

AI Product Review for Solo Practitioners and Small Firms: ChatGPT from OpenAI

By Keegan McMullan, Timothy Wiseman, Andrew Switlyk

Introduction

ChatGPT by OpenAI is almost synonymous with the new era of AI products. These products are powered by the latest advancements in generative AI models, which can generate text, images and video in response to user provided contexts – usually in the form of text prompts and document or image files. When ChatGPT launched to the public at the end of 2022, it introduced its chat user interface as the way users interact with the various models OpenAI develops and releases. Since then, the interface added additional features and abilities, and newer and more powerful models. However, its paradigm remains unchanged, where a user “chats” with the model to leverage its abilities, knowledge base and reasoning. This interface is now widely adopted as an industry standard and emulated by ChatGPT’s competitors such as Anthropic (Claude) and Google’s Bard.

To explore how ChatGPT performs and its viability as a tool for small and solo law firms, members of the Nevada bar evaluated it using criteria and a scoring rubric that considers both its pragmatic usefulness in the practice of law as well as ethical and professional responsibilities. The evaluating members’ practice areas span transactional, litigation and appellate, and they used and evaluated ChatGPT in light of these diverse practice areas in an attempt to come to a broader overall conclusion.

After our evaluation, our key takeaways are that ChatGPT is excellent as a sounding board in situations such as brainstorming and can be useful as a start for research in unfamiliar subjects when the question is best posed in natural language. However, it is best to be cautious when providing sensitive information to ChatGPT, and its responses should be meticulously fact-checked to mitigate the risks of hallucination or confabulation. The practitioner should also refrain from sharing any potentially sensitive or confidential information with it.

Each evaluation category is discussed below.

Accuracy & Reliability

Score: (3.2/5.0)

ChatGPT includes a clear static disclaimer within its interface's footer that "ChatGPT can make mistakes. Check important info." Based on paces the evaluators put ChatGPT through, this warning should be heeded by lawyers. Although, it often provided generally accurate or mostly accurate information in regard to high-level queries, it also frequently produced confabulated or hallucinated responses that were completely or mostly incorrect. The answers are usually superficial appeared correct and are frequently partially incompletely cited or not referenced if at all. For superficial or general inquiries, or as a starting point, ChatGPT could be a time saver and preferable than starting at a search engine. Also, if additional context was given, such as pdfs of cases, law, pleadings or manuals, the answers tended to be better or more constrained. However, there were many occasions where ChatGPT confidently cited references incorrectly or even fabricated them altogether. It was rare to not have results that included both accurate and completely inaccurate references. This means ChatGPT answers should never be relied upon without strict verification. As a result, using ChatGPT for legal research can often take the practitioner more time than traditional means since every answer and citation needs to be meticulously checked. In our experience, for more complex queries the erroneous or not useful references produced made us hit that time wasting threshold often.

Outside of legal research, ChatGPT did an effective job with drafting initial generic letters and summarizing and helping parse through materials. You can feed a limited number of pdfs or other document formats into it and query against those, which was a highly productive use.

Data Privacy & Security

(2.0/5.0)

OpenAI offers a few product tiers which have varying levels of data privacy and security. The enterprise and team offerings provide the customer the most granular and transparent controls over their data, with some even offering zero data retention options. However, these aren't appropriate for the solo or small firm use case due to their high initial costs or minimum seat requirements. So, our evaluation focused on either the free or plus options of ChatGPT.

By default, user data is shared with OpenAI and may be used to train their models. This can be disabled via a user's privacy portal. Also, OpenAI's privacy policy and data handling appears to comply with the requirements of NRS 603A and the California

Consumer Privacy Act (CCPA). It does not comply with HIPPA, and to get this heightened level of compliance would require a separate Business Associate Agreement with OpenAI that is not available to non-enterprise customers.

However, beyond these laws' requirements in terms of the retention, transmission, and deletion of user data, OpenAI's privacy policy contains fairly broad usage terms regarding the company's use and disclosure of personal data. The goals of these terms, which include improving OpenAI's services, conducting research, developing product features or assisting in business operations, are clearly opposed to the goals and responsibilities of lawyers to keep client's information and communications privileged and private. Given this, we would recommend against using ChatGPT with any non-public or sensitive client information or concerns.

Bias & Fairness:

(3.0 /5.0)

There have been reports by users of bias and fairness concerns by ChatGPT from all sides of the political spectrum. These issues have been attributed to potential issues with the training data, moderation policies, reinforcement learning policies or other decisions made within the company that influence the results provided by ChatGPT. To address these concerns, ChatGPT provides information on how attempts to address biases when developing its models and its guidelines that control the chat when tricky or controversial topics are involved. They also subject themselves to external reviews and audits to evaluate the model's behavior here.

