal briefs or any substantive legal research or work product. More recently, the office secured a six-to-nine-month trial of Westlaw’s CoCounsel product at no additional cost over their existing Westlaw subscription, which includes several AI features. Finally, the office is also partnering with Princeton University, NJOIT, and the New Jersey Innovation Authority to develop a “Brief Bank,” which would house all New Jersey public defender briefs on secured state infrastructure (through the state’s Office of Information Technology and with help from Amazon Web Services) and allow attorneys to generate first drafts informed by that institutional knowledge base.

Critically, each of the tools used is structured so that client queries do not train the underlying models, and have validated security controls to ensure client information remains inaccessible to anyone but the public defenders. The New Jersey OPD’s posture overall might be described as “curious and responsible”: Public Defender Jennifer Sellitti recognizes that meaningful AI adoption requires understanding not just what tools exist and the limitations of those tools, but also a willingness to engage with public and private stakeholders who are interested in developing new tools and testing them with public interest focused organizations. This posture is worth examining for public defense offices navigating similar questions about efficiency and ethics, and procurement and resource constraints.

2. *Evidence and Case Review; Legal Research and Writing*

Defenders are deploying AI to address the crushing workload that has long defined public defense, including everything from intake to evidence analysis to trial preparation. For example, the Kentucky Department of Public Advocacy uses JusticeText to “handle the surge of bodycam footage in cases,” and defenders reported that it “reduced time spent reviewing evidence by hours per case, allowing them to find key contradictions in police statements more efficiently.”⁴¹ The Colorado State Public Defender’s office has used Reduct, which allows defenders to highlight key clips, generate captions, and create court exhibits from bodycam footage and police interrogations.⁴² AI transcription and analysis tools have thus fundamentally changed how defenders manage overwhelming swaths of data and evidence. Defenders upload audio and video files, and AI generates searchable transcripts within minutes. What once required teams of staff listening to recordings in real time now happens automatically. Attorneys can search across hours of footage using keywords, jumping directly to moments where specific topics, names, or phrases appear. One Wisconsin defender used JusticeText to more efficiently reveal evidence of police misconduct that would have required substantial manual review, including instances where detectives destroyed evidence intentionally.⁴³

AI also performs substantive analysis of discovery materials. When police reports and witness statements are uploaded, AI generates organized summaries that highlight key facts such as who was involved, what evidence was collected, and where inconsistencies appear. It can identify which witnesses may be changing their stories and flags contradictions between different accounts of the same events.

In another example, the California Innocence Project relied on CoCounsel to narrowly focus on issues commonly seen in wrongful conviction cases, such as witness misidentification. It also helped comb through documents in databases to identify inconsistencies in witness statements or testimony. The managing attorney, Michael Semanchik, said: “Maybe we get to the point where it’s able to just take a look at an entire case, flag some of the wrongful conviction causes that we normally see and send it over for attorney review so we’re not spinning our wheels trying to find the good cases; the good cases come to us through AI.”⁴⁴

3. *Sentencing Advocacy*

AI also supports sentencing advocacy. Some platforms analyze federal sentencing data to generate statistical reports on likely sentencing outcomes based on historical trends.⁴⁵ By analyzing trends in judicial decisions, defenders can present data-backed arguments demonstrating disparities or inconsistencies in sentencing patterns for similarly situated defendants.⁴⁶



AI is increasingly present in the courtroom itself, and defenders should cautiously embrace using AI tools for things like sentencing memos and character statements. AI drafting tools can generate initial versions of sentencing memoranda by synthesizing case facts, identifying mitigating circumstances, and locating precedent, each requiring work that might otherwise consume hours of defender time. Some platforms — generally categorized as predictive analytics or legal analytics tools — analyze millions of federal sentencing outcomes to produce statistical reports showing how similarly-situated defendants have been sentenced, giving defenders empirical support for variance arguments. Although the specific products used are changing much like other general-use products, an example product is the Sentence Intelligence Report or Federal Sentence Predictor tools from SentencingStats.com. For character letters, AI can help family members and supporters structure their narratives effectively and address the elements judges find most persuasive while avoiding common pitfalls like disputing guilt or criticizing the prosecution.

Yet defenders should understand the risks of using AI for these uses, specifically. Hallucinations, the same error that led to sanctions against attorneys in widely publicized cases, and other errors in sentencing memos could devastate a client facing years of incarceration. Character letters present different dangers. When clients or their families use AI to draft support letters without defender review, the results often read as generic and formulaic, lacking the specific stories and concrete details that judges find persuasive. Worse, AI-generated letters from multiple supporters may share suspiciously similar phrasing, signaling to the court that the letters are manufactured rather than genuine expressions of support. Judges are increasingly adept at recognizing AI-generated content, and the perception that a defendant's mitigation package is inauthentic can undermine the entire sentencing presentation.

Some judges have approved using AI for sentencing and/or character statements in much more dramatic ways. In 2025, the family of a slain road-rage victim used AI to generate a video of the victim speaking to the court. It was considered the “first time the technology has been used in the U.S. to create an impact statement read by an AI rendering of the deceased victim.”⁴⁷ Nonetheless, a defender's obligation to scrutinize and object to outrageous uses of AI — such as the creation of wholly fabricated evidence and not simply its use for exhibit creation, generally — extends to all materials submitted to the court, as discussed further in Part II.



C. The Defense Lawyer's Duty to Use AI

1. *Adversarial Reality*

The criminal legal system operates on adversarial principles. Prosecutors and defense lawyers each advocate zealously for their respective positions, and the truth-seeking function depends on both sides presenting their strongest cases. When one side gains significant technological advantages, this balance shifts.

The Vanderbilt Project on Prosecution Policy documents current prosecutorial uses including probabilistic genotyping for DNA analysis, extraction and analysis of data from personal electronic devices, enhancement of digital evidence, evidence management and organization, automated transcription and translation, AI-assisted redactions, and pattern identification across cases.⁴⁸

Defense lawyers who are unable — due to a lack of resources or support or use restrictions from their firm or office — or unwilling to engage with AI risk falling behind. If prosecutors can delegate workloads more efficiently through AI-assisted platforms, they may better manage staff and deadlines. If prosecutors can process a cell phone extraction containing thousands of text messages in hours while defenders take weeks to review the same material manually, critical evidence may go unnoticed. And if prosecutors use AI to identify connections between cases and defendants while defenders lack similar capabilities, the informational asymmetry compounds existing resource disparities.

This does not mean defenders should develop or invest in a new AI solution without critically reviewing its capabilities and security. But it does mean defenders have an obligation — rooted in their duty of competence under Model Rule 1.1 and their duty of diligence under Model Rule 1.3 — to understand what AI tools exist, how prosecutors use them, when to object, and whether comparable tools might benefit their own clients. The ABA's comment to Model Rule 1.3 specifically states that lawyers must act with *zeal in advocacy upon the client's behalf* (emphasis added). Of course, the ability to mount a zealous defense is contingent upon having access to the same or similar tools that prosecutors access.

2. *Ethical Use and Evidence*

In 2023, the Advisory Committee on Evidence Rules, for example, amended Federal Rule of Evidence 702 to require that “expert testimony may not be admitted unless the proponent demonstrates to the court that it is more likely than not that the proffered testimony meets the admissibility requirements set forth in the rule.”⁴⁹ Likewise, the Advisory Committee has proposed a new Federal Rule of Evidence 707 to regulate the admissibility of machine-generated evidence introduced without expert testimony.⁵⁰



The proposed rule seeks to address issues raised by AI use in courts and has generated extensive discussion and comments from judges, lawyers, and other stakeholders. The NACDL submitted comments to the Advisory Committee on February 16, 2026.⁵¹

Defense lawyers must recognize that AI creates new risks of fabricated or manipulated evidence generated by clients, witnesses, and prosecutors, and it also presents a temptation for defense lawyers to follow suit. Defense lawyers should prepare to scrutinize and confront fabricated evidence generated by any and all parties in a case. AI image generators can create photographs of events that never occurred. The use of deepfakes in the courtroom has prompted responses from scholars, practitioners, bar associations, and courts alike. Professor Rebecca Delfino argued in 2023 that “no evidentiary procedure explicitly governs the presentation of deepfake evidence in court,” and was the first to propose a new Rule 901(c) in the Federal Rules of Evidence “reflecting a novel reallocation of fact-determining responsibilities from the jury to the judge, treating the question of deepfake authenticity as one for the court to decide as an expanded gatekeeping function under the Rules.”⁵² Others have more recently suggested a new Rule 901(b)(11) requiring courts to go beyond a witness statement to enable the accused party to request a hearing to require the proponent to provide corroborating sources,⁵³ and a new Rule 901(c) that holds if the challenging party successfully presents evidence that challenges the authenticity of evidence as more likely than not to be a deepfake, the proponent must show that its probative value outweighs its prejudicial effect on the party challenging the evidence.⁵⁴

The risks of manipulated evidence run in multiple directions. Prosecutors might intentionally use or unknowingly rely on AI-manipulated evidence. Witnesses might submit AI-generated materials. Defenders themselves might inadvertently present fabricated content if they fail to verify AI outputs. Defenders should approach evidence with heightened skepticism in the AI era and consider the feasibility of bringing in forensic experts to inspect both the prosecutor’s materials and audit their own materials.

3. *Standard Setting*

Defenders have an opportunity — and perhaps an obligation — to lead in establishing ethical standards for AI use in the criminal legal system. By developing and adhering to rigorous ethical frameworks, defenders can model responsible AI practices and create standards against which prosecutorial AI use can be measured. Moreover, knowledgeable and trained defenders can become advocates for parity in ethical use of AI among prosecutors. Once defenders better understand how practical uses of AI clash with ethical obligations, defenders can better articulate that conflict to key stakeholders in the public and private sector.



This standard-setting function serves multiple purposes. Internally, clear ethical guidelines protect defenders from inadvertent misconduct and help offices maintain consistent practices. Externally, documented defense standards provide a basis for challenging prosecutorial AI use that fails to meet comparable ethical benchmarks.

Defenders who understand AI well can more effectively litigate against improper prosecutorial uses. Challenging the reliability of probabilistic genotyping software, for instance, requires understanding how such algorithms function and where they might fail. Contesting facial recognition identifications demands knowledge of how those systems operate, their error rates, and their documented racial disparities. Defenders who remain ignorant of AI technology cannot effectively advocate for their clients when that technology affects the prosecution's case.

The NACDL's Task Force on Artificial Intelligence represents one effort to develop collective defense standards. Individual offices and bar associations can contribute by sharing policies, training materials, and litigation strategies. Through coordinated effort, the defense bar can ensure that AI advances ethical justice rather than undermines it.



PART II: AI Ethics and the Model Rules of Professional Conduct

A. Helpful Resources to Guide AI Use

Lawyers who adopt AI tools remain bound by the same professional responsibility obligations that have always governed the practice of law and the attorney–client relationship. The Model Rules of Professional Conduct provide the authoritative framework, and bar opinions and standing judicial orders help defenders apply that framework to the practice of law. But the Model Rules were drafted long before generative AI emerged as a practical tool for legal work. No rule mentions artificial intelligence by name. Defenders must therefore interpret existing obligations — competence, confidentiality, and candor, in particular — in light of technological capabilities the drafters never anticipated. This interpretive challenge is complicated by AI’s more problematic characteristics. Its capacity to generate plausible but fabricated information, its reliance on third-party service and monitoring that may compromise confidentiality, and its tendency to affirm the biases of its creators and its users pose ethical challenges for defenders. The sections that follow identify duties of competence, confidentiality, and candor as the fundamental obligations for defenders but acknowledge other obligations that may be found in other Model Rules and are impacted by AI use. The sections examine how each core professional obligation applies when defenders incorporate AI into their practice.

