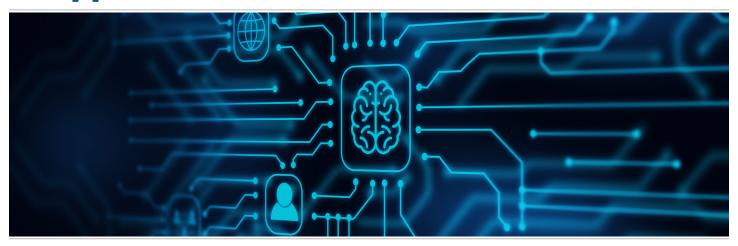# SheppardMullin

Appellate Judge Proposes Possible Use of GenAI for Contract Interpretation – Recognizes That AI Hallucinates But Flesh-And-Blood Lawyers Do Too!

By: James Gatto

By now, most lawyers have heard of judges sanctioning lawyers for misuse of generative AI, typically for not fact checking the outputs. Other judges have issued local rules governing use of or prohibiting AI. These actions have become prevalent. What has not been prevalent is judges encouraging the possible use of AI in interpreting contracts. Perhaps that will change because of a very thoughtful concurring opinion [link to attached decision] in an appeal to the 11th Circuit Court of Appeals in insurance coverage dispute matter. In that opinion, Judge Newsom penned a 31-page concurrence which focused on whether and how AI-powered large language models should be "considered" to inform the interpretive analysis used in determining the "ordinary meaning" of contract terms.

A primary issue in the district court litigation was whether claims against a landscaper were covered by an insurance policy and particularly upon whether installing an in-ground trampoline fell within the meaning of "landscaping." Because the insurance policy didn't define the term "landscaping," the court said, the coverage determination turned on whether the trampoline-related work fit the common, everyday meaning of the word. The district court reviewed multiple (disparate) dictionary definitions provided by the parties and concluded in this case, it did not. On appeal, the plain-meaning battle continued on whether the site work necessary to install the trampoline was "landscaping." The appeals court affirmed the decision but largely based on Alabama law and that the landscaper expressly denied on the insurance application that his work included any recreational or playground equipment construction or erection. Thus, the appeals court was able to reach a conclusion without delving squarely defining what it means to landscape.

Judge Newsom's concurrence notes that despite the result not turning on the definition of landscape, he "spent hours and hours (and hours) laboring over the question" whether the trampoline-installation project qualified as "landscaping" as that term is ordinarily understood. He revealed that, as part of that deliberation, he had the "disconcerting" thought: "Is it absurd to think that ChatGPT might be able to shed some light on what the term "landscaping" means?

From there he explained with remarkable transparency, rarely shared in judicial opinions, his thought process on approaching the decision. He covered a bit of history of the "ordinary meaning rule" and the traditional role of dictionaries in this test. But he noted here that the parties "served up a buffet of definitions" but those definitions "left a little something to be desired." This prompted him to query ChatGPT: "What is the ordinary meaning of 'landscaping'?" ChatGPT responded:

"Landscaping" refers to the process of altering the visible features of an area of land, typically a yard, garden or outdoor space, for aesthetic or practical purposes. This can include activities such as planting trees, shrubs, flowers, or grass, as well as installing paths, fences, water features, and other elements to enhance the appearance and functionality of the outdoor space.

Judge Newsom noted that the explanation seemed more sensible than he had thought it might and definitely less nutty than he had feared. Admittedly, his interest was piqued. He then asked it the ultimate question: "Is installing an in-ground trampoline 'landscaping'"? ChatGPT responded:

Yes, installing an in-ground trampoline can be considered a part of landscaping. Landscaping involves altering the visible features of an outdoor area for aesthetic or practical purposes, and adding an in-ground trampoline would modify the appearance and function of the space. It's a deliberate change to the outdoor environment, often aimed at enhancing the overall landscape and usability of the area.

He posed the query to several other AI chatbots with somewhat similar results.

Despite that answer not being relevant to the ultimate decision on appeal, Judge Newson noted that his process of working through the plain-language issue was a valuable one, if only because it got him thinking about what, to him, was "a previously unimaginable possibility: Might LLMs be useful in the interpretation of legal texts?" He shared that despite initially thinking the idea "positively ludicrous," he was now a pretty firm "maybe" and that it was an issue worth exploring.

