UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE: OPENAI, INC. COPYRIGHT INFRINGEMENT LITIGATION	1:25-md-03143 (SHS) (OTW) **DECLARATION OF DEFENDANTS' EXPERT TAYLOR BERG-KIRKPATRICK IN SUPPORT OF MOTION FOR SUMMARY JUDGMENT**

I, Taylor Berg-Kirkpatrick, Ph.D, declare under penalty of perjury, pursuant to 28 U.S.C. § 1746, as follows:

1. I submit this Declaration in support of OpenAI Defendants' (collectively, "OpenAI") Motion for Summary Judgment.

I. PROFESSIONAL BACKGROUND AND QUALIFICATIONS

2. I am a tenured Associate Professor at the University of California, San Diego in the Department of Computer Science and Engineering. Previously, I was an Assistant Professor in the Language Technologies Institute at the School of Computer Science at Carnegie Mellon University. I received a PhD in Computer Science from the University of California, Berkeley in 2015 and a Bachelor of Arts degree in Computer Science and Mathematics from the University of California, Berkeley in 2009.

3. I have developed machine learning techniques for understanding both structured and unstructured human data. This includes methods for taking unstructured information, such as emails or other human-written documents, and converting them into databases, tables, or other structured file formats that can be easily searched, analyzed, and used. My research

extends to analyzing human-generated data in natural language, music, images, historical ciphers, and videos.

4. My work includes extensive research examining large language models (LLMs), including for papers I have coauthored such as Towards a Unified View of Parameter-Efficient Transfer Learning[1] and Prompt Consistency for Zero-Shot Task Generalization.[2] These works address efficient adaptation of LLMs to new data and assess the consistency of LLM behavior across varied prompts. I have also examined memorization and related privacy issues in LLMs and authored papers measuring the extent to which language models reveal information about their training data.

5. I am also the author of several papers concerning text annotation, text classification, and dialog analysis. Research for these papers required overseeing the annotation of thousands of conversations to identify certain characteristics and content.

6. I am the recipient of numerous awards, including a Schmidt Science of Trustworthy AI Award in 2025, a Schmidt HAVI Award in 2025, a Google Trust & Safety Award 2023, a National Science Foundation CAREER Award in 2022, an Amazon Research Award in 2019, a Google Faculty Award in 2017, and the Dorothea Klumpke Roberts Prize in Mathematics in 2009.

II. ASSIGNMENT

7. I have been engaged on behalf of OpenAI to provide an expert opinion regarding whether ChatGPT can be induced to output verbatim or near-verbatim excerpts of the Asserted

[1] Junxian He et al., *Towards a Unified View of Parameter-Efficient Transfer Learning*, International Conference on Learning Representations (Feb. 2, 2022), https://arxiv.org/abs/2110.04366.

[2] Chunting Zhou et al., *Prompt Consistency for Zero-Shot Task Generalization*, *in* Findings of the Association for Computational Linguistics: EMNLP 2022 2613–2626 (Dec. 2022), https://aclanthology.org/2022.findings-emnlp.192/.

Works and, if so, what techniques are required to induce such outputs. I was also asked to provide opinions regarding how ordinary users use ChatGPT and how ChatGPT responds to user requests to obtain or access the Plaintiffs' Asserted Works in the ordinary course.

8. I submitted an opening expert report in this matter on April 24, 2026; a rebuttal report on May 29, 2026; a brief addendum to my rebuttal report on June 2, 2026; a reply report on June 26, 2026; and a supplemental response to Dr. Tom Goldstein's supplemental report on July 29, 2026. I was deposed in this case on August 3 and August 4, 2026. The statements in my reports are accurate and true to the best of my knowledge, information, and belief.

9. This declaration provides a concise summary of the opinions set forth in my reports concerning whether and the extent to which ChatGPT can be induced to output verbatim or near-verbatim excerpts of Plaintiffs' Asserted Works. These opinions are based on my knowledge, skill, experience, training, and education, as well as my review and analysis of the documents and data as disclosed in my reports.

III. CHATGPT DOES NOT OUTPUT THE PLAINTIFFS' WORKS IN ORDINARY USE

10. To assess the frequency with which ChatGPT outputs verbatim or near-verbatim excerpts of Plaintiffs' Asserted Works in real world interactions, I conducted a Conversation Regurgitation Study. For this study, I analyzed a stratified random sample of 20 million ChatGPT conversation logs provided to me by OpenAI. These conversation logs were sampled randomly from ChatGPT conversations during the period at issue and show the kinds of ChatGPT conversations that users have with the OpenAI Models. The sample included conversations with outputs from GPT-3.5, GPT-3.5-turbo, GPT-4, GPT-4-turbo, GPT-4o, and GPT-4o-mini.

