

UNITED STATE DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

ROBERTO MATA,

Plaintiff,

Civil Action No 2-cv-1461 (PKC)

-against-

AVIANCA, INC.,

Defendant,

**[PROPOSED] AMICUS CURIAE BRIEF OF GREGORY C. BELMONT
& CEREBEL LEGAL INTELLIGENCE
IN SUPPORT OF COURT'S CONSIDERATION OF ORDER TO SHOW CAUSE**

Gregory C. Belmont
CereBel Legal Intelligence
Pro Se
750 Lexington Avenue, 9th Floor
New York, NY 10012
(631) 506-8424
Fax (914) 462-4465
belmont@grandcentralholdings.com
Amicus Curiae

TABLE OF CONTENTS

PRELIMINARY STATEMENT 1

INTEREST OF THE AMICUS CURIAE 3

A. The Amicus’s livelihood is closely associated with the software attributed with error. 3

B. The Amicus possesses the requisite expertise to provide timely and useful information to the Court 4

TECHNOLOGY BACKGROUND 5

A. Enabling Deep Neural Network Technology and Large Language Models 5

B. Training 6

C. The Launch and “Iterative Development” of ChatGPT 7

D. Capabilities 9

E. Shortcomings 10

F. Legal Utility 11

G. Safety Messages 12

H. Illustrative Legal Prompts 12

ARGUMENT 12

I. COLORABLE REASONABLENESS OF ATTORNEYS FAITH IN GENERATED LEGAL RESEARCH 13

A. Belief Citations and Judicial Opinions “generated by” Chat GPT were bona fide may not have been unreasonable at the time of consultation. 13

B. Discrepancies in submitted chat transcripts 14

C. Use, evaluation, and perfection of chat transcripts as evidence. 14

II. SANCTIONS MUST FOCUS ON CONDUCT, NOT MISUSED TOOLS 15

A. The Court must consider effect of sanctions on the bar and client benefits. 15

B. Responsive standing orders of two other district courts demonstrate that directives could form a harmful precedent 16.

C. Neither sanctions nor courtroom rules should preempt legislative action or make new in haste. 19

D. Discipline the conduct, not the tool. 19

CONCLUSION 21

PRELIMINARY STATEMENT

Gregory C. Belmont (“Belmont”) on behalf of himself and his fully owned business CereBel Legal Intelligence (“CereBel”) respectfully submits this Amicus Curiae Brief to support the Court’s consideration of its sua sponte Order to Show Cause to issue sanctions to attorneys Peter LoDuca, Esq. and Steven A. Schwartz, Esq. and associated law firm Levidow, Levidow & Oberman, P.C. (“Subject Attorneys”) pursuant to Rule 11(b)(2) & (c) and only incidentally to support consideration of sanctions pursuant to Fed. R. Civ. P., (2) 28 U.S.C. § 1927 for variously directly and indirectly citing non-existent cases and submitting non-existent judicial opinions attributed to use of the generative artificial intelligence software application ChatGPT to conduct legal research (D.E. #32, Att. 1).

The software application alleged to have generated the submission error was developed by California non-profit OpenAI, Inc. and affiliated entities (“OpenAI”) and typifies a brand new generation of artificial intelligence technologies that have been developed and/or made popularly available by internationally significant companies Microsoft, Google, Meta, NVIDIA and Amazon and rapidly adopted since first broad public release on November 30, 2022¹. These so-called generative AI applications are widely expected to have a massive impact on the national and international economy akin to a Fourth Industrial Revolution² and are the subject of intense public interest. More prosaically, generative AI (“GenAI”) is expected to bring massive productivity gains to and thereby upend labor structures in many knowledge industries including to the practice of law³. Indeed, the instant Order to Show Cause (“OSC”) has generated

¹ [“ChatGPT is growing faster than TikTok,”](#) February 1, 2023

² [“Sui Generis Right for Trained AI Models,”](#) Intellectual Property Owners Association, November 2, 2020,

³ [“End of the Billable Hour?”](#) *Wall Street Journal*, May 11th, 2023 and [“A.I. Is Coming for Lawyers, Again,”](#) *New York Times*, April 10, 2023

enormous media coverage and public interest (“repeated in hundreds (if not thousands) of articles and online posts,” D.E. #45, p. 1; D.E. #44, Att. 1, ¶10; D.E. #48, Att. 10-12).