First, however, defenders may find these resources helpful for compilation of existing judicial orders and/or local rules and bar opinions regarding the use of AI in the profession:



1. ***Ropes & Gray Standing Orders and Local Rules on the Use of AI***⁵⁵

The interactive map includes 378 standing orders, local rules, or other decisions related to the use of AI in federal and state courts. The default map shows whether the majority of AI-related orders in a given state encompass any AI or only generative AI, and the remaining selections are color-coded to signify different types of orders, local rules, or decisions (e.g., “requires disclosure,” “suggests cautious use of AI,” “prohibits use of AI”).

2. ***American Bar Association Formal Opinion 512***⁵⁶

Formal Opinion 512 describes ethical considerations for lawyers using AI, including duties under the ABA Model Rules of Professional Conduct. It focuses on competence, confidentiality, communication with clients, maintaining candor toward tribunals, supervisory responsibilities, and charging reasonable fees.

3. ***American Bar Association Task Force on Law and Artificial Intelligence Year 2 Report on the Impact of AI and the Practice of Law***⁵⁷

The report examines AI’s transformative impact across the legal profession — including law practice, courts, access to justice, legal education, and governance — while providing practical guidance and ethical frameworks for lawyers and judges to responsibly integrate AI tools into their work. It emphasizes that AI offers significant opportunities to enhance efficiency and expand access to justice, but encourages the profession to maintain human oversight, address risks like hallucinations and deepfakes, and ensure ethical AI use.

4. ***Bloomberg’s In Focus: Artificial Intelligence (AI)***⁵⁸ ***and JUSTIA’s 50-State Survey on AI and Attorney Ethics Rules***⁵⁹

Both Bloomberg and JUSTIA offer resources for tracking state ethics opinions and judicial orders. Defenders may use these as a quick reference for various state bar association opinions but are encouraged to visit their respective state’s bar association to confirm any opinions or guidance remain current.



B. Balancing Duties of Competence, Confidentiality, and Candor

1. *Rule 1.1 – Competence*

Model Rule 1.1 requires that lawyers provide “competent representation,” which demands “the legal knowledge, skill, thoroughness and preparation reasonably necessary for the representation.” Comment 8 to Rule 1.1 elaborates that maintaining competence requires lawyers to “keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology.”

This technological competence obligation has direct implications for AI use. Defenders must understand how AI tools function before using them in client matters. This understanding need not reach the level of a computer scientist or software developer, but it must be sufficient to recognize the technology’s capabilities and limitations. A defender who uses an AI research tool without understanding that it might generate fabricated citations fails to meet the competence standard. Conversely, a defender with less experience using and understanding AI may fail to notice when a prosecutor intentionally or unknowingly misuses it.

The competence obligation also connects to what some commentators call “calibration,” or the appropriate level of AI use for given clients or circumstances.⁶⁰ As Simshaw has observed, effective legal AI requires “calibration” that accounts for the specific consumers, legal issues, and underlying processes involved in each case. Defenders should not focus on “maximizing AI use,” but instead “making more informed decision about what AI should, and should not, be used for.”⁶¹ Simshaw notes that legal service providers face increased pressure to adopt AI, and the calibration is essential for all lawyers.⁶² Significant resource disparities exist between prosecution and defense in most jurisdictions. Prosecutors’ offices typically have larger budgets, more staff, and greater access to technology than public defenders or private counsel from smaller and/or solo firms. AI may widen these disparities if prosecutors adopt sophisticated AI tools while defenders lack resources for comparable technology. Moreover, the biggest companies offering AI tools to prosecutors often build their tools with prosecutors in mind and later market them to defenders as a second (less valued) market. Thus, some defenders may not trust the systems that were designed for prosecutors and may also be concerned that their use of the same exact system may interfere with their ability to challenge the prosecution’s use.

These disparities raise fairness and competence concerns. If competence under Model Rule 1.1 requires technological sophistication that under-resourced defenders cannot afford, the ethical framework may demand what defenders cannot deliver. Clients of well-funded private counsel might receive AI-enhanced representation while indigent defendants receive manual-only defense — creating a two-



tiered system based on wealth. Addressing these disparities requires systemic responses beyond individual defenders' control. Adequate funding for public defense, resource-sharing arrangements among defense offices, and development of low-cost or free AI tools for defenders could help.

On a related note, over-reliance on AI poses risks to defenders' professional development and advocacy capabilities. When AI drafts motions, generates arguments, and identifies authorities, defenders may not develop the analytical skills that effective advocacy requires. The convenience of AI assistance can become a crutch.

This concern applies with particular force to newer attorneys. Early-career or junior defenders who rely heavily on AI may never develop foundational skills in legal research, writing, analysis, and oral advocacy. When circumstances require them to perform without AI assistance — in court, during meetings, or when technology fails — they may find themselves unprepared. Defenders and supervisors should structure AI use to preserve skill development. Junior attorneys should perform tasks manually before learning AI-assisted approaches. AI should supplement rather than replace human analysis. Regular practice without AI assistance maintains skills that might otherwise atrophy. Moreover, AI's effectiveness — at least for junior attorneys — is predicated on the junior attorney's familiarity with underlying legal principles and ability to effectively prompt the AI tool. For example, if a junior attorney asks an AI tool to assist with generating *voir dire* questions, but the junior attorney is unfamiliar with deselection, the tool will likely generate questions that actually draw out the most favorable jurors — something that is not currently a best practice for the *voir dire* process.

AI use in investigation, discovery, and evidentiary presentations raises distinct concerns about reliability and reproduction. When AI processes evidence — transcribing audio, extracting data from devices, or identifying relevant documents — the accuracy of that processing directly affects case outcomes. Defenders should understand the error rates and limitations of AI tools used in evidence processing. AI transcription, for instance, may struggle with accents, technical terminology, or poor audio quality. AI document classification may miss relevant materials or flag irrelevant ones. These errors can cause defenders to overlook helpful evidence or waste time on unhelpful materials.

When prosecutors use AI to process evidence, defenders should probe the reliability of that processing. What tools did the prosecution use? How were they trained on those tools? How were the tools trained? What are the known error rates? Was the output verified by human review? Were there quality control procedures? All of these questions can help evaluate whether both prosecutors and the tools they use are properly trained on identifying Brady material, biased charging recommendations (i.e., the prosecutor or the tool looks for evidence that fits the elements of a more serious charge), among other procedural issues. These questions parallel traditional challenges to forensic evidence and deserve similar attention when AI plays a role.



Moreover, AI systems can perpetuate and amplify biases present in their training data or design, especially when used in the context of the criminal legal system.⁶³ When AI is trained on historical data reflecting discriminatory patterns, the AI may reproduce those patterns in its outputs. This concern applies across AI applications, from facial recognition systems that perform worse on darker-skinned faces to risk assessment algorithms that may disadvantage defendants from communities of color.

Professor Brandon Garrett and Professor Cynthia Rudin have documented bias concerns in criminal justice AI. They note that “criminal justice data is often noisy, highly selected and incomplete, and full of errors.”⁶⁴ When AI systems train on this flawed data, they may encode existing biases rather than correcting them. A risk assessment tool trained on historical sentencing data, for instance, might perpetuate racial disparities present in that data.

Defenders should be alert to bias in AI tools they use and in AI tools prosecutors employ. When AI outputs seem to disadvantage clients based on protected characteristics, defenders should investigate whether bias might explain the results. Challenging biased AI evidence requires understanding how bias enters AI systems and how to demonstrate its effects.

When a defender lacks AI competence, at least three options exist. First, the defender can decline matters requiring AI assistance. Second, the defender can invest the time necessary to develop adequate AI knowledge and skills. Third, the defender can associate with another lawyer who possesses the necessary technological competence. A defender simply cannot proceed with AI-assisted representation while remaining ignorant of how the technology works and where it might fail.

2. Rule 1.6 – Confidentiality of Information

Model Rule 1.6 prohibits lawyers from revealing “information relating to the representation of a client” without informed consent or implied authorization to carry out the representation. The rule further requires lawyers to “make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.” Defenders should note that this Model Rule specifically addresses the ethical duty to keep all information about the client confidential. Attorney–client privilege is a separate evidence rule, although privileged information is considered confidential.

AI systems pose significant confidentiality risks. Depending on the system’s architecture, the information might be stored on remote servers, used to train future AI models, or become accessible to the vendor’s employees or contractors. Even if the AI provider promises confidentiality, data breaches remain possible, and the lawyer has limited ability to verify or enforce the provider’s security practices.



Moreover, there are complex contractual risks of enterprise AI tools. For example, when a firm subscribes to an AI-powered legal research platform offered by a vendor such as Thomson Reuters or RELX, it enters into a direct contractual relationship with that vendor. Yet vendors building on top of third-party models routinely incorporate, by reference or by flow-down clauses, the acceptable use policies and terms of service of their upstream model providers such as OpenAI, Anthropic, Google, or others. The result is that a law firm's subscription agreement may silently bind users to a set of third-party obligations that are never surfaced during contract negotiation, are subject to unilateral amendment by the upstream provider, and may impose restrictions on data inputs, outputs, or use cases that are difficult to reconcile with ordinary legal practice.

The confidentiality analysis under Rule 1.6 addresses the broad category of information “relating to the representation.” Attorney-client privilege presents a narrower but equally important concern. Privileged communications between lawyer and client receive special protection, and defenders must take care that AI use does not waive or compromise privilege.

“There remains a plausible argument that uploading . . . privileged client information would destroy confidentiality and privilege.”⁶⁵ Entering privileged communications into an open, insecure AI system might constitute disclosure to the AI vendor and its personnel. Defenders should exercise particular caution with privileged materials. Even when using enterprise AI systems with strong confidentiality protections, defenders should consider whether processing privileged communications through AI serves a legitimate purpose that justifies any waiver risk. When doubt exists, defenders should err on the side of keeping privileged materials out of AI systems.

Information entered into most free versions of consumer-facing tools may be incorporated into training data or otherwise retained by the provider. The terms of service for many consumer AI products explicitly permit such uses, and even products that currently do not permit such use may change terms and conditions without the user's knowledge. Defenders who enter client facts, case theories, or discovery materials into unsecure AI systems risk violating Rule 1.6 by allowing confidential information to leave their control. This is perhaps most problematic when defenders use their own personal devices and personal versions of an LLM — like using ChatGPT on a personal laptop or its mobile app on a personal cell phone — to do legal work.

Discovery materials present particular challenges. Protective orders often restrict how parties may use and store produced materials. Entering protected discovery into an AI system might violate the protective order if the system's data practices conflict with the order's requirements. Defenders should review protective orders carefully before processing discovery through AI tools and seek modifications if necessary to permit appropriate AI-assisted review.

3. *Rule 3.3 – Candor Toward the Tribunal*

Model Rule 3.3 prohibits lawyers from making false statements of fact or law to a tribunal and requires lawyers to correct false statements previously made. This duty of candor applies with full force to AI-assisted work product.

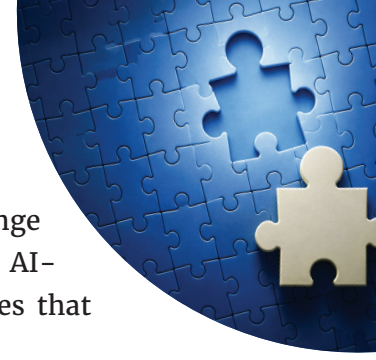
The most prominent AI-related candor failures have involved fabricated legal citations or propositions later noticed by a judge or opposing counsel. Human verification against authoritative sources remains essential. Think of this intervention as a “lawyer-in-the-loop,” a version of “human-in-the-loop” oversight for ensuring AI outputs are accurate.

