Wolters Kluwer

Strategic Perspectives

Dos and Don’ts for GenAI in Expert Work

*By Collin Starkweather, Izzy Nelken, and Jey Kumarasamy**

You would have to have been hiding under a rock not to have noticed the ongoing tectonic shift in legaltech due to the advent of generative AI technologies that have brought equal parts anticipation and trepidation to the legal profession.

A wide variety of legal professionals are actively moving to adapt their practices to the rapidly changing landscape. A recent Wolters Kluwer survey of legal professionals indicated that “73% of lawyers expect to integrate generative AI (GenAI) into their legal work in the next year.”[1]

Adapting to disruptive technology can be fraught with pitfalls, and the legal profession is in some ways particularly vulnerable. Recognizing this, some courts have required that parties to litigation include a declaration in all documents for which AI was used to generate content.[2]

While a great deal of ink has been expended on the pitfalls of the use of GenAI by legal professionals, less attention has been devoted to its use by expert witnesses, though the potential pitfalls may be even greater for expert testimony.

Here we will discuss potential applications of GenAI in expert witness work based on the current state of the art and anticipated near-term development in the technology.[3] We will then provide a set of recommendations on best practices for the use of GenAI for expert witness work.

Potential Applications of GenAI in Expert Witness Work

Expert witnesses are retained to provide opinions to assist the trier of fact on matters relevant to legal issues which require specialized knowledge. Experts will often perform analysis that serves as the basis for their opinions, and expert testimony will often include a report, or multiple reports, submitted to the court summarizing their opinion as well as any supporting analysis and details of associated methodology. The expert may be expected to testify in a deposition or at trial in lieu of a settlement.

Applications of GenAI are possible at all stages of expert witness workflows short of testimony. In the initial stages of an engagement, an expert may be provided with discovery production or other materials for review. Experts may be provided with the documents directly or via an e-discovery document management system.

Incumbent providers of e-discovery and related document management systems such as Relativity and Casepoint have moved quickly to incorporate GenAI into their feature set, while more recent market entrants such as DISCO developed their

---

* Dr. Collin Starkweather is the Director of Data Science and Jey Kumarasamy is a Senior Associate at Luminos.Law, where they assist clients to manage the privacy, fairness, security, and transparency of their AI and data. Dr. Starkweather is also the Chief Data Scientist at a stealth AI startup company. Dr. Izzy Nelken is the founder of Super Computer Consulting, Inc. and has served as a consultant or expert witness on numerous securities and other financial matters. The views expressed herein are solely those of the authors and do not necessarily reflect the views or positions of LuminosLaw, Super Computer Consulting, Inc., or their clients or affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice. No generative AI was used in the drafting of this article.

1 Wolters Kluwer, “2023 Future Ready Lawyer Survey Report,” October 25, 2023, available at https://www.wolterskluwer.com/EN/know/future-ready-lawyer-2023.

2 JD Supra, “Court-mandated AI disclosures: what you need to know when using AI in court filings,” July 13, 2023, available at https://www.jdsupra.com/legalnews/court-mandated-ai-disclosures-what-you-8934776/.

3 Technically speaking, anyone who performs a simple internet search or uses a spelling or grammar checker uses advanced AI technologies. This article focuses on recently-developed GenAI technologies, primarily large language models (LLMs) such as OpenAI’s GPT, Anthropic’s Claude, or Meta’s Llama models and associated technologies and toolsets.

products around recent advances in cloud, data analytics, and AI technology.[4] These products are utilized by many expert witnesses, especially in complex litigation matters involving a substantial volume of e-discovery, often through access provided by counsel.

Experts would largely employ use cases similar to those of legal professionals, including document search and summarization tasks, so many of the best practices and potential pitfalls that legal professionals should have on their radar would also apply to experts. Pitfalls would include "hallucinations,"[5] which legal, risk, and compliance professionals surveyed recently by Deloitte described as a threat that poses "the greatest challenge to their organization's legal workflows."[6]

There are a variety of other simpler, but potentially efficiency enhancing, use cases centered around generation of language-based content, including grammar review and suggestions for improving stylistic elements and document flow.[7] Language models can be particularly helpful for multilingual document sets, facilitating document translation[8] or even language-agnostic search.[9]

Currently, some tools are available via GenAI models that accommodate analytical tasks, such as sentiment analysis or analysis of spreadsheet content,[10] but these are currently relatively nascent technologies that remain limited in scope. Going forward, the development of more robust analytical GenAI applications will result in a proliferation of practical use cases for analysis. These developments are

4 More generally, substantial resources have been devoted to the development of products based on generative AI by established firms, while startup firms building products based on generative AI technologies have rapidly proliferated. See, e.g., The LegalTech Fund, "Early-Stage LegalTech Generative AI Landscape: Oct 2023," available at https://www.linkedin.com/feed/update/urn:li:activity:7130561732288106496/.