The OSC may be the first action in a United States court, federal or state, in which use of Generative AI software is the subject of controversy and has motivated at least two standing orders in other districts related to the use of the technology (“[Mandatory Certification Regarding Generative Artificial Intelligence](#)”, Hon. B. Starr (N.D. Tex.) and “[Standing Order for Civil Cases Before Magistrate Judge Fuentes](#),” N.D. Ill., May 31, 2023.) To the extent the Subject Attorneys use of GenAI software is found by the Court to be typical and innocent on the one hand or careless and irresponsible on the other and the related degree to which issued Rule 11 sanctions are *too severe* or *too lenient* could potentially *chill productive and responsible use* or *further stimulate development and adoption* of these technologies by legal practitioners. This is especially so since sanctions resulting from 11(b)(2) violations initiated by the Court are to be “limited to what suffices to deter repetition of the conduct or **comparable conduct by others similarly situated**” (11(c)(4)(emphasis added.) Miscalibrated sanctions could negatively impact the Legal AI software industry, depriving practitioners and their clients of the many productivity gains possible.

Given the new and novel nature of these software applications and common misunderstandings about what they are and are not capable of doing, fair consideration of sanctions in this matter requires understanding operation of the software used and/or relied upon by the Subject Attorneys and the complete context in which it was made available for public use in late November, 2022. This Brief presents timely and useful information to support these understandings by providing information about:

- The market context, including messaging by the media and the software’s developer.

- The user experience produced by the software.
- Assessing excerpted records of content generated by the software.
- The full capabilities and weaknesses of the technology when used responsibly.

INTERESTS OF THE AMICUS CURIAE

A. The Amicus’s livelihood is closely associated with the software attributed with error.

The Amicus is the individual owner, investor, data scientist, software architect, and operator of CereBel Legal Intelligence, a New York County based sole proprietorship that applies Generative AI to the needs of the legal practitioners. CereBel utilizes the same OpenAI large language models (GPT-3.5 and GPT-4, collectively the “GPTs”) and platform that powers ChatGPT (the “GPT Platform”) through use of an Application Programming Interface (“API), along with comparable Generative artificial intelligence tools provided by Google and start-ups Anthropic and Hugging Face (also New York City based, albeit located in the EDNY), to deliver proprietary software services to attorneys that generate full text, interactive, and semantically searchable summaries of statutes, rules of procedure, judicial decisions, patents, and litigation documents including deposition transcripts, medical histories, bills of particulars, and mediation briefs. In addition to pioneering methodologies to effectively summarize and extract relevant data from legal documents, CereBel provides an interactive interface that enables attorneys to utilize secure versions of the GPT Platform in a manner highly comparable to use of the eponymous Chat GPT application but with the added benefits of additional privacy, citation guardrails™, and ability to augment queries with litigation case files in a process referred to as “retrieval augmented prompting.”

The degree to which sanctions are or are not issued in this matter will impact the public’s, the legal practice industry’s, and the wider judiciary’s perception and evaluation of ChatGPT and

related software including the Amicus’ which are broadly perceived as one-and-the-same as the application consulted by the Subject Attorneys⁴. Therefore Amicus has a compelling interest in ensuring the Court be presented with evidence setting forth the potential harm created by sanctions, and non-statutory orders that could likely be thereby motivated, that focus on tools used to make professional error rather than on conduct. Directives that chill the productive use and experimentation with GenAI technologies by the bar would undercut the need for continued investment in the many safety layers engineered by CereBel and other “legal AI” software ventures. Therefore, the outcome of the Order to Show Cause could have a significant potential effect on the market for CereBel’s legal AI software, the Amicus’s related professional services, and further investments made by him and the industry.

B. The Amicus possess the requisite expertise to provide timely and useful information to the Court

The Amicus is informed by significant time spent since ChatGPT’s November, 2023 launch evaluating, generating legal analysis with, engineering software products integrating, and instructing attorneys in the Southern District of New York on use of the GPT Platform. Rooting in this intimate personal knowledge of a) how LLM’s are constructed and trained, b) how ChatGPT specifically performs with respect to legal searches like those declared to have been conducted by and on behalf of the Subject Attorneys, and c) how the general consumer and legal markets have been exposed to the software’s capabilities and limitations enables the Amicus to be of considerable help to the Court in evaluating the reasonableness of the Subject Attorneys

⁴ “‘ChatGPT’” has come to be used as shorthand for all generative AI, almost like it’s the Band-Aid or Kleenex of AI. It’s critical to understand, however, that that is inaccurate, “[The Problem With the ‘Bogus’ ChatGPT Legal Brief? It’s Not the Tech.](#)” June 2, 2023, Law.com

conduct and how sanctions, if deemed necessary, might be crafted to achieve whatever balance of discipline and persuasive impact the Court desires.