Many courts have responded to AI-related candor concerns by requiring disclosure of AI use. Federal courts in nearly two dozen states now have standing orders requiring disclosure and/or verification of AI use. Some simply require disclosure of its use in filings or other documents, but others, like Judge Brantley Starr of the Northern District of Texas, have adopted standing orders requiring attorneys to certify that *no* portion of any filing was drafted by generative AI without human review, or that any AI-generated content was checked for accuracy by a human. Defenders must distinguish between AI outputs and their own human analysis. When AI generates a draft, the defender must review that draft, verify any factual or legal assertions, and take responsibility for the final work product (and take responsibility for the work product of junior defenders or staff who may have used AI, too). If AI contributes to a filing, defenders should be prepared to disclose that contribution if required by court order or if candor otherwise demands it.

AI can generate convincing fake audio, video, and images — so-called “deepfakes” that may be difficult or impossible to distinguish from authentic recordings. This capability creates risks that fabricated evidence might enter criminal proceedings, whether through prosecutorial misconduct, witness fabrication, or third-party manipulation. Defenders must approach audiovisual evidence with heightened scrutiny. When recordings seem inconsistent with other evidence, when provenance is unclear, or when the evidence seems too convenient, defenders should consider whether manipulation might explain the anomalies. Expert analysis may be necessary to detect sophisticated deepfakes.

Defenders should also ensure their own evidence is authentic. When clients or witnesses provide recordings, letters, or other materials, defenders should verify authenticity before relying on them. Submitting fabricated evidence — even unknowingly — can devastate a client’s case and the defender’s credibility.

Beyond deepfakes, AI raises broader questions about evidence authentication. Traditional authentication requires testimony that evidence is what it purports to be. When AI generates or processes evidence,



authentication becomes more complex. Defenders should also prepare to challenge authentication of prosecution evidence involving AI. When prosecutors rely on AI-processed materials, defenders can probe whether proper authentication establishes that the AI output accurately represents the underlying evidence.

Defenders frequently help clients prepare allocution and other materials for sentencing or other proceedings. AI can assist with this process — generating drafts, suggesting content, or polishing language — but such assistance raises ethical concerns. Character letters present a common example. Courts and opposing counsel may scrutinize such materials for signs of AI generation, and defenders should advise clients about these concerns.

Authenticity matters for character materials. Courts rely on character letters as genuine expressions of the writer’s views. Candor obligations apply here as well. If a court or opposing party asks about the preparation of character materials, defenders must answer truthfully. Representing AI-generated content as the client’s or a third party’s original work may violate the duty of candor. Defenders should disclose AI assistance in character materials when authenticity might be questioned and should ensure that any person signing AI-assisted materials has meaningfully reviewed and adopted the content as their own. When doubt exists about appropriate disclosure, defenders should err toward transparency.

Courts have responded to AI misuse with sanctions and prospective restrictions. Courts in multiple jurisdictions have adopted standing orders requiring AI disclosure or certification. Defenders should know the AI-related rules and orders in their jurisdictions. Requirements vary significantly. As noted above, some courts require disclosure only of generative AI used in drafting, while others sweep more broadly to include any AI assistance. Some courts require affirmative certification that AI content was verified, while others simply require disclosure. Noncompliance can result in sanctions, rejected filings, or other adverse consequences.

4. *Rule 5.3 – Supervision; Responsibilities Regarding Nonlawyer Assistance*

Model Rule 5.3 addresses lawyers’ responsibilities when using nonlawyer assistants. Partners and supervising lawyers must ensure that nonlawyer conduct remains “compatible with the professional obligations of the lawyer.” The rule requires law firms to have “measures giving reasonable assurance” that nonlawyer assistants’ conduct comports with the lawyer’s ethical obligations.

AI vendors constitute a form of nonlawyer assistance that lawyers must supervise to ensure ethical compliance. This supervisory obligation has several implications. Defenders must evaluate AI vendors’ practices before using their services, just as they would evaluate a contract paralegal or document review company. Defenders must monitor AI outputs for accuracy and ethical compliance. And defenders must



maintain sufficient control over AI-assisted processes to intervene when problems arise. The supervisory obligation also affects how defenders structure AI use within their offices. Senior defenders and supervisors bear responsibility for ensuring that junior defenders and staff use AI appropriately. This may require establishing office-wide AI policies, providing training on approved tools and practices, and creating review procedures for AI-assisted work product.

5. *Rule 1.3 – Diligence and Rule 1.4 – Communication*

Model Rule 1.3 requires lawyers to “act with reasonable diligence and promptness in representing a client.” Model Rule 1.4 requires lawyers to keep clients “reasonably informed about the status of the matter” and to “explain a matter to the extent reasonably necessary to permit the client to make informed decisions.”

AI can support diligence by enabling defenders to work more efficiently, process information more quickly, and respond to developments more promptly. When AI helps a defender review discovery faster or draft motions more quickly, the technology serves the diligence obligation by accelerating representation.

However, although zealous advocacy requires using available and appropriate AI tools as a supplement to human qualities, AI cannot wholly substitute for zealous advocacy. The human qualities that effective defense requires center on contextual judgment — spontaneity, empathy, creativity, and morality — and remain largely beyond AI’s current capabilities.⁶⁶ A defender who uses AI to generate a competent-looking brief but fails to identify the strongest arguments for the client based on known perceptions or biases of a particular judge, for example, or a defender who relies on AI efficiency to handle an excessive caseload without meaningful engagement, likely fails the diligence standard.

The communication obligation may require defenders to inform clients about AI use in certain circumstances. When AI materially affects the representation — such as when AI tools will process sensitive client information or when AI assistance substantially shapes case strategy — clients may need this information to make informed decisions. The scope of required disclosure will vary based on the AI’s role, the sensitivity of the information involved, and the client’s expressed preferences.

6. *Rule 2.1 – Independent Professional Judgment; Environmental Responsibility*

Model Rule 2.1 provides that lawyers “shall exercise independent professional judgment and render candid advice.” Beyond traditional ethical considerations, defenders should consider the environmental impact of AI use. Large language models require enormous computational resources. Although the full extent of the environmental impact is contested among scholars and experts, it is clear that AI models can consume consequential amounts of electricity and generate significant carbon emissions. Ongoing use of AI systems continues to consume energy for each query processed.

NACDL has committed to environmental sustainability. Defenders who share this commitment should consider AI’s environmental footprint when selecting and deploying tools. Where possible, defenders should prefer smaller, task-specific models over massive general-purpose systems; select providers that use renewable energy; avoid unnecessary or duplicative AI queries; and advocate for vendors to improve their environmental practices. These considerations may not necessarily override other factors in AI selection, but they deserve weight alongside cost, capability, and security in the overall evaluation.



PART III:

Practical Guidance For Defense Lawyers

Parts I and II of this white paper established what generative AI is, how it functions, and the ethical framework governing its use. This Part translates that foundation into practical guidance that defenders can implement immediately. It addresses tool selection, risk assessment, verification protocols, and the development of comprehensive AI use policies.

A. Practical Guidance for Defense Lawyers

1. *Identification of Enterprise v. Personal Tools*

Defender offices and firms have an imperative to provide at least some AI tools to defenders. Both defenders and staff will use AI regardless of whether their offices have adopted formal policies because the tools are too accessible and the workload pressures are too acute to expect otherwise. AI has entered and will likely become a focal point of defenders' day-to-day practice. The question is whether it will permeate through sanctioned systems and tools with appropriate safeguards or through ad hoc individual use that risks confidentiality breaches and ethical violations. Providing vetted, secure AI tools is a matter of client protection. Moreover, a growing body of commentary suggests that the failure to leverage AI may itself raise competence concerns under professional responsibility rules, as attorneys who forgo efficiency-enhancing technologies may be unable to provide representation comparable to that of AI-assisted prosecutors. Defenders must distinguish between AI tools available through organizational subscriptions and those accessible for personal use. This distinction carries significant implications for confidentiality, security, and appropriate use cases.



Enterprise-level AI tools are deployed through organizational accounts with negotiated terms of service, dedicated security infrastructure, and administrative oversight. These systems typically operate in closed environments where data remains within the organization's control. Examples include Microsoft 365 Copilot deployed through a firm's existing Microsoft license, legal-specific platforms like Harvey or CoCounsel with enterprise agreements, and custom AI implementations hosted on private servers or secure cloud instances. Enterprise tools generally offer stronger confidentiality protections, audit capabilities, and compliance features⁶⁷ — but they require significant financial investment and technical infrastructure.

Personal AI tools are consumer-facing products accessible to individuals through free or low-cost subscriptions. The free versions of ChatGPT, Claude, and Gemini fall into this category. While these tools provide powerful capabilities, they typically operate as open systems where user inputs may be retained, analyzed, or incorporated into future training. Terms of service for personal accounts rarely include the confidentiality guarantees that legal work requires.

The dangers of personal tool use for client matters cannot be overstated. As discussed in Part II, entering confidential client information into open systems risks breaching Rule 1.6 obligations. Even seemingly innocuous queries may reveal case strategies, client identities, or privileged communications. A defender who uses a personal AI tool to quickly review their state's rules on expert testimony during a break in trial, for example, may inadvertently disclose client facts to OpenAI and its contractors, potentially waiving privilege and violating confidentiality duties.

Defenders should also understand that different tools excel at different tasks. Large language models like ChatGPT and Claude perform well for text generation, summarization, and analysis. Image generation tools like DALL-E and Midjourney can create visual aids for trial presentations. Specialized legal AI platforms may offer enhanced capabilities for document review, citation checking, and case law research. Matching the right tool to the task improves both efficiency and output quality.

For defenders lacking access to enterprise tools, some organizations offer subsidized access, shared subscriptions, or resource pools for certain legal-specific AI tools.⁶⁸ And in some circumstances, careful use of consumer tools for non-confidential tasks — such as researching general legal principles or drafting non-case-specific training materials — may be appropriate, provided defenders rigorously avoid entering any information relating to client matters.

2. *Vendor Selection*

Choosing appropriate AI vendors requires evaluation across multiple dimensions. Defenders must balance functionality against security, cost against capability, and convenience against compliance.



The first consideration is system architecture, or cloud-based versus “on-prem” deployment. Cloud-based AI services process data on the vendor’s servers, typically in shared infrastructure serving multiple customers. This approach reduces costs and technical complexity but requires trust in the vendor’s security practices. On-premises deployment, by contrast, runs AI models on hardware the organization controls, keeping data entirely within the defender’s environment. On-premises solutions offer maximum data control but require significant technical expertise and capital investment. Currently, a hybrid approach or carefully vetted cloud deployment will prove more practical than full on-premises infrastructure for most defenders. But even a hybrid approach requires careful attention to what file systems are (and whether the public internet is) accessible to the model. Systems with any on-premises access should only have access to those specific files or folders in the file system that have been affirmatively vetted for concerns around confidentiality, privilege, and relevance. Although the RAG-related privacy concerns addressed in the literature capture the “data leaving the building” problem (i.e., the risk that confidential information is transmitted to and processed by external systems), there is also a “data in the building” problem that arises when local or on-premises deployment concentrates sensitive information within the firm’s own infrastructure, creating new vectors for insider threat, unauthorized access, and inadequate data governance.