5 "Hallucinations" are model output that is inaccurate or nonsensical, and can be subtle, pernicious, and difficult to both detect and remedy. Cassandre Coyer, "Hallucinations Are Legal's Main Concern With Generative AI, But Maybe Not For Long," Legaltech News, November 27, 2023, available at https://www.law.com/legaltechnews/2023/11/27/hallucinations-are-legals-main-concern-with-generative-ai-but-maybe-not-for-long/. For example, in a particularly well-publicized example, attorney Steven Schwartz was sanctioned for filing a legal brief in a personal injury case with input from ChatGPT including six non-existent court decisions. Sara Marken, "Lawyer who cited cases concocted by AI asks judge to spare sanctions," Reuters, June 8, 2023, available at https://www.reuters.com/legal/transactional/lawyer-who-cited-cases-concocted-by-ai-asks-judge-spare-sanctions-2023-06-08/. In another, more recent highly-publicized example, a brief filed by attorneys for Michael Cohen was found to have nonexistent citations. Nate Raymond, "Ex-Trump fixer Michael Cohen says AI created fake cases in court filing," Reuters, December 29, 2023, available at https://www.reuters.com/legal/ex-trump-fixer-michael-cohen-says-ai-created-fake-cases-court-filing-2023-12-29/.

6 Cassandre Coyer, "Hallucinations Are Legal's Main Concern With Generative AI, But Maybe Not For Long," Legaltech News, November 27, 2023, available at https://www.law.com/legaltechnews/2023/11/27/hallucinations-are-legals-main-concern-with-generative-ai-but-maybe-not-for-long/.

7 For example, writing assistance tools such as Grammarly use AI to primarily correct grammar and spelling, and provide stylistic suggestions. However, even some of these tools are quickly starting to include document generation features which may entice current users to expand their scope of reliance on such tools beyond simple editing tasks. See, e.g., Grammarly, "AI Writing Tools," available at https://www.grammarly.com/ai-writing-tools.

8 For example, the Google Translate tool has an option for "Search Generative Experience" ("SGE") which uses Generative AI. See Molly McHugh-Johnson, "How we taught Google Translate to recognize homonyms," Google, November 10, 2023, available at https://blog.google/products/translate/google-translate-homonyms/.

9 In addition to AI assistants being provided by producers of general tools for language-based content, such as Microsoft 365 Copilot, specialized tools with domain-specific features are being rapidly introduced, such as Casetext's CoCounsel "AI legal assistant" and LexisNexis's Lexis Snapshot and Lexis Create in the legaltech space, open source tools such as Workflow are being developed. See, e.g., Jared Spataro, "Introducing Microsoft 365 Copilot – your copilot for work," Microsoft, March 16, 2023, available at https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/; Matt Reynolds, "Thomson Reuters announces new AI initiatives and CoCounsel integration," ABA Journal, November 15, 2023, available at https://www.abajournal.com/web/article/thomson-reuters-to-launch-cocounsel-integration-after-650-million-casetext-buyout; Rhys Dipshan, "LexisNexis Expands Generative AI Offerings With Lexis Snapshot and Lexis Create Integration," Legaltech News, November 14, 2023, available at https://www.law.com/legaltechnews/2023/11/14/lexisnexis-expands-generative-ai-offerings-with-lexis-snapshot-and-lexis-create-integration/; and the Workflow GitHub repository at https://github.com/poloclub/wordflow/.

10 An example would be Microsoft's Analyze Data feature in Excel. (See https://support.microsoft.com/en-us/office/analyze-data-in-excel-3223aab8-f543-4fda-85ed-76bb0295ffc4.) This and similar tools could easily be (and in all likelihood will be) incorporated into multimodal models for the analysis of document repositories or other document sets, such as email or messaging archives with attachments.

expected to prove particularly impactful to experts, with continued enhancement of LLM toolsets[11] and multimodal LLMs ("MLLMs")[12] expected to accelerate the pace of development of GenAI analytical capabilities.