TECHNOLOGY BACKGROUND

A. Enabling Deep Neural Network Technology and Large Language Models

The Large Language Model technology at the core Generative AI evolved from:

- i) a simplified mathematical model of a biological neuron described by early neuroscientists Warren McCulloch and Walter Pitts in 1943⁵, which
- ii) along with the learning theory of neuropsychologist Donald Hebb, inspired Frank Rosenblatt, with funding from the Office of Naval Research at Cornell, to build a binary classifier known as a perceptron⁶, which lead
- iii) to the development of multilayer perceptrons, or artificial neural networks, that could model more complex relationships (eg, non-linearly separable data) through the introduction of hidden layers and popularization of an error-correcting algorithm called backpropagation⁷ for adjusting the weights of the connections between artificial neurons, enabling
- iv) neural networks to be were reinvigorated in the 1990's and 2000s by, among others, Geoffrey Hinton (who also co-authored the backpropagation paper), Yan Lecun, and Yoshua Bengio,⁸ variously at times at Bell Labs, University of Toronto, and University of Montreal, and more recently also Google and Meta, initially through the success of a commercial system to recognize handwritten numbers on bank checks and then with more specialized network

⁵ “[Logical Calculus of the Ideas Immanent in Nervous Activity](#),” *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133, 1943

⁶ “[The perceptron: A probabilistic model for information storage and organization in the brain](#),” *Psychological Review*, 65(6), 386–408, 1958

⁷ “[Learning representations by back-propagating errors](#),” October 9, 1986

⁸ Awarded 2018 Turing Award (often referred to as the "Nobel Prize of Computing")

architectures, such as Convolutional Neural Networks (CNNs) for image recognition tasks and Recurrent Neural Networks (RNNs) for sequenced data which lead to revolutionary breakthroughs in image recognition⁹, speech recognition, and language translation, and then

v) to the first “transformer model¹⁰” utilizing a so-called attention mechanism to process input data in parallel, rather than sequentially, which, by enabling words to be weighed against one another in context, enables large and multi-layered networks of rudimentary artificial neurons to guess the next word in a sequence, from which ability

vi) when combined with increased computational power, cloud computing, big data, generously open sourced deep learning software libraries, specialized hardware like graphics processing units, and investments in basic research by industry and a small number of venture capital investors (largely Google, Meta, and a group of well-known Silicon Valley based individuals), emerged the power of Large Language Model transformers like ChatGPT to generate incredibly human-like text and perform a range of useful natural language processing tasks, including those at controversy in this matter.

B. Training

Text Corpus. After particular network architecture and training algorithms are selected, transformer LLMs are developed through a two-step process. In a pre-training stage, the models “learn” from exposure to a diverse corpus of digitized text (“Corpus”),¹¹ which include a broad variety of topics meant to represent useful human knowledge including law-related texts¹², to

⁹ “[ImageNet Classification with Deep Convolutional Neural Networks](#)”, Communications of the ACM 60 (6), 84-90, 2017

¹⁰ “[Attention is all you need](#),” NIPS 2017

¹¹ “[Inside the secret list of websites that make AI like ChatGPT sound smart](#),” Washington Post, April 19, 2023

¹² While non-profit OpenAI does not disclose the source of GPT-4’s training text, a comparable corpus used by Google includes content from law & government websites: patents.google.com, patents.com, caselaw.findlaw.com, publications.parliament.uk, freepatentsonline.com, openjurist.org, and uscourts.gov.

predict the next word in a sentence. Once a particular Corpus is assembled and the model pre-trained, the model's principal weights are essentially locked and no new, more current text is added to the corpus, freezing the "knowledge" to which the model was exposed and therefore the timeliness of any data it might be able to replicate from its collective memory. The model then undergoes supervised fine-tuning (SFT) and reinforcement learning from human feedback (RLHF), laborious processes in which human reviewers following specific guidelines review and rate possible outputs, helping the model align with utility, ethical, and safety considerations (a process referred to as "alignment" or "safety alignment.")

Large Language Models can also be made conversational and positioned as "chatbots", seeming to "understand" context and maintaining coherence in successive responses, despite not fully understanding the conversation in a human sense through the various tuning processes and methods that store data about individual "chat sessions". To date chatbots, which can also be deployed as customer service, technical support, and virtual assistance agents, have become the most popular way users access the capabilities of LLMs.

C. The Launch and "Iterative Development" of ChatGPT

ChatGPT, the most well-known chatbot, was publicly released via web browser access on November 30, 2022, gaining one million users within five days and 100 million in two months, including among members of the bar¹³, only three months before filing of the first papers reflecting its troublesome utilization in this matter (D.E. #21, filed March 1, 2023.) Prior to that time OpenAI's technology was not marketed to consumers. A relevant timeline:

¹³ ["End of the Billable Hour? Law Firms Get On Board With Artificial Intelligence. Lawyers start to use GPT-4 technology to do legal research, draft documents and analyze contracts,"](#) *Wall Street Journal*, May 11th, 2023 and ["A.I. Is Coming for Lawyers. Again. Previous advances in A.I. inspired predictions that the law was the lucrative profession most likely to suffer job losses. It didn't happen. Is this time different?,"](#) *New York Times*, April 10, 2023