The landscape continues evolving. Research suggests that some organizations are moving away from “cloud-first to strategic hybrid: cloud for elasticity, on-premises for consistency, and edge for immediacy.”⁶⁹ What seems prohibitively expensive today may become feasible within years. Defenders should monitor these developments and reassess their options periodically.

Security guarantees warrant careful scrutiny. Defenders should examine (1) whether the vendor encrypts data in transit and at rest (i.e., information actively moving between systems — such as from an email on a laptop to a cloud-based storage system — versus information stored in a fixed location, like on a server, hard drive, or cloud database), (2) who can access stored data within the vendor’s organization, (3) what third-party service terms, acceptable use policies, data retention policies, and/or any other policies apply, (4) whether the vendor uses customer data for model training, (5) what audit and logging capabilities exist, and (6) how the vendor handles security incidents and breach notifications. Vendors marketing to legal customers should provide clear, written answers to these questions. Vague assurances or refusals to document security practices suggest inadequate protection.

Business and contractual terms matter as well. When practical, defenders with the necessary resources should negotiate for terms prohibiting vendor use of submitted data for training or other purposes, requiring prompt data deletion upon request, indemnifying the defender against security breaches caused by vendor negligence, and specifying the jurisdiction and venue for any disputes. Larger organizations may have leverage to negotiate custom terms; smaller offices and solo practitioners may need to accept standard terms but should review them carefully before committing.



The vendor evaluation process differs between established companies and emerging startups. Established providers like Microsoft, Google, and Thomson Reuters bring reputational stakes and institutional resources to security and compliance. However, their products may be more generic and their terms less negotiable. AI startups may offer more tailored solutions and greater flexibility but carry higher risk of business failure, security immaturity, or inadequate resources to honor contractual commitments. Defenders should weigh these tradeoffs carefully and consider requiring startups to demonstrate security certifications, insurance coverage, or escrow arrangements protecting against business discontinuity.

3. *Low-Risk vs. High-Risk Tasks*

Not all AI applications carry equal risk. Defenders should calibrate their AI use based on the sensitivity of the task and the consequences of error.

Low-risk applications involve tasks where AI errors are unlikely to harm clients and where confidential information need not be disclosed to the AI system. Examples include summarizing general legal principles not specific to any case, drafting administrative documents like office policies or internal memoranda, generating templates for routine correspondence, creating educational materials for staff training, and brainstorming approaches to common legal issues. These tasks can often be performed with consumer-grade AI tools, though enterprise systems remain preferable where available. Human review of the output can eliminate the risk associated.

Medium-risk applications involve work product that affects client matters but can be thoroughly verified before use. Drafting initial versions of motions or briefs falls into this category. AI can provide a starting point that the defender then reviews, revises, and validates. Research assistance, such as identifying potentially relevant authorities or summarizing long documents, similarly offers efficiency gains while preserving human oversight. The Vanderbilt Project on Prosecution Policy framework identifies “litigation content” as one category where AI has shown promise for prosecutors, including to draft jury selection questions, opening statements, direct examination questions, motions, and appellate arguments.⁷⁰ Defenders engaging in medium-risk applications must verify AI outputs rigorously, as discussed below, and should prefer enterprise systems with appropriate confidentiality protections.

High-risk applications include, but are not limited to, tasks involving the use of privileged and/or confidential information and where AI errors could directly harm clients or where the work requires judgment, strategy, or analysis that AI cannot reliably provide. Case-specific factual analysis, strategic decision-making, client counseling, and courtroom advocacy fall into this category. AI should support rather than supplant human judgment in these areas.



The Vanderbilt white paper states that “decision-making in charging, sentence recommendations, and other outcomes” could be “the most consequential use of AI powered systems in prosecution.”⁷¹ The white paper itself notes that “there is a major risk in lack of validation, and by extension, the potential for perpetuating biases” should AI replace prosecutorial judgment about case screening, charging, bail recommendations, plea negotiations, sentencing recommendations, and post-conviction assessments.⁷² It is essential for defenders to anticipate even minimal AI use by prosecutors for these high-risk tasks, and defenders should likewise consider their use of AI for these tasks as high-risk. Of course, there are some lower-risk tasks contained in those high-risk tasks that prosecutors are already using AI for, and defenders should do the same. Those include using AI (LLMs or legal-specific tools) to summarize case files. One recommendation for using AI in these specific instances would be to have an “independent researcher review a random sample of decisions and determine whether [] disparities exist in the outcomes produced.”⁷³

4. *Confidentiality Protocols*

Part II’s discussion of Rule 1.6 established the confidentiality framework. This section translates that framework into protocols.

The basic rule is that defenders should use only secure, approved systems for work involving confidential information. What constitutes “secure” and “approved” requires office-level determination based on available resources and risk tolerance, but certain minimum standards apply universally. Any system handling client data should operate in a closed environment, either as a cloud-based instance or an on-prem environment. The system should encrypt data in transit and at rest. Access should be limited to authorized personnel with legitimate need. And the vendor should contractually commit to confidentiality protections at least as strong as the defender’s own obligations.

Defenders should establish clear categories of information and corresponding protocols for other types of non-confidential information. Some information — such as general legal research unconnected to any client — may appropriately be processed through less secure systems, including open-sourced and open environment systems like a free model of ChatGPT or Claude. Other information — including case facts, client communications, and discovery materials — should never enter any AI system without enterprise-grade security. The most sensitive information — details that could endanger witnesses, privileged strategy discussions, or materials subject to restrictive protective orders — may warrant exclusion from AI systems entirely, even secure ones.

Practical implementation requires training and enforcement. Staff must understand which systems are approved for which purposes. Workflow design should channel work toward appropriate tools and make inappropriate uses inconvenient. Regular audits can verify compliance. And a culture that supports reviewing each other’s AI practices also reinforces good habits.

5. *Verification Expectations*

AI outputs require human verification before use in client matters. This white paper repeatedly emphasizes this principle.

The verification challenge varies by output type. Factual assertions require checking against authoritative sources. For example, if a defender uses a transcription tool, the defender should review the original audio or video against the AI-generated transcript. Likewise, if AI states that a particular statute applies or that a case holds a particular proposition, the defender must confirm these assertions by reading the statute or the case. The hallucination problem documented in Part II means that even confident-sounding AI assertions may be entirely fabricated, propositions coming from actual cases may still be incorrect, and even legal-specific tools like CoCounsel and Lexis AI can hallucinate. Every citation in AI-assisted work product should be verified — and not by asking the AI whether its citations are real, which simply invites the AI to affirm its own errors.

Legal analysis requires a different kind of verification. Defenders must evaluate whether AI-generated arguments are logically sound, whether they accurately characterize authorities, whether they address the strongest counterarguments, and whether they serve the client's interests. This evaluation requires legal judgment that AI cannot provide about its own work.

Of course, practical constraints impact verification expectations for all defenders. A solo practitioner with overwhelming caseloads cannot verify every sentence with the same rigor as an associate at a well-resourced firm. But this constraint argues for using AI judiciously rather than bypassing verification entirely. Defenders should focus AI assistance on tasks where verification is feasible and avoid using AI for tasks where verification would be impractical. When time permits only spot-checking rather than comprehensive verification, defenders should acknowledge this limitation to themselves and consider whether the risk is acceptable.

The white paper uses “Verification Expectations” rather than “Verification Standards” intentionally. The defender's verification obligation likely changes depending on the particular type of task or output, and reasonable defenders may disagree about what verification a particular task or output requires. What matters is that defenders approach verification thoughtfully rather than treating AI outputs as presumptively reliable.

6. *Client Communication*

Model Rule 1.4 requires lawyers to explain matters to clients as needed for informed decision-making. AI use may trigger this obligation in several circumstances.

When AI will process sensitive client information, clients may have legitimate interests in understanding how their information will be handled. Some clients may have strong preferences about data security or may work in industries with regulatory constraints on information handling. One interpretation of Model Rule 1.6 could be that clients must provide informed, explicit consent before a defender may use confidential, privileged, and/or any personal information in an AI tool. Another interpretation could be that defenders should merely explain how a defender intends to use AI in the case. Explaining that the defender uses enterprise AI systems with specified security features — or, conversely, that the defender has chosen not to use AI for their matter — respects client autonomy. The client may also request that the defender not use AI at all during their representation. A defender should be prepared to explain how foregoing AI could be beneficial or detrimental to the case, especially if opposing counsel is likely to use it.

When AI substantially shapes case strategy or work product, transparency may be appropriate. If AI analysis identifies a promising defense theory or AI-generated drafting forms the backbone of a brief, some clients may wish to know. The scope of required disclosure will depend on the AI's role, the client's sophistication, and any preferences the client has expressed.

Defenders do not necessarily need to obtain consent for every AI use, and there is currently no ABA opinion or other guidance demanding it. Routine applications that do not materially affect representation (e.g., using AI to format documents or schedule appointments) likely fall within the implied authorization that accompanies the representation. But when AI plays a significant role in how the defender approaches the client's matter, disclosure and consent to utilize AI tools promotes trust, protects the attorney-client relationship, and satisfies the most conservative reading of the ethical rules.

7. *Responding on Your Feet*

Part II identified the risk that over-reliance on AI could erode or “flatten” defenders' advocacy skills. Courtroom practice demands capacities that AI cannot replicate, including reading a witness's body language, responding to unexpected testimony, adjusting strategy based on a judge's reactions, and advocating with the passion and authenticity that effective defense requires. Defenders should guard against skill erosion. Several strategies may help:

- a. Junior attorneys should develop foundational skills before relying heavily on AI. Early-career defenders benefit from performing research, drafting, and analysis manually, developing instincts and capabilities that AI assistance might otherwise prevent from developing. Once those skills are established, AI can enhance them, but AI makes a poor substitute for skills never learned.
- b. Regular practice without AI maintains skills that might otherwise atrophy. Defenders who typically use AI for initial drafts should periodically draft from scratch, identifying whether their unassisted capabilities have declined and where refresher practice might help.
- c. Courtroom and/or trial preparation should emphasize human skills alongside AI efficiency. AI might generate a cross-examination outline, but the defender should internalize that outline, practice executing it, and prepare for departures from the script. The goal is readiness to perform effectively with or without AI assistance.
- d. Offices should structure expectations to preserve skill development. If supervisors evaluate attorneys based solely on productivity, the pressure to maximize AI assistance may override professional development concerns. Evaluation criteria should value analytical quality, courtroom performance, and professional growth alongside raw productivity.

8. *Sustainability Best Practices*

The environmental impact of AI warrants practical attention. Defenders committed to responsible resource use can take several steps to minimize AI's carbon footprint.

Smaller, task-specific AI models generally consume less energy than massive general-purpose systems and, in low-resource settings, “smaller, fine-tuned models can provide significant energy and performance benefits for specific applications.”⁷⁴ When a specialized tool can accomplish a task, using it rather than a larger model reduces environmental impact. The emerging landscape of smaller, efficient models makes this approach increasingly practical.

Cloud providers vary in their environmental practices. Some have committed to powering data centers with renewable energy; others have not. Defenders selecting cloud-based AI services should consider environmental practices alongside security and functionality.

Efficient prompting reduces unnecessary AI queries. Thoughtful, well-structured prompts that accomplish tasks in fewer interactions consume less energy than iterative, trial-and-error approaches. This efficiency also saves time and cost, aligning environmental and practical interests. Contextual



prompting — providing the AI with relevant background information upfront rather than through extended back-and-forth — similarly reduces query volume. A single well-constructed prompt often outperforms multiple exchanges that gradually clarify the task. Some prompt generators like Prompt Cowboy⁷⁵ can help generate efficient prompts. Moreover, efficient prompting is an indication that the defender has developed and maintained the human-legal skills to inform the prompt.

