Responsible Use of Artificial Intelligence in Legal Practice: Emphasizing Confidentiality through Self-Hosted Large Language Models

Author: Teddy Tennant

Date Published: November 1, 2025

Abstract

Artificial Intelligence (AI) has revolutionized various sectors, including law, by enhancing efficiency, accuracy, and accessibility in legal processes. However, its integration raises significant ethical, practical, and confidentiality concerns. This paper explores the responsible application of AI in legal practice, drawing on ethical guidelines from professional bodies and real-world implementations. It emphasizes the importance of competence, oversight, and data protection. A key focus is on self-hosted Large Language Models (LLMs) as a strategy to safeguard client confidentiality, mitigating risks associated with third-party AI services. By examining benefits, challenges, and best practices, this paper provides a framework for lawyers to adopt AI responsibly while upholding professional standards.

Introduction

The advent of AI technologies, particularly generative AI (GenAI) and LLMs, has transformed the legal profession [1]. From automating document review to aiding in legal research and predictive analytics, AI offers tools that can streamline workflows and reduce costs. However, the responsible use of these technologies is paramount to maintain the integrity of legal practice. Ethical dilemmas arise from issues such as bias, accuracy, and confidentiality, necessitating adherence to professional rules like those outlined by the American Bar Association (ABA) Model Rules of Professional Conduct [2].

This paper delineates strategies for responsible AI adoption in law, with particular emphasis on self-hosted LLMs to protect client information. It draws on recent ethical opinions and practical guidance to propose a balanced approach that leverages AI's potential while mitigating risks [3].

Benefits of AI in Legal Practice

AI's integration into law firms has led to significant advancements in efficiency and decision-making. Tools powered by AI can handle repetitive tasks such as contract analysis, e-discovery, and case prediction, allowing lawyers to focus on complex strategic work [4]. For instance, GenAI assists in drafting legal documents, summarizing case law, and identifying patterns in large datasets, which traditionally consumed substantial time and resources. In commercial litigation, AI enhances document review processes, improving accuracy and speed while reducing human error [5].

Moreover, AI promotes access to justice by enabling pro bono initiatives and supporting under-resourced legal aid organizations through automated tools for case triage and legal advice generation [6]. When used responsibly, these technologies can democratize legal services, making them more affordable and accessible. However, realizing these benefits requires a foundation of ethical awareness and technical competence.

Ethical Considerations in AI Use for Law

Ethical challenges are at the forefront of AI adoption in legal practice. Key concerns include competence, confidentiality, bias, and accountability. Lawyers have a duty under Model Rule 1.1 to maintain technological competence, which extends to understanding AI tools' capabilities and limitations [7]. This includes verifying AI-generated outputs to avoid inaccuracies or "hallucinations"—fabricated information presented as factual—which have led to sanctions in several high-profile cases [8].

Bias and fairness represent another critical area. AI systems trained on historical data may perpetuate systemic inequalities, leading to unfair outcomes in legal predictions or recommendations [9]. Ethical guidelines stress the need for transparency and human oversight to mitigate these risks. Additionally, accountability ensures that lawyers remain responsible for AI-assisted work; they cannot delegate ultimate judgment to machines.

Client confidentiality, governed by Model Rule 1.6, is particularly vulnerable when using cloud-based AI services that may store or process sensitive data [10]. Unauthorized data sharing for AI training purposes can breach privilege, highlighting the need for secure alternatives. Ethical opinions from bodies like the ABA and state bars emphasize obtaining informed client consent and conducting due diligence on AI providers [11].

Responsible Practices for AI Integration

To use AI responsibly, lawyers should adopt a structured approach. First, conduct thorough evaluations of AI tools, including reviewing terms of service for data handling practices. Training and education are essential; firms should invest in programs to build AI literacy among staff [12].

Second, implement oversight mechanisms, such as cross-verification of AI outputs with human review and auditing for biases. Courts and regulatory bodies, like the National Center for State Courts, advocate for principles such as transparency, accountability, and inclusivity in AI deployment [13].

Third, prioritize attorney well-being by integrating AI in ways that support mental health, such as automating mundane tasks to reduce burnout [14]. Finally, stay abreast of evolving regulations and ethical opinions to adapt practices accordingly.

Self-Hosted LLMs for Client Confidentiality

A pivotal aspect of responsible AI use in law is safeguarding client confidentiality, especially when employing LLMs for tasks like legal drafting or analysis. Publicly available LLMs, such as those from third-party providers, often involve data transmission to external servers, raising risks of breaches or unintended use in model training [15]. This contravenes ethical duties under confidentiality rules, as sensitive client information could be exposed without consent.