When first launched, via a web browser interface, Chat GPT provided access to a version of OpenAI's GPT-3.5 LLM called text-davinci-003 which was tuned for conversational instruction following (the first version of a GPT-3 LLM was released in 2020 and GPT-2 in 2019, but these were not marketed to consumers.) Since then, the following upgrades have been made¹⁴:

- **2017:** Google publishes foundational "Attention is all you need" paper
- **2019:** GPT-2 made available for research.
- **2020:** GPT-3 made available for research and to select developers.
- **January, 2022:** InstructGPT, the first models to be trained with human SFT to be better at following user intentions is released for research and to select developers.
- **November 30, 2022:** ChatGPT launched via a website interface and waisted listed API to provide access, respectively, to the GPT-3.5 LLM later named GPT-3.5 turbo and text-davinci-003, the former of which was trained with RLHF and the latter of which was not. The application becomes the most rapidly adopted in history and the pace at which established and start-up software companies like CereBel are granted the ability to incorporate the GPT Platforming powering ChatGPT into existing and new applications through use of an Application Programming Interface ("API") increases.
- **January 9, 2023:**¹⁵ model improvement that "should be generally better across a wide range of topics and has improved factuality."
- **January 30, 2023:** Upgraded model with "improved factuality and mathematical capabilities."
- **February 1, 2023:** ChatGPT Plus \$20/month subscription plan with faster processing and early access to new features.
- **March 1, 2023:** GPT-3.5 Turbo model made available to API developers.
- **March 14, 2023:** GPT-4, reported to have been trained with RLHF and other techniques for six months and delivering a significant upgrade in capabilities, is made available to ChatGPT Plus subscribers and API developers. Among the capabilities of , GPT-4 is touted as capable of passing state bar exams in the 90th percentile¹⁶.
- **March 23, 2023:** Experimental web browsing capability added along with Plugins¹⁷ by waitlist for subscribers enabling queries to access live internet and thus for augmented information prompts to include text created after the Corpus lock.
- **May 3, 2023:** Legacy (GPT-3.5) model, with which the subject errant citations may have been generated, is deprecated.
- **May 18, 2023:** A mobile interface is made available via Apple's App Store.

¹⁴ "[GPT-4, GPT-3, and GPT-3.5 Turbo: A Review Of OpenAI's Large Language Models](#)," *Ankur's Newsletter*, April 10, 2023

¹⁵ "[ChatGPT — Release Notes](#)" OpenAI accessed June 6, 2023

¹⁶ "[GPT-4 Passes the Bar Exam](#)," *SSRN*, March 15, 2023

¹⁷ "[ChatGPT Plugins](#)," OpenAI accessed June 6, 2023

OpenAI CEO Sam Altman described the process by which the company “decide[s] whether or not a model is safe enough to deploy and safe enough to have been built” (question asked by Senator Christopher Coons, Subcommittee Hearing) as “iterative development.”

“I’d like to frame it by talking about why we deploy at all. Like why we put these systems out into the world. There’s the obvious answer about there’s benefits and people are using it for all sorts of wonderful things and getting great value, and that makes us happy. But a big part of why we do it is that we believe that iterative deployment and giving people in our institutions and you all time to come to grips with this technology to understand it, to find its limitations, it benefits that the regulations we need around it, what it takes to make it safe. That’s really important. Going off to build a super powerful AI system in secret and then dropping it on the world all at once, I think would not go well. So a big part of our strategy is while these systems are still relatively weak and deeply imperfect, to find ways to get people to have experience with them, to have contact with reality and to figure out what we need to do to make it safer and better.” (Sam Altman, Ibid.)

The rapid rise of ChatGPT took OpenAI and most other industry participants and observers by surprise. Other companies rushed their LLM-driven products to market including Microsoft’s Bing (incorporating GPT-4), Google’s Bard, and Google subsidiary Quora’s Poe (incorporating OpenAI and Anthropic’s models.)

D. Capabilities

The natural language capabilities of LLMs like ChatGPT can be categorized as: Composition (fiction), Composition (non-fiction), Coding, Conversation, Search, Summarization, Editing, Translation, Classification, and Sentiment Analysis. While all of these are accomplished by synthesizing strings of words, the later five typically constrain output to provided text and therefore may not be considered under all definitions “strictly generative” tasks. Accordingly, the Amicus differentiates the functionality of LLM-facilitated tasks as :

“De Novo” (Pure Generation)

Generation of content or interaction with users (eg, composition, coding, conversation, coding, searching), where LLMs provide the benefit of creative output, conversational engagement, or unaided information prompting (meaning not augmented by in-context text.)

