



Generating . . . Client Confidentiality Concerns in the Use of Generative AI Technology

Since the emergence of low- to no-cost public access to more sophisticated uses of artificial intelligence (“AI”), there has been lively discussion regarding how lawyers may ethically use various artificial intelligence software and services. The discussion has primarily focused on the duties of competence, communication, confidentiality, detecting conflicts of interest, and supervising.¹ “The potential applications for AI in the legal world are immense and include composing client briefs, producing complex analyses from troves of documents, and helping firms with limited resources compete with the largest groups.”² At, perhaps, its best, AI can provide access to justice tools to help narrow the justice gap.³ No matter the manner of AI use, upholding ethical duties in this new technological frontier will require proficiency in data privacy standards and programming. It also mandates a reasonable understanding of ever-changing professional norms when it comes to the use of technology and the lawyer’s duty to safeguard confidential information.

Arguably, the principle of maintaining confidentiality between the client and lawyer is one of the most important duties that lawyers possess—without it, the client-lawyer relationship cannot be fully functional because clients may be reluctant to share information necessary for representation out of fear that their confidences will not be protected.⁴ Therefore, it is very important to examine how the use of AI implicates the duty of confidentiality. Specifically, upholding the duty while using certain generative AI platforms may require an examination of the duty to protect client confidences from inadvertent or unauthorized disclosure.⁵

The duty to protect client confidences under the Model Rules of Professional Conduct (“MRPC”) is twofold. First, MRPC 1.6(a) commands that: “A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation or the disclosure is permitted by [an exception in] paragraph (b).” Comment 3 to Rule 1.6 goes on to explain that “the principle of client-lawyer confidentiality is given effect by related bodies of law: the attorney-client privilege, the work product doctrine, and the rule of confidentiality established in professional ethics.”⁶ Importantly, the attorney-client privilege and work product doctrine apply in formal judicial and administrative proceedings where a lawyer may be required to provide evidence or testimony concerning communications with a client or notes related to representation. The duty of confidentiality under MRPC 1.6, however, applies in all other situations. The Restatement (Third) of the Law Governing Lawyers § 59 defines confidential client information as “all information relating to the representation of a client, whether in oral, documentary, electronic, photographic, or other forms.”⁷ Additionally, “[i]t includes work product that the lawyer develops in representing the client, such as the lawyer’s notes. . . .”⁸ Thus, the scope of confidential information protected by the duty of confidentiality is broad and includes almost all information concerning the representation of a client.

The second component of the duty to protect client confidences is mandated in MRPC 1.6(c). It compels 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.”⁹ Essentially, this paragraph assigns an objectively reasonable duty of competence to secure and safeguard confidential client information.¹⁰ More specifically, lawyers are required to employ methods that protect electronic documents from “unauthorized access by third parties,” which may happen in the case of a data breach or other cyber-attack.¹¹ If a lawyer, firm, legal aid, or government office experiences unauthorized access or disclosure of confidential client information, this provision is not violated if the lawyer “made reasonable efforts to prevent the access or disclosure.”¹² Factors that can be considered when determining whether reasonable efforts have been made include: “the sensitivity of the information, the likelihood of disclosure if additional safeguards were not employed, the cost of employing additional safeguards, the difficulty of implementing the safeguards, and the extent to which the safeguards solely affect the lawyer’s ability to represent clients (e.g., by making a device or important piece of software excessively difficult to use).”¹³ Moreover, . . . the

transmission of confidential client information, a lawyer is required to take “reasonable precautions to prevent the information from coming into the hands of unintended recipients.”¹⁴

The use of AI software and services directly implicates the lawyer’s duty to protect client confidences under MRPC 1.6(c). The risk of unauthorized access to confidential client information varies based on the artificial intelligence functions used. One of the most popular uses of artificial intelligence includes widely accessible generative AI programs, such as ChatGPT, Scribe, GPT-4, GitHub Copilot, and Claude. These platforms typically employ machine learning processes to create a variety of content, including audio, code, images, text, and videos, based on the input and requests of the user.¹⁵ Its use truly has the potential to transform the way we request and deliver legal services.

Although generative AI has the potential to revolutionize the practice of law, it poses some risks to protecting client confidentiality due to the nature of the programming employed. Relevant to lawyers, some generative AI platforms use what is known as large language models (LLM), which are “a specialized class of AI model that uses natural language processing (NLP) to understand and generate humanlike text-based content in response.”¹⁶ “These large models achieve contextual understanding and remember things because memory units are incorporated in their architectures. They store and retrieve relevant information and can then produce coherent and contextually accurate responses.”¹⁷ Essentially, generative AI platforms allow a user to input written prompts and instructions in order to generate highly specific and individualized content. One can imagine a situation where a lawyer inputs transcripts and interview notes and requests an output of discovery requests or other litigation documents.