As analytical capabilities advance, LLMs may eventually be able to accommodate queries with complex analytical requirements such as: "Identify all employees of the accounting department, then produce a network graph of all email exchanges between accounting staff with nodes representing email addresses, node sizes representing the total number of emails sent and received, and edges labeled with the number of emails exchanged between the respective nodes;" or "List the Bates stamps of any documents that appear to be work papers for the balance sheet released in the fiscal year 2023 10-K filing, and based on those documents, plot the total liabilities reflected in the work papers over time based on the dates of the work papers clearly distinguishing actual from projected figures."

In some cases, as analytical methods become more accessible, lawyers may be tempted to perform some analytical work that would normally be performed by experts. However, attorneys should exercise caution in their use of such applications lest they inadvertently introduce assertions that may later prove problematic for their experts, especially with complex litigation with arguments that rely on domain expertise and sophisticated numerical analysis.

Caveats in the Application of GenAI to Expert Analysis

When using GenAI for analytical work, a host of caveats apply above and beyond those associated with more traditional analytical tools.

The use of GenAI tools does not excuse an expert from looking "under the hood." Because GenAI can introduce an additional degree of separation between analytical results and the underlying data, if anything, there is an even greater need to understand the details that form the basis of any analytical results derived from GenAI that contribute to an expert's opinions.

As discussed earlier, at present, GenAI is notorious for issues with reliability, and the court's confidence in the reliability of an expert's testimony is integral to the expert's ability to persuade the trier of fact of the merits of his or her opinions. The use of GenAI may not only introduce potential issues with accuracy or reliability, but may also provide inroads for adversarial parties to call into question the integrity of the expert's analysis regardless of its merits.

In order to defend his or her opinions, an expert may be required to produce backup materials for any supporting analysis.[13] The backup materials should be sufficient for opposing experts to be able to reproduce the expert's analytical results.[14] Without a proper understanding of the limitations of any GenAI tools employed and the appropriate configuration of those tools, even the production of prompts and other supporting documentation may be insufficient to ensure reproducible results.[15]

---

11 In the context of LLMs, tools are functionality that is available to the LLM to perform specific tasks, such as a calculator that allows an LLM to perform mathematical operations that the LLM itself is unable to perform accurately. See, e.g., Yujian Qin et al., "ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs," arXiv, October 3, 2023, available at https://arxiv.org/abs/2307.16789.

12 Multimodal LLMs combine textual capabilities of LLMs with other modes of information processing, such as optical character recognition ("OCR") or image and video processing. See, e.g., Shukang Yin et al., "A Survey on Multimodal Large Language Models," arXiv, June 23, 2023, available at https://arxiv.org/abs/2306.13549.

13 For example, Rule 26(a)(2) of the Federal Rules of Civil Procedure specifies that experts must include "the facts or data considered by the witness in forming" their opinion in their report. Similarly, Rule 705 of the Federal Rules of Evidence provides that on cross-examination an expert may be required to disclose the facts or data underlying their opinion. See, e.g., "Rule 26. Duty to Disclose; General Provisions Governing Discovery," Cornell Law School Legal Information Institute, available at https://www.law.cornell.edu/rules/frcp/rule_26; "Rule 705. Disclosing the Facts or Data Underlying an Expert," Cornell Law School Legal Information Institute, available at https://www.law.cornell.edu/rules/fre/rule_705; Andrew McLure Toft, "Expert Disclosures Must Include Facts or Data 'Considered,' Even if Not 'Relied On'," ABA Journal, January 31, 2018, available at https://www.americanbar.org/groups/litigation/resources/newsletters/pretrial-practice-discovery/expert-disclosures-must-include-facts-or-data-considered-even-if-not-relied/.

14 "The foundation of Daubert is being able to demonstrate that the expert's opinion by using sufficient data that utilizes reliable principles and methodologies that can be reproduced by independent testing. If another expert cannot replicate your expert's analysis, you will not survive a Daubert challenge. The importance of this cannot be overstated." Ebony S. Morris, "Daubert-Proofing Your Expert," ABA Journal, December 29, 2021, available at https://www.americanbar.org/groups/litigation/resources/newsletters/mass-torts/daubert-proofing-your-expert/.

15 LLMs typically use hyperparameters which govern the degree to which randomness is introduced in the generation of model output. For example, temperature is a hyperparameter that may be used to increase or decrease randomness in model responses, with a higher temperature reflecting greater randomness, but also greater creativity, in a model's response. (See, e.g., Ayesha Saleem, "How to tune LLM Parameters for optimal performance," Data Science Dojo, September 11, 2023, available at https://datasciencedojo.com/blog/llm-parameters/.) While reproducibility may be possible where such hyperparameters are set to ensure no randomness in the model response, those hyperparameters may not be exposed to users or even desirable to use at a setting that does not admit randomness, depending on the use case.