During our review and usage, the majority of our legal focused topics were non-ideological or non-opinion questions, and they didn't produce patently biased chat results. Our main concern was when we experience model hallucinations, wherein the model produces a result or part of a result that is clearly inaccurate - and the model had an indifference toward the truth. This can produce an unfair result if the model is representing something or someone in an inaccurate way. We experienced this fairly often, which makes manually verifying AI results a necessity before relying on them.

Ethical & Legal Implications

(3.0/5.0)

A lot of the recent concerns and discussions regarding the ethical and legal implications of AI, by us and within some of the broader legal community, revolve around

the lawyer's duty of competence and of confidentiality. Based on the privacy and security discussion above, we cannot recommend using ChatGPT with any confidential client information or work product. Also, as cautioned within the accuracy and reliability section above, any legal research done with ChatGPT must be thoroughly verified. Instances of lawyers who blindly used ChatGPT outputs have infamously led to them being sanctioned for including fake cases or citations in their court pleadings for acting in bad faith. To avoid similar ignominy and to stay on the right side of the minimal competence bar, any ChatGPT legal research output should be assumed unreliable and requiring verification and follow-up research. Being aware of these limitations will help avoid ethical pitfalls.

Integration & Usability

(3.5/5)

In terms of usability, the platform is an industry benchmark. It is cloud based, has a polished and straightforward UI and is highly reliable in terms of availability. It's also continually adding product features that just work, such as adding files to the context window, external web searches as part of your query and improved reasoning capabilities like with the o1 model. However, in terms of integration, there really is none. ChatGPT is a standalone chat platform and does not integrate with external software or systems.

Support & Training

(3.0/5)

In terms of product ease and intuitiveness, ChatGPT is straightforward and for basic consumer usage probably requires little support. For free or plus users, there is no personal support or training offered by OpenAI but they have a decent help center with articles that range from using the product to billing inquiries. The articles are geared toward generic usage and there isn't anything that is legal industry specific in terms of support or recommended best practices.

Cost Effectiveness

(4.1 / 5)

For non-enterprise users, there's free, plus individual plans of \$20.00 per month and business / team plans for \$25.00 per month per user (billed annually). They also released a recent "pro" plan for \$200.00 per month which give the subscriber unlimited access to their

most advanced features and models like o1. We were able to evaluate o1 with the plus plan, and the model at this time doesn't provide a significant improvement for our legal use cases over GPT-4o. Therefore, the plan we recommend and where we derived the most value was from either the free or plus ones. For brainstorming, summarizing, parsing public documents and drafting help it's an affordable price point.

Compliance & Regulatory Impact

(2.9 / 5.0)

OpenAI adheres to a few compliance standards and laws, such as SOC 2 Type 2 Compliance, CSA STAR level 1, GDPR and CCPA. These levels are below services such as Microsoft 365 cloud, which comply with more stringent levels of assurances in terms of cloud security and system and organizational control standards. However, OpenAI's compliance to data security is on-par or above its common competitors. Also, as discussed above, HIPPA compliance is available but only for enterprise clients and with a business associate agreement. For non-confidential topics or work product, the compliance frameworks OpenAI adheres to should be sufficient. If you have specialized use-cases or work with data subject to specific regulations, then you should enquire about heightened privacy offerings.

Innovation & Scalability

(3.0/5.0)

ChatGPT continually updates its product and is usually the market leader, for general chat AI offerings, in terms of features and model ability. During our evaluation period, OpenAI released or upgraded its models and released a handful of product capabilities, such as its canvas feature which is useful for drafting or editing a piece of writing. The product cannot be scaled beyond its current abilities, which have strict hard limits as to its context window and size and number of files that can be attached to a conversation. Also, when different versions of models are asked identical legal questions, the answer quality of the more recent or advertised more advanced models did not improve that much, if at all.

User Feedback & Market Reputation

(4.0/5.0)

ChatGPT is the market leader and is the best-known generative AI product available. Some of its biggest criticisms pertain to models have old information and their tendency to “hallucinate”. Product features such as integrated web search was added to address some of those concerns. The general feedback fluctuates depending upon OpenAI’s latest model release as compared to competitors. However, OpenAI tends to remain consistently at the top in terms of performance and remains in the lead for adoption.

Total Score: 31.7 (Satisfactory)