“Ab Initio” (In-context Processing)

Processing and analyzing furnished text data (eg, summarization, editing, translation, classification, sentiment analysis), where LLMs provide the benefit of extracting meaningful information, improving text quality, or categorizing content.

Both forms of Generative AI offer benefits to the legal community including, among many others: Time saving efficiencies with demonstrated ability to streamline writing processes and, in the right circumstances and especially with the addition of value added layers provided by companies like CereBel, quickly analyze some complex legal issues such as identifying potential discrepancies in testimony. Enhanced Accuracy and Consistency: Legal writing can require a high degree of consistency; under the right use cases, LLMs can enhance the overall quality and reliability of legal documents. Augment, not replace: By reducing time spent on rote tasks that do not require the nuanced and complex critical thinking and judgment of trained attorneys, can allow more focus on higher-level analysis and client advocacy. Creativity: The tool also proves helpful facilitating exploration of creative argument and persuasive writing.

Information and Retrieval Augmented Prompts, in which CereBel specializes, are techniques in which additional information is retrieved from a knowledge base or external sources to augment the prompt, or pseudo-query, provided the language model. This additional information provides context, facts, or references that provide the basis for more accurate and relevant responses. Using this technique enables LLMs to be engineered to provide natural language query interface to traditional document stores. For example, enabling queries of case files containing discovery documents.

E. Shortcomings

The text generated by LLMs, since they are not driven by traditional faithful databases or prescribed rules, can range from eerily precise and accurate to extremely random, performance compared to a “stochastic parrot” or “blurry JPEG.”¹⁸ Frequent problems include: hallucinations,

¹⁸ “[On the dangers of stochastic parrots: Can language models be too big?](#),” *In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*. 610–623, March 3, 2021 and [ChatGPT is a Blurry JPEG of the Web](#),

confident errors and false confidence (seemingly plausible statements with no contrary indicators); variance (different outputs from the same prompt); bias (reflecting any in the Corpus such as, to some critics, prevalent cultural hegemonies); errant reasoning; and, very rarely, jumble. While not shortcomings in their utility, open legal questions not directly relevant to the issue in controversy here, include potentially unresolved open issues related to possible infringement of copyrights of Corpus content owners, potential for libelous output (one critic referred to LLMs as “Large Libel Models”), attorney-client privilege¹⁹, and other confidentiality concerns.

F. Legal content

The more the pre-training Corpus includes judicial opinions, statutes, regulatory codes, contracts, and other legal documents, and the less these types of texts are discouraged during the alignment process, the more likely the output will helpfully mirror authentic citations or harmfully comprise de nouveau text that only resembles the format of the legal content while depicting fictional judicial citations or opinions, artifacts referred to as hallucination, or perhaps more precisely, ‘hallucinations’.²⁰

Several practical and theoretical methods are being engineered and explored to increase the quantity and salience of domain-specific output like accurate and faithful citations. Changing the mix of domain content in pre-training and tuning is one direction being explored. Another involves pre-processing text to substitute specialized tokens for reference entities like case

The New Yorker, February 9, 2023 which also characterizes GenAI output as *paraphrases* vs internet search engine’s as *quotes*.

¹⁹ “[ChatGPT and Ethics: Can Generative AI Break Privilege and Waive Confidentiality?](#),” January 26, 2023

²⁰ A neologism some have attributed to AI scholar [Kate Crawford](#), but which predates her use by at least one other [tweeter](#).

citations which tokens can be generated by the models²¹. The Amicus's business post processes LLM output to authenticate statute references against internet databases and converts authenticate references to clickable links; the business is in the process of doing the same with judicial citations.

G. Legal safety messages

Part of the safety alignment measures engineered in ChatGPT pertain to the inclusion of disclaimers such as the following illustrative examples which were excerpted from actual chat sessions. Such disclaimers appear more regularly at the current time than they likely did before the subject submission was made on March 1, 2023, presumably as a result of the described upgrades to ChatGPT:

The cases referenced in my previous response, such as "People v. Samuels," "Matter of Johnson," "In re Ramirez," and "Doe v. Smith," were fictional examples created for illustrative purposes. They were not intended to represent actual published decisions in New York State or any other jurisdiction. If you require legal precedent specific to New York State, I recommend consulting reputable legal research platforms, law libraries, or qualified legal professionals who can provide accurate and up-to-date information on relevant cases and their citations.

I'm sorry for any confusion, but as an AI developed by OpenAI, I don't have access to specific legal databases or the ability to retrieve specific court cases. My training data includes a wide range of data sources, but does not extend to specific case law databases.

H. Illustrative Legal Prompts

With appropriate experience, know-how, and value added software layers, users can learn to reduce the number of hallucinations and increase the quality of research content produced by GenAI. A user generated database of legal prompts and conversational chat session transcripts

²¹ [Tweet](#), April 6, 2023, Google AI Research Scientist Margaret Mitchell; [quote tweeting](#) relevant perspective of Princeton computer scientist Arvind Narayanan

that illustrate the type of legal research that can be effectively performed, along with instructive bloopers, is being made available at <https://www.cerebel.law/promptdb>.

