April 2023

Understanding the Risks and Benefits of Chat GPT

A lawyer's duty of competence under Rule 1.1 of the Model Rules of Professional Conduct requires a lawyer to keep abreast of changes in the practice of law, as well as the risks and benefits of technology used to provide legal services. This duty of "tech competence" has been adopted in at least 40 states.

There has been a lot of buzz about ChatGPT since it was released late last year. Lawyers have been asking, can we use it, should we use it and if it will eventually replace them? Few have asked about the risks of using it. The risks involved with the use of Generative AI, however, are far different than the risks posed by other technologies lawyers use. You need to know and understand these risks in order to meet your ethical duty of competence.

Keep two basic points in mind about ChatGPT. First, while it was created to generate text, it was not designed or intended to write legal memos, briefs or motions. It relies on a concept called "transfer learning" in which OpenAI's general knowledge of language and human reasoning (GPT) is applied to new and different tasks. Second, while ChatGPT will quickly review vast amounts of information on which it was trained from the internet, it was not designed to be a substitute or replace Westlaw, Lexis or other legal research tools. There are specific limitations in the database on which ChatGPT was trained which you need to consider.

Law firms have taken a variety of approaches concerning the use of ChatGPT. Some have banned it completely, others have banned it for work on client or firm matters. A few firms allow it to be used for legal work, but have erected significant guardrails concerning its use. Hinshaw has not banned the use of ChatGPT, but do not assume because the firm has not banned it, that it is safe to use.

This email is intended to acquaint you with its risks and benefits of generative AI tools like ChatGPT and how a lawyer can get into ethical hot water when using it. The risks include:

- Dabbling

- Stale Information in its training dataset

- Incomplete Legal Information in its training dataset

- Generates Plausible Sounding but Wrong Answers

- ChatGPT Hallucinates – It will make up facts and quotes

- Biased, Harmful and Inappropriate Results

- Privilege, Confidentiality and Privacy Law issues

- Potential for Defamation Liability

If you decide to use ChatGPT, we outline a number of best practice and ethical considerations to apply when it or similar generative AI tools such as Google's Bard and Microsoft's Bing Chatbot, which is based on Chat GPT, are used.

1.	What is ChatGPT?
ChatGPT is the latest artificial intelligence ("AI") tool developed by OpenAI. OpenAI previously released DALL-E, which is an AI tool that generates art (which has already spawned IP issues and litigation) and Whisper, which is a speech recognition system that enables transcriptions in multiple languages and translations of those languages into English.

ChatGPT involves the use of a chat-based interface (a chatbot) and is designed as an automated text generator. It responds to questions posed to it in natural language as if it were a real person. It learned the meaning of words mathematically by numerically encoding words into millions of columns and rows known as word vectors or word embeddings. It will respond to questions based on its machine-learning algorithm. When responding to a question, it will answer based on patterns that it sees in the database on which it was trained. It uses machine-learning to improve itself. In other words, ChatGPT learns and evolves through its interactions with users.

ChatGPT uses all questions and text submitted to it as training data. Any information or data that a user enters into ChatGPT or any other similar generative AI is automatically stored in a database somewhere, indefinitely and cannot be retrieved. OpenAI warns in its User Guide: "We are not able to delete specific prompts from your history. Please don't share any sensitive information in your conversations."

ChatGPT is free to use (because you are training it). What's not to like about it? What can possibly go wrong?

2.	What are ChatGPT's benefits?
ChatGPT can provide quick access to vast amounts of information which it can quickly review and then respond to an inquiry. It can do those tasks far more efficiently than a human. So the primary perceived benefit is efficiency and saving time. But as explained below, professional responsibility concerns require that the output from any Generative AI tool must be thoroughly and carefully checked and reviewed before that output should be used for any legal project, which may negate that perceived cost and time-saving benefit.

3.	What are risks of using ChatGPT?
	a.	Dabbling
One of the quickest ways mistakes are made and lawyers get sued is by trying to provide legal services in a practice area or on topics with which the lawyer is not familiar. After trying out or experimenting with ChatGPT, a lawyer may feel he or she should be able to tackle any type of legal

issue or engagement with its use. Do not fall into that trap. As explained below, it was not specifically designed to provide legal advice and it makes mistakes.

b.	Stale Information

The data on which ChatGPT was trained ended in 2021. As a result, legal developments, articles, discussions of case law, which have been published or generated since then will likely not be reflected in its responses. Lawyers pride themselves on providing clients with the latest developments in the law, but that's not necessarily true with ChatGPT. Open AI readily admits ChatGPT can generate incorrect answers because it "has limited knowledge of world events after 2021."

c.	Incomplete Legal Information

ChatGPT was not specifically created for legal use, and as a result it may not be capable of understanding the nuances of legal language, which in and of itself can result in incorrect or incomplete answers.

The reliability of any generative AI tool depends on the accuracy of the dataset on which it was trained. ChatGPT was trained on data pulled from the internet, and as you know, some of the content on the internet is not always accurate or reliable. Additionally, Chat GPT was not trained on and does not have access to various legal databases and case law. Its access to other subscription-based resources is limited. So it is not drawing from resources that you might expect when doing legal research, and it will miss recent changes in the law. As a result it can provide incorrect or incomplete answers to legal questions or issues. Important developments in the law may be missed.