The use of this generative AI technology by lawyers poses a confidentiality risk under MRPC 1.6(c). First, due to the nature of LLMs and NLPs, AI platforms collect and regularly use massive amounts of data to improve their processes and train their algorithms. When utilizing and inputting information into generative AI platforms, lawyers risk providing confidential information to unauthorized third parties, i.e., employees and developers. For example, “ChatGPT chat history is accessible and reviewable by ChatGPT employees. . . .”¹⁸ Legal commentators have opined that the use of these programs may “effectively waive attorney-client privilege.”¹⁹ It is important, however, to distinguish between the attorney-client privilege and the duty of confidentiality and the implications of a lawyer’s failure to meet the duties under each principle.

The attorney-client privilege is an evidentiary privilege that shields clients and lawyers from compelled disclosure of confidential communications. It is narrower in scope than the duty of confidentiality, and clients have the ultimate authority to determine whether to assert or waive privilege.²⁰ Ultimately, an inquiry regarding whether a lawyer waives a client’s privilege through the use of generative AI will likely be complex and fact-specific and turn on whether the lawyer’s use of generative AI is considered an authorized, unauthorized, or inadvertent disclosure of privileged communications.²¹ The answer to that question will likely be answered in the scope of an evidentiary dispute in the course of litigation, and does not resolve the question of whether the duty of confidentiality under MRPC 1.6(a) and (c) is being upheld.

Thus, the principal privacy risk when using generative AI tools likely lies with the duty to protect client confidences from unauthorized disclosure.²² Not only do generative AI tools store and utilize information inputs by users, but the platforms have also been vulnerable to data breaches and data leaks.²³ For instance, in 2023, ChatGPT experienced a data breach where users were able to see the chat history of other users.²⁴ Indeed, some bad actors are using generative AI to elevate and improve their cyber-attacks.²⁵ There was another instance where sensitive personal information of users was accessible to other users.²⁶ While these intrusions could be considered minor, some have noted that “the incidence could be a harbinger of the risks that could impact users and users in the future.”²⁷ Unsurprisingly, when asked whether a user’s input data is confidential, ChatGPT’s chatbot responded with: “OpenAI may collect

and use user data for research and privacy purposes, as described in its privacy policy.” It followed with: “To ensure the confidentiality of your data, it is important to follow best practices, such as not sharing sensitive personal information or confidential data when using AI models like me.”

If confidential client information of lawyers were to be accessed during one of these sophisticated intrusions or collapses of generative AI software security structures, it would certainly move the boundary line of what is considered “reasonable efforts to prevent the access or disclosure.”²⁸ It has the potential to be embarrassing to both the lawyer and the client and, if the disclosure is large enough, could result in lawsuits, financial consequences, and professional discipline.²⁹ It also can trigger an attorney’s duty to report a data breach, even if it involves the breach of a third-party vendor.³⁰ Therefore, when using generative AI programs, lawyers have a greater duty to inquire into privacy standards and policies and ascertain the methods by which input and output data are protected. Due to the risk of inadvertent disclosure of client confidences, lawyers should be very selective when utilizing generative AI platforms, and developers should be familiar with the needs of those providing legal services when developing and marketing their programming. Importantly, those who are tempted to input confidential client information to streamline tasks should ensure the platform’s compliance with MRPC 1.6(c) obligations and standards.

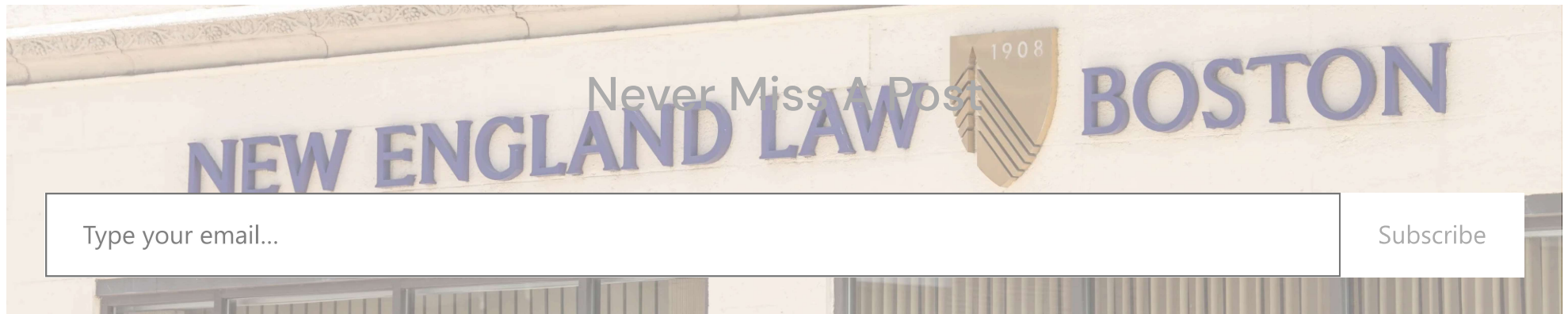


Chanal McCain

Professor McCain is a faculty fellow at New England Law | Boston.

Tagged in : [AI](#), [Artificial Intelligence](#), [Ethics](#), [Professional Responsibility](#)





[Home](#) [About](#) [Submissions](#) [Masthead](#) [Forum](#) [Blogs](#) [Lex Per Sonus](#) [Symposia](#) [Archives](#)

New England Law Review v 1.0