Finally, defenders should avoid unnecessary AI use. When a task can be accomplished efficiently without AI, choosing the non-AI approach eliminates the associated environmental cost. AI is a tool, and deploying it only where it genuinely helps conserves resources while maintaining effectiveness.

B. AI Use Policies

Individual judgment cannot substitute for systematic policy. Defenders should adopt written AI use policies that establish clear expectations, promote consistent practices, and provide a framework for training and accountability.

1. *Policy Purpose and Scope*

An effective AI use policy begins by stating its purpose and ensuring that AI use supports ethical, zealous advocacy while safeguarding client information. The policy should apply to all personnel who might use AI in connection with the office's work, including attorneys, paralegals, administrative staff, and contractors. It should cover all AI tools, not only large language models but also image generators, transcription services, analytics platforms, and any other AI-powered, legal applications (e.g., CoCounsel or Lexis+ AI).

The policy should acknowledge that it will require updating as technology evolves. Establishing a review schedule (annually at minimum, with interim reviews when significant technological or regulatory changes occur) ensures the policy remains current.

2. *Governance*

Assigning clear roles and responsibilities promotes accountability. Offices should designate an individual, task force, or committee responsible for AI policy development, implementation, and oversight. In larger organizations, this might be a formal governance committee with diverse membership, and in smaller offices, a single attorney might assume the role.⁷⁶ Governance responsibilities include evaluating and approving AI tools for office use, developing and updating policies and procedures, overseeing training



and compliance, responding to incidents or concerns, and coordinating with IT and security functions where applicable. Clear assignment of these responsibilities prevents gaps where issues fall through organizational cracks.

3. *Permissible vs. Prohibited Uses*

The policy should clearly distinguish between uses that are permitted, uses that require supervisory approval, and uses that are prohibited entirely.

Permissible uses might include drafting routine administrative documents, researching general legal principles, generating templates and checklists, and other low-risk applications as discussed above. The policy should specify which approved tools may be used for these purposes.

Uses requiring supervisory approval might include drafting case-specific documents like motions or briefs, analyzing client-provided information, processing discovery materials, and other medium-risk applications. The approval requirement ensures oversight without prohibiting beneficial uses.

Prohibited uses should include any use of non-approved AI tools for client matters, entering confidential information into open systems, relying on AI outputs without verification, and any application that would violate ethical rules or court orders. Moreover, firms and offices should prohibit the use of any personal AI tools for professional work and the commingling of any personal use with professional use. The policy should state these prohibitions clearly and unambiguously.

4. *Confidentiality and Data Security Requirements*

The policy must address confidentiality with specificity. Requirements should include restricting AI use to approved enterprise systems for any work involving confidential client information, prohibiting entry of privileged communications into AI systems unless specifically authorized, establishing protocols for handling discovery materials subject to protective orders, requiring review of AI vendor security practices before approval, and specifying data handling requirements for any in-house AI development.

The policy should also address what happens when things go wrong. If confidential information is inadvertently entered into an unapproved system, staff should have clear guidance on how to report and mitigate client damage. The mitigation system should promote immediate disclosure. This section may look similar to Business Continuity Plan documents or Data Breach Notification / Security Incident Response Plans. Prompt reporting, assessment of harm, client notification where appropriate, and remediation measures should all be addressed.

5. *Accuracy and Verification Protocols*

Written verification expectations establish minimum standards and create accountability. The policy should require attorney verification of all AI outputs before submission to courts, specify that citations must be checked against official sources (with a specific callout that many legal-specific AI tools still hallucinate), establish documentation requirements for verification processes, and define what “verification” means for different types of outputs. For example, verification expectations may be less burdensome for general legal research conducted by a junior attorney to learn more about a particular legal issue common to the firm or office, but more burdensome when using AI to make strategic decisions about a client’s case or to draft a motion or brief that will be submitted to the court. The strongest verification expectations should correlate to high-risk tasks, such as the submission of evidence or exhibits that were created or modified by AI. The policy might distinguish between verification requirements for different risk levels. High-stakes submissions — like sentencing memos — should require documented verification by at least one, if not multiple, human reviewers. Lower-stakes documents might permit less elaborate verification while still requiring human review.

6. *Supervision and Training*

Effective AI use requires training, and compliance requires supervision. The policy should mandate initial training for all personnel before they use AI tools, require periodic refresher training as tools and policies evolve, establish supervisory review requirements for AI-assisted work product, and create mechanisms for answering questions and addressing concerns.

Training content can be developed internally or provided by the various AI vendors (some vendors provide robust training and onboarding support).⁷⁷ The content should cover the office’s AI policies and procedures, the capabilities and limitations of approved tools, confidentiality and verification requirements, recognition of AI errors and hallucinations, and compliance with court orders regarding AI disclosure. Training should be practical, including hands-on exercises with actual tools rather than instruction alone, when practical. This training could include how to effectively and efficiently generate prompts, reduce wasteful use, and otherwise use a specific tool, or how to spot hallucinations, overstatements of law, and other pitfalls of AI-generated content.

7. *Vendor and Procurement Standards*

Policies should establish criteria for evaluating AI vendors and guide procurement decisions. Required vendor characteristics might include demonstrated security practices and certifications, contractual commitments to confidentiality and data protection, transparency about data handling and retention, financial stability and business continuity planning, and responsiveness to security inquiries and incident notification. Others might include sustainability efforts, training and onboarding offerings,



and customer support. The policy should specify who may commit the office to AI vendor relationships and what approval process applies, which requires a careful review of any third-party service terms either directly imposed by or flowed down from underlying model providers. This is particularly important because upstream obligations are often unenforceable by the attorney yet fully enforceable against the firm or office, creating an asymmetric risk exposure that most users do not recognize. Solo practitioners often make these decisions personally; larger organizations need procurement procedures that ensure appropriate review before adoption. Offices committed to sustainability should prefer vendors with demonstrated environmental practices, including renewable energy commitments and carbon and/or water reduction goals.

8. *Disclosure and Transparency*

Many courts now require disclosure or certification regarding AI use in filings. NACDL acknowledges this and strongly believes that court rules about disclosure and transparency should apply equally to both the defender and prosecution functions (and their various agents). With that caveat, policies should require compliance with all applicable court orders regarding AI disclosure, establish procedures for tracking and certifying AI use, identify responsibility for monitoring new disclosure requirements, and create documentation practices that support accurate certification. For example, the policy might have specific language the firm or office uses for AI disclosures, either as one single disclosure or as multiple models for different uses or outputs. Some offices may choose to adopt disclosure practices beyond what courts require. Voluntary transparency about AI use can demonstrate good faith and preempt concerns that might otherwise arise.

9. *Monitoring and Enforcement*

Policies without enforcement become suggestions. Effective policies establish mechanisms for monitoring compliance, auditing AI use against policy requirements, investigating potential violations, and imposing appropriate consequences for violations. Monitoring does not have to be intrusive. Technical controls can restrict access to approved tools and flag potentially problematic uses, although this often requires the vendor to configure the monitoring unless the firm or office has sophisticated information technology support or other qualified professionals. Review of work product can identify AI-related issues. Perhaps most important, a culture of compliance in which defenders and staff view ethical and efficient AI use as a shared professional responsibility reduces the need for heavy-handed enforcement.

When violations occur, proportionate responses maintain policy credibility. Minor inadvertent violations might warrant coaching and remediation. Serious or repeated violations might justify formal discipline. The policy should provide enough guidance that personnel understand the stakes without creating rigidity that discourages using AI altogether.



C. Additional Policy Recommendations

Beyond office-level policies, defenders should pursue collective approaches to AI governance. Moreover, no model policy can anticipate every organizational context, and this one is no exception. A large law firm deploying AI across dozens of practice groups faces a fundamentally different set of risks than a solo practitioner using an AI drafting tool for correspondence, and a policy calibrated for one will almost certainly be over – or under-inclusive for the other. The provisions will require careful tailoring: what counts as an acceptable use case, how broadly or narrowly to define permitted tasks, and where to require human review are questions that turn on facts specific to each organization and its size, its practice areas, its client base, and its existing data governance infrastructure. Organizations should treat the framework offered here as a starting point for internal deliberation rather than a finished compliance solution and should expect that meaningful implementation will require revisiting and refining these provisions as their own AI use matures.

1. *Timeline for Adoption*

Although this white paper recommends defense lawyers and offices should implement formal AI use policies within one year, it recognizes that not all defender offices, firms, or practices operate the same. Indeed, it may be impossible to implement an AI use policy for some systems within a year. At the very least, setting a deadline for development and implementation of an AI policy is essential. Given the pace of AI adoption and the risks of unguided use, delay exposes clients and practitioners to unnecessary harm. Interim measures should begin immediately. Even without a comprehensive written policy, defenders can establish informal guidelines, restrict use of non-secure and/or personal AI tools, and ensure that any AI use receives appropriate oversight. The model policy provided in Appendix A offers a template that defenders can adapt to their specific firm or office.

2. *National Coordination*

NACDL is well-positioned to coordinate AI policy development across the defense community. Consistent policy language may help reduce confusion when defenders collaborate across organizations. Shared resources — including model policies, training materials, and vendor evaluations — reduce the burden on firms and offices, especially those without adequate resources to develop policies and procure tools from scratch. NACDL's Task Force on Artificial Intelligence has compiled a list of helpful Champion® articles and other resources about adopting AI in criminal defense, including commentary from the NACDL President, techniques for mastering prompt engineering, and third-party resources (e.g., books, articles, webinars) that may be helpful to defenders.⁷⁸

3. *Collaboration and Knowledge Sharing*

The defense community should share lessons learned about AI use. What tools work well? What pitfalls should others avoid? How have offices addressed particular challenges? Informal networks, conference presentations, published articles, and online forums can all facilitate knowledge sharing. Coordinated responses to prosecutorial AI use amplify individual efforts. When prosecutors use AI in ways that disadvantage defendants — whether through inadequate disclosure or intentional or unintentional misuse — defenders who share strategies for challenging such use strengthen the entire community’s advocacy.

Defenders should also lead in advocating for environmentally responsible AI practices. By adopting and sharing sustainable practices, defenders demonstrate that environmental responsibility and effective representation can coexist. Defenders may also create market pressure for greener AI infrastructure by demanding environmental accountability from vendors, especially those providing legal-specific tools. And by publicly committing to sustainability, the defense community models commitments that the broader legal profession can emulate.



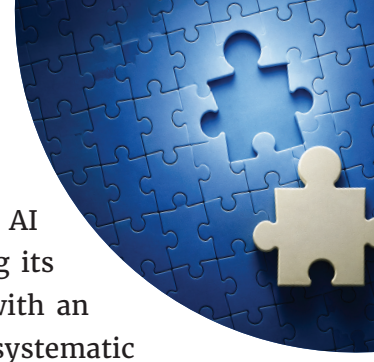
Conclusion

Generative artificial intelligence is already an increasingly integrated facet of the criminal legal system, and defenders have an ethical duty to use it to enhance their advocacy in and out of the courtroom.

The technology offers genuine benefits. Ethical and efficient AI can help overburdened defenders manage growing workloads, extract insights from mountains of digital evidence, and allow defenders to focus on the judgment-intensive work — the human skills — machines cannot perform. For defenders struggling to provide adequate representation amid chronic underfunding and excessive caseloads, AI represents a force multiplier that might help close the gap between what clients need and what overwhelmed attorneys can deliver. Defenders must acknowledge the adversarial nature of practice, too, and understand that prosecutors are already using sophisticated AI tools for both mundane research and high-risk tasks.