Some GenAI tools may even lack the ability to produce replicable results.[16]

The use of GenAI may even expose experts to potential Daubert or Rule 702 challenges.[17] Among other things, the Daubert standard applied to determine the admissibility of expert witness testimony specifies that any technique or theory employed by the expert "can be, and has been" tested as well as have a "known or potential error rate." Given the "move fast and break things" culture of AI generally, these and related criteria need to be considered when employing GenAI for expert work.

In addition, experts are often bound by protective orders or NDAs. As discussed further below, some GenAI systems may incorporate user-provided information, including documents and data and even the prompts themselves, into responses provided to other users, putting confidentiality and potentially even privilege at risk.[18]

Best Practices for the Use of GenAI by Expert Work

Expert witnesses looking to avoid falling afoul of these potential pitfalls from the use of GenAI should consider the following recommendations for the use of GenAI.

First and foremost, experts should be transparent about their use of GenAI. Even in jurisdictions which do not require disclosure of the use of GenAI, experts should consider getting ahead of the ball and proactively provide disclosures. It may serve to resolve any questions that might otherwise not be explicitly articulated by the court, and where questions are not articulated by the court, experts may very well find those questions articulated by opposing counsel. For example, it is not uncommon for experts to be asked while giving testimony whether they formed or drafted the opinions expressed in their reports themselves, and it is not much of a leap to extend that line of inquiry to the use of GenAI.

In order to properly disclose the use of GenAI, experts must be aware of how it is being used. Especially with complex litigation, where experts are often supported by a sizable team which itself may be coordinating behind the scenes with a legal team, there may be less visibility into exactly how the sausage is made.[19]

An expert's junior staff may use GenAI tools for a variety of tasks such as document search and summarization and will be increasingly using it for tasks involving numerical analysis such as those discussed above. Inexperienced staff may be unaware of potential risks associated with content generated by GenAI or the broader context of their work, where the veracity of even otherwise seemingly benign assertions is integral to the efficacy of the expert's testimony. They also may be more prone than more experienced practitioners to attribute unwarranted credibility to content that may appear perfectly plausible, be presented with a confident tone, and even be defended by the GenAI itself when prompted.[20]

The expert may rely on the integrity of the work product produced by their staff, and for practical or even strategic reasons be unaware of how the work product is produced at a high level of detail. Where the expert exercises such reliance, they would be well advised to clearly communicate expectations regarding the use of GenAI to engagement managers and other senior staff in supervisory roles.

---

16 For example, with proprietary tools that are not configurable by the user, such as ChatGPT, there is no guarantee that any given prompt will produce identical output when posed a second time. Charles Ross, "Does ChatGPT Give the Same Answer to Everyone?," Medium, March 20, 2023, available at https://medium.com/@charles-ross/does-chatgpt-give-the-same-answer-to-everyone-521e3e9355a4.

17 Rule 702 of the Federal Rules of Evidence, for example, requires that an expert witness's testimony be demonstrated to be more likely than not, among other things, "the product of reliable principles and methods" and that it "reflects a reliable application of the principles and methods to the facts of the case." Non-trivial uses of GenAI in preparing an expert's opinion may limit the proponent's ability to overcome an admissibility challenge under this standard. See, e.g., "Daubert Standard," Cornell Law School Legal Information Institute, available at https://www.law.cornell.edu/wex/daubert_standard; and "Rule 702. Testimony of Expert Witnesses," Cornell Law School Legal Information Institute, available at https://www.law.cornell.edu/rules/fre/rule_702.

18 For example, "ChatGPT saves all of the prompts, questions, and queries users enter into it, regardless of the topic or subject being discussed." Aaron Drapkin, "Does ChatGPT Save My Data? OpenAI's Privacy Policy Explained," Tech.co, June 29, 2023, available at https://tech.co/news/does-chatgpt-save-my-data.

19 The results of a Salesforce survey of over 14,000 global workers indicated that 55% of those who had used GenAI tools at work were not using approved GenAI tools and 40% were even using banned tools. Moreover, 64% had passed off content produced by GenAI as their own. "More than Half of Generative AI Adopters Use Unapproved Tools at Work," Salesforce, November 15, 2023, available at https://www.salesforce.com/news/stories/ai-at-work-research/.

20 Alternatively, GenAI models may even agree that they have produced factually incorrect responses when in fact those responses are correct. Tatyana Woodall, "ChatGPT often won't defend its answers – even when it is right," Ohio State News, December 7, 2023, available at https://news.osu.edu/chatgpt-often-wont-defend-its-answers--even-when-it-is-right/.