ARGUMENTS

I. COLORABLE REASONABLENESS OF ATTORNEY’S BELIEF IN LEGAL RESEARCH

A. Belief Citations and Judicial Opinions “generated by” Chat GPT were bona fide may not have been unreasonable at the time of consultation.

The attorney conduct at issue in this matter— submission of bogus case law citations and textual references— has arisen from personal use of the ChatGPT application to conduct legal research on case law relevant to the impact of bankruptcy and international aviation treaties on the statute of limitations in a personal injury matter (D.E. #32, Att 1 and D.E. #42). Belief citations and judicial opinions “generated by” ChatGPT were bona fide may not have been unreasonable at the time of Attorney Schwartz’s consultation. ChatGPT gained extraordinarily wide user base, fanned by wide hype and praise in general and technology press (D.E. #45, page 16)²², including among members of the bar, for months following its public release in November, 2022, only three months before filing of the first papers reflecting its troublesome utilization in this matter (D.E. #21 filed March 1, 2023.) Attorney Schwartz used the application before many of the incremental safety and accuracy enhancements described in the aforementioned timeline were implemented. At the time of his use, only the GPT-3.5 language model powered the ChatGPT application, not the vastly improved GPT-4 language model. The 3.5 model implements fewer safety guardrails such as disclaimers about Corpus cut-off date of 2021, warnings to consult other sources, and was more prone to provide fictitious judicial citation.

²² “End of Billable Hour,” *New York Times* and “A.I. is Coming for Lawyers,” *Wall Street Journal*

B. Discrepancies in submitted chat transcripts

The discrepancy between the chat transcript of Schwartz's research for his opposition to the motion to dismiss (D.E. #32, Att 1 presented as partial screenshots and D.E. #46 Ex. 1 described in D.E. #46, ¶18) and the chat transcripts depicting similar prompts submitted by his attorney (D.E. #49, Att. 9 and described in D.E. #48 ¶5) have almost certainly arisen from the latter's production by the GPT-4 language model²³. The Court may wish to confirm the same in query to counsel.

C. Use, evaluation, and perfection of chat transcripts as evidence.

The transcripts described above may be the first time in a US Court, federal or state, that work product consisting of text produced by generative artificial intelligence software, is being evaluated for evidentiary value. The more complete impression provided by the longer printed transcripts compared to the partial screenshots demonstrates the higher weight the Court should accord them. Evaluating partial screenshots of partial transcripts is analogous to providing a partial bank ledger or partial deposition testimony.

Perfected evidence of a Generative Artificial Intelligence chatbot would consist of an entire chat transcript, the name of the language model (eg, GPT-3.5, GPT-4), the name of the software application (ChatGPT, Bard, Poe, Bing, CoCounsel, Spellbook, CereBel) the date of production, and, in some instances, other parameters that could vary or otherwise been known (eg, temperature; completion length; presence, frequency or other penalties; top-p; etc.) Even then, not all chat transcripts can be recreated verbatim even when the chatbot configuration is

²³ The order of prompts can also have a significant impact on generated text, but here the difference between the 3.5 and 4 models is almost certainly responsible for the vast discrepancies.

replicated for a variety of reasons including small discrepancies in the operation of floating point numbers²⁴.

II. SANCTIONS MUST FOCUS ON CONDUCT, NOT ENABLING TOOLS

If the Court determines that Attorney Schwartz had a good faith belief ChatGPT performed the function of a traditional legal research service or could not envision the possible generation of a non-existent court opinion, the conduct here is akin to an error that could be produced by any “headnote lawyer” which Black’s Law Dictionary disparages as “A lawyer who relies on the headnotes of judicial opinions rather than taking the time to read the opinions themselves” and UCLA Distinguish Law Professor Eugene Volokh modernized as a “Chatbot Lawyer.”²⁵ Although Northern District of Illinois Magistrate Judge Gabriel A. Fuentes describes a presumption that a Rule 11 certification represents a presumption that filers have “**read and analyzed all cited authorities** to ensure that such authorities actually exist” (emphasis added), the Court might consider the degree to which it is common and accepted practice for the actual reading of all citations to depend on factors such as relevance and volume of cited authorities, certifier’s familiarity with the cited authorities, and dependence on legal research tools, colleagues, law clerks, and/or paralegals.

A. The Court must consider effect of sanctions on the bar.

The Second Circuit has found that Rule 11 sanctions can have the effect of dissuading counsel from engaging in ethical conduct that is deemed in the best interest of clients if sanctions issued to third party attorneys create a threat that attorney conduct could be construed as improper by associating that conduct with penalties. In RE: Pennie & Edmonds LLP, No.