Yet this potential comes with risk. The same tools that promise efficiency can compromise confidentiality when defenders enter client information into unsecure systems. The same tools that generate persuasive prose can fabricate citations, facts, and work product that sounds authoritative but is pure fiction (or exaggerated truth). The same tools that help defenders process discovery, draft eloquent motions or briefs, and craft courtroom exhibits, can erode the foundational skills that effective advocacy requires from defenders, especially early career defenders. And the same tools that offer convenience consume vast quantities of energy, contributing to environmental harms that NACDL has specifically committed to addressing.

The risks extend beyond individual practice to systemic concerns about fairness and equal justice. Prosecutors have moved aggressively to adopt AI tools, leveraging government budgets and institutional resources to integrate technology into high-risk practices and tasks, such as charging decisions, bail recommendations, and sentence recommendations. If defenders cannot achieve comparable capabilities, the adversarial balance tips further against the accused. Indigent defendants represented by under-resourced public defenders may receive fundamentally different — and inferior — representation than wealthy defendants who can afford counsel with access to sophisticated AI tools. This two-tiered system would compound existing inequities in a criminal legal system already marked by profound disparities, and requires the type of “calibration” envisioned by Professor Simshaw for the legal profession, generally.⁷⁹



The Model Rules of Professional Conduct demand neither uncritical embrace of AI nor willful ignorance of it. Defenders who rush to adopt AI without understanding its limitations risk harming their clients, and those who reject it out of discomfort with an unfamiliar technology risk falling irretrievably behind. They leave their clients at a systematic disadvantage against AI-equipped prosecutors.

This white paper advocates for informed, ethical engagement with AI. That means understanding what AI can and cannot do, deploying it where benefits outweigh risks, implementing safeguards that protect against known harms, and improving the human skills that effective defense ultimately requires. This approach treats AI as a valuable tool when used appropriately, and defenders' professional responsibilities and ethical obligations demand its use in and out of the courtroom.

Endnotes

1. This white paper is intended for all criminal defense attorneys in both public and private sectors. For brevity, criminal defense lawyers will generally be referred to as “defenders” for the remainder of the white paper.
2. “Flattening” refers to a shift away from developing and using distinctly human skills and reasoning toward increased use of and reliance on automation and AI assistance in daily lawyering tasks. This white paper is not suggesting that using automation and AI assistance is inherently harmful to a lawyer’s professional development; rather, it argues that the risk of “flattening” is something all defenders — all lawyers — should be cognizant of as they learn about and use AI tools in their practice.
3. See Alissa Marque Heydari, *AI & Prosecution: Mapping the Current and Future Roles of Artificial Intelligence in Prosecution* Vanderbilt Project on Prosecution Policy & Policing Project at N.Y.U. Sch. of Law (Dec. 2024).
4. *Id.* at 5.
5. Krystal Hu, *ChatGPT Sets Record for Fastest-Growing User Base*, REUTERS (Feb. 2, 2023), <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.
6. Andrew Perlman, *The Implications of ChatGPT for Legal Services and Society*, THE PRACTICE, (Mar./Apr. 2023), <https://clp.law.harvard.edu/knowledge-hub/magazine/issues/generative-ai-in-the-legal-profession/the-implications-of-chatgpt-for-legal-services-and-society/>.
7. Jamie Bernstein, *Use of Artificial Intelligence-Based Tools by Lawyers in Criminal Advocacy*, 31 HOFSTRA LAW STUDENT WORKS 1, 9 (2024).
8. Brandon L. Garrett & Cynthia Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice*, 109 CORNELL L. REV. 561, 565 (2024).
9. Angela Jin & Niloufar Salehi. 2024. *(Beyond) Reasonable Doubt: Challenges that Public Defenders Face in Scrutinizing AI in Court*. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI ’24). ASSOCIATION FOR COMPUTING MACHINERY, New York, NY, USA, Article 467, 1–19. <https://doi.org/10.1145/3613904.3641902>.
10. Shaila Dewan, *Prosecutor Used Flawed A.I. to Keep a Man in Jail, His Lawyers Say*, N.Y. TIMES, Nov. 28, 2025, <https://www.nytimes.com/2025/11/25/us/prosecutor-artificial-intelligence-errors-lawyers-california.html>.
11. *People v. Smith* (2025 IL 130067), Oral Arg. at 40:48–42:39, available at <https://www.illinoiscourts.gov/courts/supreme-court/oral-argument-audio-and-video/>.



12. *Id.* at 42:17-31.
13. *Id.* at 36:50-37:27.
14. For a recent example involving the representation of MyPillow CEO Mike Lindell, see Jaclyn Diaz, *A recent high-profile case of AI hallucination serves as a stark warning*, NATIONAL PUBLIC RADIO, July 10, 2025, <https://www.npr.org/2025/07/10/nx-s1-5463512/ai-courts-lawyers-mypillow-fines>.
15. *Id.*; see <https://www.damiencharlotin.com/hallucinations/>.
16. Drew Simshaw, *Access to A.I. Justice: Avoiding an Inequitable Two-Tiered System of Legal Services*, 24 YALE J. L. & TECH. 150, 154-56 (2022).
17. *Id.*
18. *Id.*
19. David Nutt, *'Roadmap' shows the environmental impact of AI data center boom*, Cornell Chronicle, Nov. 10, 2025, <https://news.cornell.edu/stories/2025/11/roadmap-shows-environmental-impact-ai-data-center-boom>.
However, there is at least some evidence to suggest AI data center water usage has been grossly overstated, and data centers may use less water than other manufacturing plants and developments. See Molly Taft, *How Much Water Do AI Data Centers Really Use?*, UNDARK, Dec. 16, 2025, <https://undark.org/2025/12/16/ai-data-centers-water/>.
20. National Institute of Standards and Technology (NIST), *Secure Software Development Practices for Generative AI and Dual-Use Foundation Models*, Special Publication 800-218A (July 2024).
21. "Machine learning" means development and use of computer systems that adapt and learn from data with the goal of improving accuracy. NIST, *Measurement Guide for Information Security*, Special Publication 800-55v1 (Dec. 2024).
22. "Data analytics" includes, but is not limited to, data collection followed by the imposition of a model, and the analysis, estimation, and testing that follow focus on the parameters of that model. *Id.*
23. An "algorithm" means a set of rules that, if followed, will give a prescribed result. NIST, *Recommendation for Applications Using Approved Hash Algorithms*, Special Publication 800-107 (Rev. 1, Aug. 2012).
24. Sam Altman, Chief Exec. Officer, OpenAI, Written Testimony Before the S. Comm. on the Judiciary, SUBCOMM. ON PRIV., TECH. & THE L., 118th Cong. (May 16, 2023).
25. Debra Cassens Weiss, *Latest version of ChatGPT acs bar exam with score nearing 90th percentile*, ABA JOURNAL, Mar. 16, 2023, <https://www.abajournal.com/web/article/latest-version-of-chatgpt-acs-the-bar-exam-with-score-in-90th-percentile>.



26. Neal Feigenson and Brian Carney, *Generative AI as Courtroom Evidence: A Practical Guide*, 52 Mitch. Ham. L. Rev. 1, 54-5 (2025).
27. Chuck Kellner, *The End of Reality? How to Combat Deepfakes in Our Legal System*, A.B.A. J. (Mar. 10, 2025), <https://www.abajournal.com/columns/article/the-end-of-reality-how-to-combat-deepfakes-in-our-legal-system>.
28. *Open vs Closed Source AI: Key Differences & Impact Explained*, EDITORIALGE, March 10, 2024, <https://editorialge.com/open-vs-closed-source-ai/>.
29. ABA STANDING COMM. ON ETHICS & PROF. RESPONSIBILITY, Formal Op. 512 (2024) (stating that lawyers using generative AI “must make ‘reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of the client’” and must understand how AI tools handle and store data before inputting confidential information).
30. *The Open or Closed AI Dilemma*, BIPARTISAN POLICY CTR., May 2, 2024, <https://bipartisanpolicy.org/article/the-open-or-closed-ai-dilemma/>.
31. Sachin Gopal Wani, Tanisha Khurana, David Ellison, Matthew Ziegler, Jarrett Upton, *On-Premise vs Cloud: Generative AI Total Cost of Ownership*, LENOVO PRESS, May 23, 2025, <https://lenovopress.lenovo.com/lp2225-on-premise-vs-cloud-generative-ai-total-cost-of-ownership>.
32. See CAL. STANDING COMM. ON PROF. RESPONSIBILITY & CONDUCT, *Practical Guidance For the Use of Generative Artificial Intelligence in the Practice Of Law* (2024) (a lawyer should “understand to a reasonable degree how the technology works, its limitations, and the applicable terms of use and other policies governing the use and exploitation of client data by the product.”); FLA. BAR ADVISORY OP. 24-1 (2024) (building on previous ethics opinions related to cloud-computing and extending those recommendations to use of third-party generative AI).
33. Kevin James Doran, *AI and You: When Lawyers Should Think Twice about Using AI*, ABA GPSOLO EREPORT, Nov. 2024, <https://www.americanbar.org/groups/gpsolo/resources/ereport/2024-november/ai-you-when-lawyers-should-think-twice-using-ai>.
34. *Id.*
35. Varun Magesh, Faiz Surani, Matthew Dahl, Mirac Suzgun, Christopher D. Manning & Daniel E. Ho, *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, 22 J. EMPIRICAL LEGAL STUD. 216 (2025). This study was conducted in 2025 on models available in 2024. Readers should note that AI tools evolve faster than the research. By the time this white paper is published, tools will have changed again. Readers should focus on patterns “that persist across studies rather than exact percentages that will be outdated in months.” Rebecca Fordon, *What the Science Says About Hallucinations in Legal Research*, Mar. 9, 2026, <https://www.llrx.com/2026/02/what-the-science-says-about-hallucinations-in-legal-research/>.



36. *Id.*; see Bob Ambrogi, *Legal Aid Organizations Embrace AI at Twice the Rate of Other Lawyers, New Study Reveals*, LAW SITES, September 25, 2025, <https://www.lawnext.com/2025/09/legal-aid-organizations-embrace-ai-at-twice-the-rate-of-other-lawyers-new-study-reveals.html>; see also Kristen Sondag, *AI for Legal Aid: How to empower clients in need*, THOMSON REUTERS, October 15, 2025, <https://www.thomsonreuters.com/en-us/posts/legal/ai-for-legal-aid-empowering-clients/>.
37. Sondag, *supra* 38.
38. Keely Quinlan, *L.A. County's public defender uses AI to improve client management*, STATESCOOP, July 12, 2023, <https://statescoop.com/la-county-public-defender-ai-aws/>.
39. *Id.*
40. **eDefender**, JOURNAL TECHNOLOGIES, INC., [HTTPS://WWW.JOURNALTECH.COM/EDEFENDER](https://www.journaltech.com/edefender).
41. *Existing AI Tools for Criminal Defense*, UC BERKELEY LAW, <https://www.law.berkeley.edu/research/criminal-law-and-justice-center/our-work/ai-for-public-defenders/existing-ai-tools/> (last accessed Dec. 12, 2025).
42. *Id.*
43. *How Wisconsin attorney Jessa Nicholson Goetz used JusticeText to uncover police misconduct buried within 40 hours of footage*, JUSTICETEXT, <https://justicetext.com/nicholson-goetz-otis-case-study/>.
44. Matt Reynolds, *California Innocence Project harnesses generative AI for work to free wrongfully convicted*, ABA JOURNAL, Aug. 14, 2023, <https://www.abajournal.com/web/article/california-innocence-project-harnesses-generative-ai-for-work-to-free-wrongfully-convicted>.
45. *Existing AI Tools for Criminal Defense*, UC BERKELEY LAW, *supra* note 42.
46. *Id.*
47. Juliana Kim, *Family shows AI video of slain victim as an impact statement — possibly a legal first*, NATIONAL PUBLIC RADIO, May 12, 2025, <https://www.npr.org/2025/05/07/g-s1-64640/ai-impact-statement-murder-victim>. See also Sarah Martinson, *AI Video Pushes Boundaries Of Victim Impact Statements*, LAW360, May 30, 2025, <https://www.law360.com/pulse/articles/2340191/ai-video-pushes-boundaries-of-victim-impact-statements> (claiming that generative AI use for victim impact statements potentially poses both legal and ethical issues, and suggesting that generative AI video creation has advanced to the point where forensics examiners can't tell the difference between real and fake videos, and AI detection tools are being rendered obsolete).
48. Heydari, *AI & Prosecution: Mapping the Current and Future Roles of Artificial Intelligence in Prosecution*, *supra* note 5.