In remainder of the concurrence, he did just that.

In my opinion, he wrote a very thoughtful exploration of some benefits and risks of using LLMs. He asserted that LLMs can be one implement among several in the textualist toolkit, to inform ordinary-meaning analyses of legal instruments. I will not try to detail here all of his "exploration" but encourage readers to spend the time to read it for themselves. I believe, this concurrence will stimulate much debate among the judiciary and bar.

One of the more interesting points (to me) and one of the potentially most thought-provoking is his assertion that LLMs hold advantages over other empirical interpretive methods. He noted that some empiricists have begun to critique the traditional dictionary-focused approach to plain-meaning interpretation based on surveys that suggest they don't always capture ordinary understandings of legal texts. He further notes that others have turned to "corpus linguistics," which aims to gauge ordinary meaning by quantifying the patterns of words' usages and occurrences in large bodies of language. He suggests that on balance "reliance on LLMs seems to me preferable to both." He stressed that he is only suggesting that "it's at least worth considering whether and how we might leverage LLMs in the ordinary-meaning enterprise … not as the be all and end all, but rather as one aid to be used alongside dictionaries, the semantic canons, etc."

He also addressed potential drawbacks of using LLMs for this purpose. Of course, one of the obvious drawbacks, is hallucinations, which he recognizes "is among the most serious objections to using LLMs in the search for ordinary meaning." And as noted above, this has been the cause of various lawyers being sanctioned for not fact checking the outputs. But Judge Newsom thoughtfully tackles this thorny issue and suggests that this alone is not a showstopper. He notes LLM technology is "improving at breakneck speed, and there's every reason to believe that hallucinations will become fewer and farther between." He also suggests that hallucinations would be most worrisome when asking a specific question that has a specific answer, but less so when more generally seeking the "ordinary meaning" of some word or phrase. In a forthright observation, he states:

Flesh-and-blood lawyers hallucinate too. Sometimes, their hallucinations are good-faith mistakes. But all too often, I'm afraid, they're quite intentional—in their zeal, attorneys sometimes shade facts, finesse (and even omit altogether) adverse authorities, etc. So at worst, the "hallucination" problem counsels against blind-faith reliance on LLM outputs—in exactly the same way that no conscientious judge would blind-faith rely on a lawyer's representations.

He further addresses some practical questions. One question he poses is that if it's at least worth considering whether LLMs have a role to play in the interpretation of legal instruments, how might we maximize their utility? He answers this, in part, by suggesting that an LLMs "highest and best use is (like a dictionary) helping to discern how normal people use and understand language, not in applying a particular meaning to a particular set of facts to suggest an answer to a particular question." A second question is how can we best query LLMs? He suggests trying different prompts, reporting the prompts used and the range of results obtained and querying multiple models to ensure that the results are consistent.

In conclusion, he notes that these are just preliminary thoughts about whether and how LLMs might aid lawyers and judges, but that plenty of questions remain. Nonetheless, he thinks that LLMs have promise and that he no longer thinks it "ridiculous to think that an LLM… might have something useful to say about the common, everyday meaning of the words and phrases used in legal texts."

For further details, please contact:

James Gatto
Artificial Intelligence Team Co-Leader
bio
202.747.1945
jgatto@sheppardmullin.com

James Gatto is a partner in the D.C. office of Sheppard Mullin and is Co-Leader of its Artificial Intelligence Team. For over 35 years he has been a thought leader on legal issues with emerging technologies and business models, including over 20 years of experience with AI and open source. He provides strategic advice on all IP, tech transactions and tech regulatory issues. He is an adjunct professor at Ole Miss Law School where he teaches Legal Issues with AI. He has been an invited speaker for the US Copyright Office Listening Session on AI Authorship and the USPTO Listening Session on AI Inventorship Issues. He is an appointed member of the AI/Machine Learning Task Force of the American Bar Association's IP Section. He is also a member of the AI committees of the American Intellectual Property Law Association (AIPLA) and the International Technology Law Association.