²⁴ “[Observing discrepancy in completions with temperature = 0](#),” OpenAI Community Forum, February 24, 2023

²⁵ “[I Propose a New Term: “Chatbot Lawyer”](#),” *The Volokh Conspiracy*, March 3, 2023

02-7177, 2003 WL 1191197 (2d Cir. Mar. 14, 2003), counsel defending an alleged trademark infringer were sanctioned for their client’s submission of a fraudulent document prior to their appearance and even though their client was untruthful to them. Although the District Court found the law firm did not act in bad faith by “permitting [their client] to file an affidavit containing statements that the law firm could not have objectively believed were true,” the court issued a non-monetary sua sponte sanction to the firm. The Circuit Court found that the objective unreasonableness standard was too severe for a sua sponte sanction and that it was necessary to find the least stringent possible standard to apply under the rules because, the majority reasoned, Rule 11 sanction regimes can have the effect of deterring legitimate actions by counsel in the interest of clients “out of apprehension that their conduct will erroneously be deemed improper.” and too lenient regimes can “embolden them to [take] improper [actions] on behalf of clients” (in that case, to “make improper submission”). Here a harsh regime could have the effect of dissuading attorneys from adopting technologies similar to those used by the Subject Attorney that promise to increase the quality and efficiency of counsel provided to clients. Such apprehension would harm the Amicus’s livelihood.

B. Responsive standing orders of two other district courts demonstrate that directives could form a harmful precedent by motivating nonconstructive judicial rules.

The first business day after this Court’s OSC was issued and reported nationally and presumably as a result of the notoriety, Judge Brantley Starr (N.D. Tex) issued his standing order mandating appearances in his court be accompanied by negative certifications regarding use of generative artificial intelligence, which the order both praised as powerful and useful and maligned as lacking honor, conviction, and principle. Already, this judge-specific requirement,

which permits use of GenAI while effectively imposing a financial tax on doing so (use of a traditional legal database or reporter, most of which are expensive, prohibitively so for many pro se parties, and the largest of which are commercialized by entrenched providers subject to credible antitrust claims), has spawned national and industry headlines characterizing the rule as a complete prohibition on use of this powerful technology. Misinterpreted or not, it is evident from its face that the order does not define what is meant by a) generative artificial intelligence or, more problematically, what is meant by b) being “drafted by” generative artificial intelligence given the variety of text-driven tasks GenAI can perform to support attorneys participating in litigation.

A day later, Magistrate Judge Gabriel A. Fuentes (N.D. Ill.), quoting from this docket and Judge Brantley’s order, issued a standing order which improves on the later’s malalignment of generative AI platforms’ drafting ability by recognizing “AI as a tool for legal research and drafting.” While the updated certification rules of both courtrooms do explicitly allow GenAI, in one case in a qualified manner, the explicitly, but imprecisely stated judicial requirements to vet the tools may pose compliance headaches that dissuade use and inhibit vital experimentation and education. Attorneys should be trusted to make choices concerning how they practice and represent their clients. And legal tech vendors should not be unnecessarily impeded in efforts to win their trust.

1. Further concerns about opt-in certification.

Other problematic aspects of the N.D. Tex order include: a) discriminatory to include selective brand names (eg, “ChatGPT, Harvey.AI, or Google Bard”) and not general description of product or service provider categories, b) prejudicial to one set of tool providers (traditional

legal databases > generative AI models), c) redundant recapitulation of Rule 11, other Federal Rules of Civil Procedure (FRCP), and rules of professional conduct while not being least restrictive way of achieving compelling court interest (a warning notice clarifying interpretation of Rule 11 obligations might be), d) only four mentioned use cases (“form divorces, discovery requests, suggested errors in documents, anticipated questions at oral argument”) amount to endorsement by court of a restrictive list favoring named task types over others for which platforms can demonstrate at least equal utility, e) hallucination criticism disregards many productive and responsible uses which do not tend to hallucinate such as Ab Initio tasks comprising information augmented prompts (eg, summarization tasks where generation is restricted to provided texts), f) criticizes GenAI for bias, yet traditional databases have also been found to have bias²⁶, g) asserts GenAI doesn’t swear an oath where neither do headnote writers, legal reporter editors, or executives of multinational legal research database companies, h) criticizes GenAI for lack of allegiance to any law, yet providers are fully subject to all US laws, including product liability and civil suit, i) distinguishes computer code from conviction, yet coded algorithms like backpropagation are the ultimate conviction machines (following instructions without sentiment) and lawtech innovators are demonstrating how automation can benefit the administration of law, lead to more efficient and equitable resolution of legal issues with fewer court proceedings to free up resources to focus on matter where human judgment is most indispensable²⁷, and j) burdensome opt-ins to use lawful advocacy tools improperly asserts

