

LITIGATION

INFORMATION MEMO

OCTOBER 15, 2024

“Garbage In, Garbage Out” – New York Judge Throws Out Expert Damages Calculation Generated by Artificial Intelligence

What happens when an expert relies on Artificial Intelligence (“AI”) in preparing a report? Saratoga County Surrogate Judge Jonathan G. Schopf says that the use of AI as a tool to assist in preparing an expert damages calculation should be the subject of a *Frye* hearing to determine its admissibility prior to trial.

In the Matter of Weber (Saratoga County Surrogate 1845-4/B) involved a beneficiary’s objections to a trust accounting, including a breach of fiduciary duty claim against the trustee relating to the retention of a parcel of real property owned by the trust. During a three-day bench trial, the objectant’s fiduciary expert offered a supplemental damages report in an attempt to prove that the trust would have gained significant profits had the trustee sold the property earlier in her trusteeship and invested those funds in a Vanguard Balanced Index Fund. The expert testified that in preparing a supplemental damages report, he used Microsoft CoPilot, a large language model similar to ChatGPT, “in cross-checking his calculations.” However, the Court observed that on cross-examination, “Mr. Ranson could not recall what input or prompt he used to assist him with the Supplemental Damages Report. He also could not state what sources Copilot relied upon and could not explain any details about how Copilot works or how it arrives at a given output. There was no testimony on whether these Copilot calculations considered any fund fees or tax implications.”

Observing that “[t]he Court has no objective understanding as to how Copilot works, and none was elicited as part of the testimony,” the Court took it upon itself to test out CoPilot’s accuracy. The Court asked CoPilot: “Can you calculate the value of \$250,000 invested in the Vanguard Balanced Index Fund from December 31, 2004 through January 31, 2021” on three different Unified Court System computers. CoPilot returned three different values: \$949,070.97, \$948,209.63, and a little more than \$951,000.00. Each of these values was different than the expert’s calculation.

Seemingly asking CoPilot to evaluate itself under the *Frye* standard, the Court then asked CoPilot “are you accurate” and “are you reliable.” Not satisfied with CoPilot’s responses, the Court critiqued: “This brings to mind the old adage, ‘garbage in, garbage out.’ Clearly a user of Copilot and other artificial intelligence software must be trained or have knowledge of the appropriate inputs to ensure the most accurate results.”

Analyzing how the expert’s use of CoPilot in his damages report impacted admissibility of the report under the rules of evidence, the Court relied on the familiar *Frye* standard (articulated in *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923)) and found that in “the instant case, the record is devoid of any evidence as to the reliability of Microsoft Copilot in general, let alone as it relates to how it was applied here. Without more, the Court cannot blindly accept as accurate, calculations which are performed by artificial intelligence.”

The Court recognized that artificial intelligence has utility and its use “is a rapidly growing reality across many industries,” but stated “the mere fact that artificial intelligence has played a role, which continues to expand in our everyday lives, does not make the results generated by artificial intelligence admissible in Court.” The Court held:

due to the nature of the rapid evolution of artificial intelligence and its inherent reliability issues that prior to evidence being introduced which has been generated by an artificial intelligence product or system, counsel has an affirmative duty to disclose the use of artificial intelligence and the evidence sought to be admitted should properly be subject to a Frye hearing prior to its admission, the scope of which should be determined by the Court, either in a pre-trial hearing or at the time the evidence is offered.

A curious postscript:

In connection with preparing this article, the author input the Surrogate’s Decision into CoPilot and asked: “Can you please summarize the Court’s findings with respect to the expert’s use of CoPilot in his Supplemental Damages Report?”

CoPilot enthusiastically responded (albeit after admitting that it had trouble reading the Decision): “The Court found that the use of CoPilot was appropriate and did not compromise the integrity of the report. The expert’s methodology was deemed reliable, and the findings were accepted as valid evidence in the case” and concluded: “[o]verall, the Court’s findings supported the expert’s use of CoPilot, highlighting its role in improving efficiency and accuracy in the analysis of damages.”

Perhaps CoPilot took offense to Judge Schopf’s “garbage in, garbage out” comment and decided to tell its side of the story.

If you have any questions about the information presented in this memo, please contact [Kathleen H. McGraw](#), [Mario F. Ayoub](#), [Mara D. Afzali](#) or any attorney in Bond’s [litigation practice](#), [artificial intelligence practice](#) or the attorney at Bond with whom you are regularly in contact.