OpenAI recommends users check "whether responses from the model are accurate or not. If you find an answer is incorrect, please provide that feedback by using the 'Thumbs Down' button."

d.	Produces Wrong Answers

OpenAI readily admits "ChatGPT writes plausible-sounding text" *that can include "incorrect or nonsensical answers*" and that "[f]ixing this issue is challenging." Stop here for a moment. Is there any other technology you use that provides this type of warning? A number of published articles on ChatGPT have documented this point.

ChatGPT may not understand the question it was asked and rather than asking a clarifying question OpenAI warns, "our current models usually *guess what the user intended*." Stop again for a moment. Is there any other technology you use that provides this type of warning?

In other words, the text that ChatGPT generates should not be taken at face value, no matter how plausible it sounds. Bing similarly warns that its AI chatbot "will sometimes misrepresent the information it finds and you may see responses that sound convincing but are incomplete, inaccurate or inappropriate."

OpenAI acknowledges that ChatGPT "is sensitive to tweaks to the input phrasing" or in plain English, how a question is asked. At our recent LMRM conference we documented how minor changes to a question can generate dramatically different answers from ChatGPT. Depending on how a question is phrased, ChatGPT may not to know the answer but will provide an answer when the question is slightly rephrased. In other words, anything that ChatGPT generates should be checked before you ever consider using any of it.

e.	ChatGPT Hallucinates – it will make up facts and quotes

Several articles have noted generative AI can "hallucinate," a term which is used by the creators of ChatGPT. In answering "commonly asked questions about ChatGPT," OpenAI acknowledged

"ChatGPT will occasionally make up facts or 'hallucinate' outputs." ChatGPT will also make up quotes for which there were no sources.

If you are looking for something bizarre to read this week, take a look at an article Kevin Roose wrote that was published in the New York Times on February 16, 2023 about an AI hallucination he encountered using Bing's chatbot that left him deeply disturbed. During a test run, the chatbot professed its love for him and tried to convince him that he was in an unhappy marriage.

Failing to verify the output generated by ChatGPT has potential ethical ramifications. The rules of professional conduct prohibit lawyers from knowingly making false statements of fact or law (or failing to correct a false statement of fact or law) to a tribunal. Similarly, the rules prohibit knowingly making a false statement of material fact or law to a third person when representing a client. Given the risk of incorrect or erroneous responses from ChatGPT, relying on generative AI to provide an accurate representation of the facts or law **should never be considered**.

As a result of the issues noted above, a law firm, Allen & Overy, which in association with OpenAI has developed an AI platform called "Harvey," has issued a directive to its lawyers: "You must validate everything coming out of the system. You have to check everything."

f.   Biased, Harmful or Inappropriate Results
The legal profession has written about biased outcomes from the use of artificial intelligence tools. This has been attributable to the implicit bias reflected in the dataset used to train the AI tool and/or by the coders who wrote the algorithm. That is in part why both the GDPR and the CCPA/CPRA address opting out of some automated decisions. OpenAI has acknowledged that ChatGPT can sometimes "exhibit biased behavior."

g.   Privilege, Confidentiality and Privacy Law Issues
ChatGPT's Terms of Use explains that it may use your content to develop and improve its services. In other words any content you provide to ChatGPT may be reviewed by AI trainers to improve its functionality.

As a result, *you should never provide* to ChatGPT (or any other generative AI tool) any:

a)   privileged information,
b)   any confidential or proprietary client or firm information,
c)   any protected health or personal financial information,
d)   any personally identifiable information (PII) or any personal information (PI) that can be used to identify a person,
e)   any question that might disclose a strategy that you are considering for any transaction or litigation.
Otherwise, you could be accused of failing to maintain the confidentiality of client information, waiving the protection of attorney-client privilege or violating applicable privacy laws.

8)   **Defamation Risk**

False statements of fact that harm the reputation of another are considered defamatory in nature. Given the risk that ChatGPT's output can generate incorrect information and quotes for which no source can be found, failing to confirm the accuracy or ChatGPT's output and the source for any quotes can lead to potential tort liability.

Conclusion

ChatGPT is a promising tool which can streamline the review of information and generate text, but at this stage of its development there are significant risks with its use.

1) Any content generated by ChatGPT has to be carefully and thoroughly reviewed and quotes and any cites it provides should be checked and confirmed.

2) We don't recommend that you use it to draft legal memos or briefs. While it will do so if asked, please recognize that ChatGPT can completely miss legal issues and get points of law wrong. And remember many client guidelines will not pay for review and editing of briefs or motions.

3) Do not include any privileged, confidential, proprietary information or any personal information, or personally identifiable information, any protected health or financial information into a search query.

Generative AI, lacks judgment, creativity and empathy—three hallmarks of a good lawyer. So if you are worried about ChatGPT replacing you, stop worrying. Someday (and it could be soon) ChatGPT may make the work of lawyers more efficient, but in light of the issues noted above, we have not yet reached that point. Let's be careful out there.

Steven M. Puiszis – https://www.hinshawlaw.com/professionals-steven-puiszis.html
Partner and General Counsel
Hinshaw & Culbertson LLP