²⁶ “[The Algorithm as a Human Artifact: Implications for Legal \[Re\]Search](#),” *Colorado Law School Faculty Commons*, 109 Law Libr. J. 387, 2017 and “[Understanding the Technical Bias of Westlaw, Lexis Advance, Fastcase, Google Scholar, and Casetext](#),” *Geeks and a Law Blog*, December 12, 2016

²⁷ “[Law As Code: A Legal System Shaped by Software](#),” Future.com, June 15, 2023 and “[The expansion of algorithmic governance: from code is law to law is code](#),” *Field Actions Science Reports*, December 31, 2017 describing forms of regulation wherein technology is used to enforce rules.

courts in the role of advocate by influencing how attorneys choose to plead for their clients, potentially depriving represented parties of the utility, productivity gains and cost savings associated with these “powerful” technologies.

C. Neither sanctions or courtroom rules should preempt legislative action or make new law in haste.

The United States Senate has held at least one subcommittee hearing to express and hear concerns about artificial intelligence²⁸ and two days ago, June 5, 2023, US House of Representative Member Ritchie Torres of the 15th Congressional District located within Bronx County, NY introduced the "AI Disclosure Act of 2023"²⁹ proposing to require labeling of AI generated content with a proscribed disclaimer (“Generative artificial intelligence shall include on any output generated by such artificial intelligence the following: ‘Disclaimer: this output has been generated by artificial intelligence.’”)

This Court and other federal courts might wish to provide the Article I deliberative law-making process time to fully contemplate these issues before establishing extra-statutory dictates that may not fully consider all parties with relevant interests at stake including the bar, the legal tech industry, access to justice groups, legal consumers, etc.

D. Discipline conduct, not tools.

“As has been the case since the dawn of the profession, lawyers are responsible for the work they provide to clients and courts. No technology, no matter how impressive or magic-seeming, absolves lawyers of their professional or ethical duties,” writes Editor-in-Chief

²⁸ “[Oversight of A.I.: Rules for Artificial Intelligence](#)”, United States Senate Subcommittee on Privacy, Technology, and the Law, May 16, 2023 ([Transcript](#))

²⁹ [H.R.3831 \(118th Congress\)](#), accessed June 6, 2023

of ALM Legaltech News Stephanie Wilkins.³⁰ The use of technology doesn't diminish the importance of a lawyer's judgment, understanding of the law, and ability to advocate for their clients. Ethical duties like competence, confidentiality, diligence, and client loyalty remain paramount, irrespective of the tools a lawyer uses. Professional conduct rules have evolved to address new challenges arising from the advent of technology in law practice. The American Bar Association's Model Rules of Professional Conduct "Rule 1.1: Competence" suggests imposition of a professional responsibility on attorneys to be competent in practice and comment 8 to the rule³¹ makes clear the intended competence includes technology competence:

"To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology, engage in continuing study and education and comply with all continuing legal education requirements to which the lawyer is subject."

As Judge Herbert B. Dixon Jr. (Ret.) describes his experience exploring legal subjects, "The chatbot, ChatGPT, notwithstanding its 'artificial intelligence' label, is merely a tool to be used wisely and carefully."³² To remind attorneys of possible limitations of AI assisted documents, the Amicus voluntarily prefaces all AI-generated reports with guidance in the following form:

The use of documents produced by Large Language Models, including this summary report, should be reviewed carefully and all information should be verified by a qualified legal professional before being used.

³⁰ "[The Problem With the 'Bogus' ChatGPT Legal Brief? It's Not the Tech,](#)" Law.com, June 2, 2023 and "[Judges, Attorneys React to Federal Court Generative AI Order: Too Hot, Too Cold, or Just Right?](#)," Law.com, June 1, 2023.

³¹ [American Bar Association Rule 1.1 comment #8](#), American Bar Association

³² "[My 'Hallucinating' Experience with ChatGPT,](#)" *The Judges' Journal*, Spring 2023

CONCLUSION

The Amicus recommends the Court be mindful of the potential impact sanctions determination will have on the public's, the legal software market's, and the wider judiciary's perception of a globally important innovation driving a nationally significant industry. Generative artificial intelligence— inclusive of the enabling neural network mechanics, the layers of value-added software being built upon it, and the skill and knowledge of its many lay and professional users— is improving at a rapid pace, the Court should not draw final judgements that assume it is a static target or which could impede continued progress.

By: 

Gregory C. Belmont
dba CereBel Legal Intelligence
Amicus Curiae, pro se
750 Lexington Avenue, 9th Floor
New York, NY 10021
(631) 506-8464
belmont@grandcentralholdings.com

Date: June 7, 2023