Self-hosted LLMs address these concerns by allowing firms to deploy models on their own infrastructure, ensuring data remains within controlled environments. Open-source LLMs, such as those based on frameworks like LLaMA or Mistral, can be customized and hosted locally, eliminating reliance on cloud services and reducing costs over time [16]. This approach enhances security through encryption, access controls, and compliance with data protection laws like GDPR or similar standards applicable in legal contexts [17].

Self-hosting preserves attorney-client privilege by preventing data leakage to external entities. For small to mid-sized firms, it avoids recurring subscription fees while maintaining control over updates and customizations tailored to legal needs. Challenges include initial setup costs and technical expertise requirements, but these are offset by long-term benefits in confidentiality and ethical compliance [18]. Firms should evaluate open-source versus proprietary options, prioritizing those with strong privacy features and transparent development processes.

Implementation steps include: (1) selecting a suitable open-source LLM; (2) setting up secure on-premises or private cloud hosting; (3) training the model on anonymized legal data; and (4) regularly auditing for security vulnerabilities. By adopting self-hosted solutions, lawyers can responsibly harness LLMs without compromising client trust.

Data Training and Leakage Risks in Third-Party Services

Online AI services like ChatGPT exemplify the confidentiality pitfalls of cloud-based LLMs. These platforms may use user prompts for model training unless explicitly opted out, potentially incorporating sensitive legal data into broader datasets [19]. Conversely, due to memorization vulnerabilities, adversaries or unrelated users could extract this data through carefully crafted prompts, leading to unintended disclosure [20]. Such risks directly violate Model Rule 1.6's requirements to protect client information from unauthorized access.

Self-hosting mitigates these risks by keeping all processing in-house, ensuring compliance with ethical rules on data protection. Firms utilizing third-party services must conduct rigorous vendor assessments, implement data anonymization protocols, and obtain informed client consent regarding potential risks [21].

Conclusion

Responsible use of AI in law requires commitment to ethical principles, continuous education, and innovative solutions like self-hosted LLMs to protect confidentiality. As AI evolves, legal professionals must balance its transformative potential with safeguards against risks. By following established guidelines and prioritizing client interests, the legal field can integrate AI to enhance justice delivery while upholding core professional values. Future research should explore comprehensive regulatory frameworks and interdisciplinary collaborations to further guide this integration and address emerging challenges in AI governance for legal practice.

References

[1] Surden, H. (2019). Artificial intelligence and law: An overview. *Georgia State University Law Review*, 35(4), 1305-1337.

[2] American Bar Association. (2023). *Model Rules of Professional Conduct*. American Bar Association.

[3] American Bar Association. (2024). Formal Opinion 512: *Generative Artificial Intelligence Tools*.

[4] Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2023). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 381(2251).

[5] Remus, D., & Levy, F. (2017). Can robots be lawyers? Computers, lawyers, and the practice of law. *Georgetown Journal of Legal Ethics*, 30(3), 501-558.

[6] Prescott, J. J., Pyle, J., & Pierson, E. (2022). Understanding civil justice problems and strategies to address them. *Journal of Empirical Legal Studies*, 19(2), 271-338.

[7] American Bar Association. (2012). *Model Rules of Professional Conduct: Rule 1.1 Comment 8*. https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/

[8] *Mata v. Avianca, Inc.*, No. 22-cv-1461 (S.D.N.Y. 2023).

[9] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica*.

[10] American Bar Association. (2024). Model Rule 1.6: Confidentiality of Information. In *Model Rules of Professional Conduct*. American Bar Association.

[11] New York State Bar Association. (2024). Ethics Opinion 1234: *Use of Artificial Intelligence in Legal Practice*.

[12] Calo, R. (2018). Artificial intelligence policy: A primer and roadmap. *UC Davis Law Review*, 51(2), 399-435.

[13] National Center for State Courts. (2024). *Principles for the Use of AI in Courts*.

[14] Krill, P. R., Johnson, R., & Albert, L. (2021). The prevalence of substance use and other mental health concerns among American attorneys. *Journal of Addiction Medicine*, 10(1), 46-52.

[15] Carlini, N., Tramer, F., Wallace, E., et al. (2021). Extracting training data from large language models. *Proceedings of the 30th USENIX Security Symposium*, 2633-2650.

[16] Touvron, H., Lavril, T., Izacard, G., et al. (2023). LLaMA: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

[17] European Parliament. (2016). *General Data Protection Regulation (GDPR)*. Regulation (EU) 2016/679.

[18] Bommasani, R., Hudson, D. A., Adeli, E., et al. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

[19] OpenAI. (2023). *ChatGPT Data Usage and Training Policies*. https://openai.com/policies

[20] Carlini, N., Ippolito, D., Jagielski, M., et al. (2023). Quantifying memorization across neural language models. *arXiv preprint arXiv:2202.07646*.

[21] American Bar Association. (2017). Formal Opinion 477R: *Securing Communication of Protected Client Information*.