49. Eric Friedman, Adrienne Busby, and Erica Moran, *Understanding the Rule 702 Amendments*, AM. BAR ASS'N, Apr. 15, 2024, <https://www.americanbar.org/groups/litigation/resources/newsletters/corporate-counsel/understanding-rule-702-amendments/>.
50. Michael L. Huggins and Hannah L. McMeans, *Proposed New Federal Rule Regarding AI-Generated Evidence*, MEYERS NAVE, Feb. 18, 2026, <https://www.meyersnave.com/proposed-new-federal-rule-regarding-ai-generated-evidence/>.
51. Nat'l Ass'n Criminal. Def. Lawyers, Comments to Proposed Federal Rule of Evidence 702, <https://www.nacdl.org/getattachment/8474211d-c05b-43a1-ae95-9dbe2b65b19c/nacdl-comments-on-proposed-amendments-to-federal-rule-of-evidence-707-02162026.pdf>.
52. Rebecca A. Delfino, *Deepfakes on Trial: A Call to Expand the Trial Judge's Gatekeeping Role to Protect Legal Proceedings from Technological Fakery*, 74 HASTINGS L. J. 293 (2023).
53. John P. LaMonaga, *A Break From Reality: Modernizing Authentication Standards for Digital Video Evidence in the Era of Deepfakes*, 69 Am. Univ. L. Rev. 1945 (2020).
54. Judge Paul Grimm (Ret.) & Dr. Maura Grossman, ADVISORY COMM. ON EVIDENCE RULES, Agenda, Proposed New Rule 901(c) to Address "Deepfakes," at 18 (Apr. 19, 2024).
55. Standing Orders, Local Rules, and Decisions on the Use of AI, Ropes & Gray LLP, last updated Jan. 8, 2026, <https://www.ropesgray.com/en/sites/artificial-intelligence-court-order-tracker>.
56. *Supra* note 31.
57. ABA TASK FORCE ON LAW AND ARTIFICIAL INTELLIGENCE, *Year 2 Report on the Impact of AI on the Practice of Law* (Dec. 2025), https://www.americanbar.org/groups/leadership/office_of_the_president/ai-task-force/.
58. *In Focus: Artificial Intelligence (AI)*, BLOOMBERG LAW, Mar. 9, 2026, https://www.bloomberglaw.com/external/page/infocus_artificial_intelligence.
59. *AI and Attorney Ethics Rules: 50-State Survey*, JUSTIA, Apr. 2025, <https://www.justia.com/trials-litigation/ai-and-attorney-ethics-rules-50-state-survey/>.
60. Simshaw, *Access to A.I. Justice*, *supra* note 18 at 157.
61. *Id.* at 181.
62. *Id.* at 182; see Agnieszka McPeak, *Disruptive Technology and the Ethical Lawyer*, 50 U. TOL. L. REV. 457 at 471 (2019) (writing, "[T]echnology has been integrated into the very act of practicing law. The core activity of lawyering—that of thinking like a lawyer—is expressed through the technology lawyers use. At its core, technology is not merely a tool of the trade, but it is wrapped up intrinsically in the very thought processes lawyers employ.").



63. Garrett & Rudin, *The Right to a Glass Box*, *supra* note 10, at 567.
64. *Id.* at 568-69.
65. Nicholas Daniel Seger, *Understanding the Risks of Uploading Client Information to Generative AI Platforms*, ABA YOUNG LAWYERS DIVISION, PRACTICE MGMT., Jan 16, 2024, https://www.americanbar.org/groups/young_lawyers/resources/tyl/practice-management/risks-uploading-client-information-generative-ai-platforms/.
66. For a detailed overview on public defenders' perspectives on AI's compatibility with human representation, see Irene Y. Cheong et al., *How Can AI Augment Access to Justice? Public Defenders' Perspectives on AI Adoption*, *arXiv preprint arXiv:2510.22933* (2025), available at <https://arxiv.org/abs/2510.22933>.
67. Andrew (A.J.) Tibbetts, *Navigating Confidentiality Risks in Third-Party AI Tools*, GREENBERG TRAURIG LLP, July 11, 2025, <https://www.gtlaw.com/en/insights/2025/7/published-articles/navigating-confidentiality-risks-in-third-party-ai-tools>.
68. Kristin Bergman, *Access to Justice 2.0: How AI-Powered Software Can Bridge the Gap*, A.B.A. J. (2025), <https://www.americanbar.org/groups/journal/articles/2025/access-to-justice-how-ai-powered-software-can-bridge-the-gap/>.
69. Kelly Raskovich, *Tech Trends 2026*, DELOITTE INSIGHTS, Dec. 10, 2025, <https://www.deloitte.com/us/en/insights/topics/technology-management/tech-trends.html>.
70. Heydari, *AI & Prosecution: Mapping the Current and Future Roles of Artificial Intelligence in Prosecution*, *supra* note 5 at 16.
71. *Id.* at 5.
72. *Id.* at 18-9.
73. *Id.* at 20.
74. Leona Verdadero, et. al., *Smarter, Smaller, Stronger: Resource-Efficient Generative AI & the Future of Digital Transformation*, UNESCO (2025), at 18-9.
75. <https://www.promptcowboy.ai/>.
76. Firms and offices are encouraged to identify the people on staff who understand the tools and bring them into the strategizing and decision-making processes when drafting an AI use policy.
77. *The Essential Questions to Ask Your AI Vendor Before Deploying Artificial Intelligence at Your Organization*, FISHER PHILLIPS LLP, Nov. 19, 2024, <https://www.fisherphillips.com/en/news-insights/essential-questions-to-ask-ai-vendor-before-deploying-artificial-intelligence.html>.



78. www.nacdl.org/ai

79. Simshaw, *Access to A.I. Justice*, *supra* note 18.

80. SOC 2 (System and Organization Controls) is a compliance and privacy standard that specifies how organizations should manage data and systems to ensure confidentiality, integrity, and availability. The standard is designed for service organizations like cloud providers, software as a service (SaaS) vendors, and other organizations that provide web-based services. A Type II compliance audit shows how an organization complies with SOC 2 for a stated period of time.



Appendix A

Model AI Use Policy

Effective Date: [Insert Date]

Last Revised: [Insert Date]

I. PURPOSE AND SCOPE

A. Purpose

This Policy establishes guidelines for the responsible and ethical use of artificial intelligence (AI) tools, including generative AI, in the practice of criminal defense. The purpose of this Policy is to ensure that AI use supports ethical, zealous representation while safeguarding client rights. Specifically, this Policy:

- Ensures compliance with applicable Model Rules of Professional Conduct and ethical obligations;
- Protects client confidentiality and attorney-client privilege;
- Maintains the quality and accuracy of legal work product;
- Enables attorneys to leverage AI tools to enhance representation while mitigating associated risks;
- Complies with court rules and orders regarding AI use and disclosure; and
- Provides guidance specific to the unique ethical obligations of criminal defense practice.



B. Scope

This Policy applies to all attorneys, law clerks, paralegals, legal assistants, investigators, support staff, and any other personnel who perform work on criminal defense matters for [Organization Name]. It governs the use of all AI technologies, including but not limited to:

- Generative AI tools (e.g., ChatGPT, Claude, Gemini, Copilot);
- Legal-specific AI platforms (e.g., Harvey, Lexis+ AI, CoCounsel);
- Evidence review and discovery tools with AI capabilities;
- Transcription and translation AI services;
- Video and audio analysis tools; and
- AI-enhanced document drafting and proofreading tools.

II. GOVERNANCE

A. AI Policy Committee

[Organization Name] shall establish an AI Policy Committee responsible for overseeing the implementation and ongoing administration of this Policy. The Committee shall include representation from leadership, practicing attorneys, information technology, and administrative staff.

B. Committee Responsibilities

The AI Policy Committee shall:

- Review and approve AI tools for organizational use;
- Maintain a current list of approved AI tools and their permitted uses;
- Develop and update training materials and protocols;
- Receive and respond to questions regarding AI use;
- Monitor developments in AI technology, ethics guidance, and court rules;
- Conduct periodic audits of AI use within the organization;
- Investigate potential policy violations; and
- Recommend policy amendments as needed.



C. Designated AI Compliance Officer

The [Chief Public Defender / Managing Partner / Executive Director] shall designate an AI Compliance Officer to serve as the primary point of contact for AI-related questions, coordinate training efforts, maintain records of approved tools, and report to the AI Policy Committee on compliance matters.

D. Role Assignments

The following roles and responsibilities are assigned under this Policy:

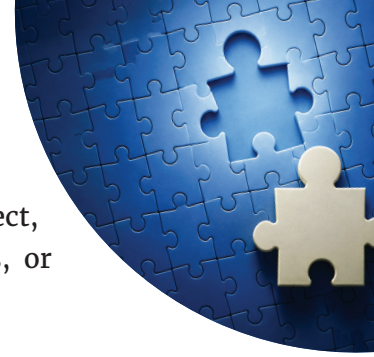
- Supervisory Attorneys: Responsible for ensuring compliance among subordinates, approving AI use in sensitive matters, and periodically reviewing AI-assisted work product to “spot-check” its use;
- Individual Attorneys: Responsible for complying with this Policy, verifying all AI outputs, and maintaining professional responsibility for all work product;
- Support Staff: Responsible for using only approved AI tools for authorized purposes and promptly reporting any concerns to supervisors;
- IT Personnel: Responsible for implementing technical safeguards, maintaining security of AI systems, and supporting the AI Policy Committee in evaluating tools.

III. DEFINITIONS

Artificial Intelligence (AI): Computer systems designed to perform tasks that typically require human intelligence, including learning, reasoning, problem-solving, and natural language processing.

Generative AI (GenAI): AI systems capable of generating text, images, code, or other content in response to prompts, including large language models (LLMs) such as ChatGPT, Claude, Gemini, and similar tools.

Approved AI Tool: An AI tool that has been reviewed and authorized for use by [Organization Name] in accordance with this Policy.



Hallucination: AI-generated content that appears plausible but is factually incorrect, fabricated, or nonexistent, including fictitious case citations, invented quotations, or false legal propositions.

Confidential Client Information: All information relating to the representation of a client, regardless of source, including but not limited to personally identifiable information, case facts, legal strategy, privileged communications, and work product.

Open AI System: Consumer-facing AI tools that may retain user inputs, use data for model training, or lack enterprise-grade security protections.