In addition to performing more sophisticated analytical work, senior staff may also contribute to drafting of the expert's report. In some cases, they may even be responsible for substantively drafting the report itself, though the expert will typically be responsible for drafting the language of his or her opinions. Counsel may contribute language to a report, often indirectly through comments or suggestions where direct contributions would be subject to challenge.

Any of the parties that contribute language to an expert report may use GenAI for various drafting functions, where in addition to the various AI assistants, an ever-increasing number of GenAI add-ins and other tools are becoming available for word processors such as Microsoft Word and Google Docs. These tools may be as readily accessible as the spell checker,[21] and may be employed without the direct knowledge of the expert or any obvious indication that those tools have been used. However, even where there are no obvious indications that GenAI has been used in drafting, sophisticated readers are increasingly able to discern AI-generated content,[22] and GenAI tools for that very purpose are proliferating.[23]

The expert should also be aware that the use of GenAI may even inadvertently introduce risks in certain contexts over and above those widely reported for publicly-released tools such as ChatGPT. As large law and litigation consulting firms and their technology vendors become increasingly capable, they will be incentivized to exploit their in-house corpora of documents and data. We anticipate efforts to fine tune or otherwise enhance publicly available models to create proprietary tools to augment their service offerings. Without rigorous controls in place, cross-contamination in which content from one matter may find its way into another in a way that proves problematic, may give rise not just to reputational risks, but risks associated with confidentiality, conflict of interest, and other process considerations.

Experts should be particularly wary of applications of GenAI to numerical work or analytical tasks requiring complex reasoning. While the capabilities of GenAI are rapidly advancing in these areas, they are not yet ready for prime time for most use cases. Particular use cases may be reevaluated as the technology develops, but for the time being, experts should be careful not to rely on analysis contributed by GenAI, particularly numerical analysis, that has not been thoroughly vetted by a human.

Many of the controls an expert would be advised to put in place are not necessarily specific to GenAI, but even greater diligence should be exercised where there are substantive contributions by GenAI. For example, rigorous internal audits of numerical and other analyses should be performed and all citations should be manually checked against source material by humans prior to finalization of the report. These controls should be in place regardless of the use of GenAI, but when GenAI has been employed for drafting, supervisory staff should be especially thorough in ensuring the rigorous performance of these audits.

Finally, experts should eschew the use of GenAI for any activities related to the drafting of their opinions. Even if its use is innocuous and would not substantively alter the opinions as conceived by the expert, such as checking grammar on a section of the document including the opinions, it may open the expert to potential challenges.

Conclusion

It is still early days for the adoption of GenAI for expert work, with little guidance from courts on boundaries for its usage. It will eventually find its way into expert work, whether through application by the experts themselves or by support staff, so experts should continue to actively monitor developments in the field. Experts should exercise caution when relying on content produced by GenAI and take steps, such as educating staff and implementing process controls, to insulate themselves from potential challenges that may arise from its usage.

---

21 See, for example, Talarian's GPT for Excel Word at https://appsource.microsoft.com/en-us/product/office/wa200005502; Talarian's GPT for Sheets Docs at https://workspace.google.com/marketplace/app/gpt_for_sheets_and_docs/677318054654; Smart Barn's Ghostwriter for Excel, PowerPoint, and Word at https://appsource.microsoft.com/en-us/product/office/wa200005107; GhostWryter's GhostWryter for Google Docs at https://workspace.google.com/marketplace/app/ghostwryter/24175974325.

22 See, e.g., James Presbitero Jr., "These Words Make it Obvious That Your Text is Written By AI," Medium, December 31, 2023, available at https://medium.com/practice-in-public/these-words-make-it-obvious-that-your-text-is-written-by-ai-9b04f399d88c; and Ellie LaPosha, "Decoding the Code: How to Spot an AI-Generated Resume ," Merito Group, August 16, 2023, available at https://www.meritogroup.com/decoding-the-code-how-to-spot-an-ai-generated-resume/; "iThenticate 2.0: Advancing research integrity with AI writing detection," Turnitin, November 1, 2023, available at https://www.turnitin.ca/blog/ithenticate-2-0-advancing-research-integrity-with-ai-writing-detection.

23 See, e.g., Jurgita Lapienytė, "ChatGPT creator reveals new tool to spot AI-written homework, admits it's not perfect," cybernews, November 15, 2023, available at https://cybernews.com/tech/chatgpt-anticheating-feature/.