Enterprise-Secure AI Tool: AI tools that provide contractual guarantees regarding data handling, do not use inputs for training, and meet organizational security requirements.

IV. ETHICAL FRAMEWORK

The use of AI in criminal defense practice implicates several fundamental ethical obligations. This Policy is designed to ensure compliance with applicable Rules of Professional Conduct, including:

A. Competence (Model Rule 1.1)

Attorneys must provide competent representation, which requires the legal knowledge, skill, thoroughness, and preparation reasonably necessary for the representation. Comment [8] to Model Rule 1.1 provides that lawyers should keep abreast of the benefits and risks associated with relevant technology. Personnel using AI tools must:

- Understand the capabilities and limitations of any AI tool used;
- Recognize that AI tools are prone to hallucinations and errors;
- Exercise independent professional judgment in evaluating AI-generated output; and
- Ensure that reliance on AI does not diminish the quality of representation.



B. Confidentiality (Model Rule 1.6)

Attorneys must protect confidential client information. Using AI tools may implicate confidentiality obligations because:

- Many AI platforms retain user inputs for training or quality improvement purposes;
- Information entered into AI tools may be accessible to third parties;
- Disclosure of privileged information to third-party AI providers may waive attorney-client privilege; and
- Criminal defense matters often involve highly sensitive information requiring heightened protection.

C. Communication (Model Rule 1.4)

Attorneys must reasonably consult with clients about the means by which their objectives are to be accomplished. Depending on the circumstances, this may require informing clients about the use of AI tools in their representation.

D. Supervision (Model Rules 5.1 and 5.3)

Supervisory attorneys must ensure that subordinate attorneys and nonlawyer staff comply with professional obligations when using AI tools. Supervisors should establish clear policies, provide training, and implement appropriate oversight mechanisms.

E. Candor to the Tribunal (Model Rule 3.3)

Attorneys must not knowingly make false statements of fact or law to a tribunal. This obligation requires careful verification of all AI-generated content before submission to any court.



V. PERMISSIBLE AND PROHIBITED USES

A. Permissible Uses

Subject to compliance with the requirements of this Policy, AI tools may be used for the following purposes:

1. ADMINISTRATIVE TASKS

- Summarizing case files and documents;
- Preparing internal memoranda and case notes;
- Drafting non-sensitive correspondence;
- Translating documents and communications; and
- Organizing and categorizing files.

2. LEGAL RESEARCH

- Identifying potentially relevant cases, statutes, and legal authorities;
- Generating initial research summaries and outlines;
- Cross-jurisdictional surveys of legal issues; and
- Exploring novel legal arguments and theories.

3. DOCUMENT DRAFTING ASSISTANCE

- Generating initial drafts of motions, briefs, and memoranda;
- Proofreading and editing for grammar, style, and clarity;
- Formatting and organizing documents; and
- Creating templates for routine filings.

4. EVIDENCE REVIEW AND DISCOVERY

- Transcribing and searching audio and video recordings;
- Organizing and categorizing discovery materials;

- Identifying relevant documents and communications;
- Synchronizing multi-camera video evidence; and
- Analyzing patterns and inconsistencies in evidence.

B. Prohibited Uses

The following uses of AI tools are strictly prohibited:

1. CONFIDENTIALITY VIOLATIONS

- Entering confidential client information into any non-approved AI tool;
- Uploading privileged documents to public or open AI platforms;
- Entering discovery materials subject to protective orders into any AI system unless expressly permitted by the order and compliant with its terms;
- Using AI tools that lack adequate confidentiality and security protections for client-related work; and
- Entering personally identifiable information, protected health information, or other sensitive data without proper safeguards.

2. UNVERIFIED SUBMISSIONS

- Filing any document with a court without independent verification of all facts and legal citations;
- Submitting AI-generated content without attorney review and approval; and
- Relying solely on AI-generated legal research without verification through traditional legal databases.

3. IMPROPER DELEGATION

- Using AI to make strategic case decisions that require professional judgment;
- Substituting AI analysis for independent attorney evaluation of case merits;

- Using AI to conduct client communications without attorney oversight; and
- Allowing AI to replace the attorney–client relationship or essential client contact.

VI. CONFIDENTIALITY AND DATA SECURITY REQUIREMENTS

A. Restriction on Open AI Systems

The use of open, consumer–facing AI systems for any work involving client information is prohibited. Personnel may not enter confidential client information into tools such as the free versions of ChatGPT, Gemini, or similar platforms that may retain inputs, use data for training, or lack enterprise security guarantees.

B. Enterprise-Secure Tools Required

Where AI use involving client information is permitted, personnel must use only enterprise–secure tools that have been approved by [Organization Name]. Approved tools must:

- Provide contractual guarantees that user inputs will not be used for model training;
- Offer data encryption in transit and at rest;
- Maintain SOC 2 Type II⁸⁰ certification or equivalent;
- Support role–based access controls and audit logging; and
- Comply with applicable data protection regulations.

C. Privilege Preservation

Entering privileged communications or attorney work product into non–secured AI platforms may constitute disclosure to a third party and could waive privilege protections. To preserve privilege:

- Do not enter verbatim privileged communications into any AI tool without supervisor approval;



- Treat AI-generated summaries of privileged information as privileged and protect accordingly;
- Document the precautions taken to maintain confidentiality when using AI tools; and
- Consult with supervisory counsel before using AI tools for any matter involving particularly sensitive privileged information.

VII. ACCURACY AND VERIFICATION PROTOCOLS

A. Mandatory Human Review

All AI-generated content must be independently reviewed and verified by a qualified attorney before use. AI output may never be submitted to a court, shared with opposing counsel, or provided to a client without thorough human review.

B. Citation Verification

All legal citations generated or suggested by AI must be independently verified. Attorneys must:

- Confirm that each cited case, statute, rule, or regulation exists using traditional legal research databases (e.g., Westlaw, Lexis, official court websites);
- Verify that cited authorities have not been overruled, reversed, or otherwise negatively treated;
- Confirm that cited authorities actually support the propositions for which they are cited;
- Review the full text of any cited authority to ensure accurate representation; and
- Document the verification process for each court filing.

C. Factual Verification

All factual assertions generated by AI must be verified against primary sources, including case files, discovery materials, and other authoritative records.



D. Evidence Review Verification

When using AI tools to review evidence in criminal cases, personnel must:

- Recognize that AI transcriptions and summaries may contain errors or omissions;
- Manually verify critical portions of transcribed audio or video evidence;
- Not rely solely on AI summaries for evidentiary determinations;
- Review original evidence before making arguments based on AI-generated summaries; and
- Be alert to potential AI bias in evidence analysis that could affect case assessment.

VIII. SUPERVISION AND TRAINING

A. Supervisory Obligations

Supervisory attorneys must:

- Ensure that all personnel understand and comply with this Policy;
- Provide adequate training on the ethical and practical aspects of AI use;
- Establish appropriate review and oversight mechanisms;
- Approve AI tool use in sensitive matters when required; and
- Address any violations of this Policy promptly.

B. Training Requirements

All personnel who use AI tools must complete training that covers:

- The requirements of this Policy;
- The capabilities and limitations of AI tools;
- Ethical obligations applicable to AI use;
- Risks associated with AI use, including hallucinations, bias, and confidentiality concerns;

- Verification procedures for AI-generated content;
- Court disclosure requirements; and
- Best practices for prompt engineering and effective AI use.

C. Ongoing Education

Training should be updated regularly to reflect changes in technology, ethical guidance, and legal requirements. Personnel should receive refresher training at least annually and whenever significant changes to this Policy are adopted.

IX. VENDOR AND PROCUREMENT STANDARDS

A. Vendor Evaluation Criteria

Before any AI tool is approved for use, [Organization Name] shall evaluate the vendor based on the following criteria:

- Data handling practices, including whether user inputs are retained or used for training;
- Security certifications and compliance (SOC 2 Type II, CJIS compliance where applicable);
- Encryption standards for data in transit and at rest;
- Access controls and audit logging capabilities;
- Data residency and jurisdictional considerations;
- Incident response and breach notification procedures; and
- Financial stability and long-term viability of the vendor.

B. Contractual Requirements

Contracts with AI vendors must include:

- Confidentiality provisions prohibiting vendor use or disclosure of client data;
- Prohibition on using organizational data to train AI models;
- Data deletion obligations upon termination;

- Security commitments and audit rights;
- Breach notification requirements; and
- Indemnification provisions for security failures.

C. Sustainability Considerations

When evaluating AI tools for adoption, [Organization Name] should consider the provider's commitment to energy efficiency and sustainable practices. AI tools require substantial computational resources with associated environmental impacts, and these factors should be weighed alongside functionality and cost.

X. DISCLOSURE AND TRANSPARENCY

A. Compliance with Court Orders and Rules

An increasing number of courts have issued standing orders, local rules, or case management orders requiring disclosure of AI use in court filings. Before filing any document, attorneys must:

- Review applicable standing orders for the assigned judge;
- Review local rules of the court;
- Review any applicable case management orders;
- Check for any jurisdiction-wide rules or guidance on AI disclosure; and
- Comply fully with all applicable disclosure and certification requirements.

B. Disclosure Statement

When required by court order or rule, filings must include a disclosure statement that:

- Identifies whether AI was used in preparing the filing;
- Identifies the specific AI tool(s) used, if required;
- Identifies which portions of the filing were AI-assisted, if required; and
- Certifies that a human attorney reviewed the filing and verified all citations and factual statements.



C. Client Disclosure

Attorneys should consider whether disclosure of AI use to clients is appropriate or required based on the nature and extent of AI use, the sophistication of the client, client instructions regarding technology, the risks associated with AI use in the particular matter, and applicable ethical rules.

XI. MONITORING AND ENFORCEMENT

A. Auditing

The AI Policy Committee shall conduct periodic audits of AI use within [Organization Name] to assess compliance with this Policy. Audits may include:

- Review of AI tool usage logs and access records;
- Spot-checks of court filings for proper verification and disclosure;
- Surveys of personnel regarding AI use practices;
- Review of vendor compliance with contractual requirements; and
- Assessment of training completion records.

B. Reporting Requirements

Personnel must promptly notify their supervisor when:

- AI has been used to assist with high-risk tasks, including but not limited to those requiring the input of confidential, privileged, and/or personal client information, or the creation or modification of courtroom exhibits.
- There is any question about whether a particular AI use complies with this Policy;
- A potential breach of confidentiality has occurred; or
- An error in AI-generated content has been discovered before or after submission.



C. Enforcement

Violations of this Policy may result in disciplinary action, including:

- Verbal or written warning;
- Mandatory additional training;
- Restriction or revocation of AI tool access;
- Performance-related consequences; and
- In serious cases, termination of employment or engagement.

D. Policy Updates

This Policy shall be reviewed at least annually and updated as necessary to reflect advances in AI technology, new ethical guidance from bar associations, changes to court rules and orders regarding AI use, lessons learned from implementation, and feedback from personnel.

XII. ACKNOWLEDGMENT

All personnel subject to this Policy must acknowledge receipt and understanding of the Policy. By signing below, I acknowledge that I have read and understand this Policy and agree to comply with its requirements.

Signature

Printed Name

Date



For more information contact:

National Association of Criminal Defense Lawyers®

1660 L Street NW, 12th Floor, Washington, DC 20036

Phone 202-872-8600 [NACDL.org](https://www.nacdl.org)

This publication is available online at [NACDL.org/ParityInPractice](https://www.nacdl.org/ParityInPractice)

A version of this article is forthcoming in the inaugural edition of
The Journal of Advocacy and Litigation (Vol. 1, 2026).