11. From these 20 million ChatGPT conversation logs, I selected all user-visible textual ChatGPT outputs to analyze. In total, the analysis included approximately 95 million user-visible messages from ChatGPT.

12. I compared these user-visible textual ChatGPT outputs with copies of all the Plaintiffs' Asserted Works from the datasets Plaintiffs provided in this case. These datasets include approximately 7.5 million total Asserted Works provided by the five News Plaintiffs and 204 provided Asserted Works from 14 Class Plaintiffs. While the New York Times is asserting over 6 million total works, they produced the texts of fewer than 3 million of them.[3]

13. Before searching for verbatim matches between the ChatGPT conversation logs and Plaintiffs' Asserted Works, I pre-processed the raw text in both datasets. Pre-processing text is important to ensure that the matching process is not hindered by non-substantive differences between documents such as capitalization, punctuation, or spacing.

14. For my studies, I decomposed texts into overlapping sequences of words to assess the similarity by the extent to which those sequences are shared between two corpora. An effective approach to checking whether two texts contain identical overlap is to break each text into smaller sequences and then search for matches for those substrings within the comparison text corpus. These contiguous sequences of items extracted from a text are also referred to as N-grams where "N" refers to a number of items extracted from a text, typically either words or tokens. Small values of "N" identify more instances of overlap because of many common sequences of words in human language. As the chosen value of

[3] Berg-Kirkpatrick Opening ¶ 122.

N increases, the number of identified overlaps between two pieces of unrelated text will decrease.

15. Another important metric is the "edit distance" between two strings. "Edit distance" measures the similarity between two strings by counting the number of insertions, deletions, and substitutions needed to convert one string to another. I performed robustness checks for my analyses using a 2-word edit distance.

16. I employed a 25-word match methodology for this analysis. This value is consistent with academic literature on LLM memorization and regurgitation. I reviewed a number of academic studies finding that an N of 25 is likely to capture language potentially reused from another source, rather than language that is similar by chance. I therefore determined that applying a 25-word match would capture instances where ChatGPT may have regurgitated text from the Plaintiffs' works.

17. I defined a "ChatGPT output of a Plaintiff's Asserted Work" (which Plaintiffs in this case have referred to as a "regurgitation") as any sequence of text in a ChatGPT output with the following conditions: 1) is visible to the user as a ChatGPT-generated message; 2) is at least 25 words long; 3) is a verbatim match for a 25-word sequence of text from at least one of the Plaintiff's Asserted Works; 4) is not a boilerplate sequence or Third-Party material; and 5) did not appear in any prompt or uploaded file provided by the user before ChatGPT generated the output.[4]

18. To begin my analysis, I first divided each output and each Asserted Work into every consecutive sequence of 25 words, called a "25-gram," and identified sequences that appeared in both a ChatGPT output and an Asserted Work.

[4] Berg-Kirkpatrick Opening ¶ 184.

19. After identifying the 25-word matches, I applied a sequential automated filtering process to remove raw matches that were not considered instances of ChatGPT regurgitating a portion of an Asserted Work.

    a. First, I employed automated filters to remove boilerplate sequences, such as repeated words, lists, dates, numbers, and Lorem Ipsum.

    b. Second, I used automated filters to remove public-domain and other text written by someone other than the Plaintiffs, including religious and historical texts, famous quotes, text in the public domain, and lyrics/poetry.

    c. Third, I removed matches that appeared in a user's prompt or uploaded file, because those matches showed ChatGPT repeating text supplied by the user rather than independently producing it.[5]

20. After applying these automatic filters, I instructed annotators to manually review all remaining ChatGPT outputs with 25-gram matches with the Plaintiffs' Asserted Works. The purpose of this manual analysis was to determine whether there existed a separate source for the matched 25-grams in the output or whether the match was a true case of ChatGPT outputting an excerpt of the Plaintiffs'Asserted Works. Through this manual analysis, I identified and filtered out many instances of text not authored by Plaintiffs, including from religious and legal texts, famous speeches, and non-news media content like books, podcasts, and movies.

21. I considered all matches remaining after automatic filtering and manual review (those that did not have an identifiable third-party source) to be an instance of ChatGPT outputting an excerpt of Plaintiff-authored text from an Asserted Work.

---

[5] Berg-Kirkpatrick Opening ¶¶184-200, 377-379.

22. As a robustness check, I also searched for near-verbatim matches. I allowed up to two word insertions, deletions, or substitutions within a 25-word sequence in order to capture small wording changes.

23. Across all approximately 95 million ChatGPT-authored, user-visible messages in the 20 million ChatGPT conversation logs, I found that ChatGPT output verbatim 25-word spans from Plaintiffs' Asserted Works 31 times. I also identified 19 additional near-verbatim outputs, for a total of 50 verbatim and near-verbatim matches.[6]

24. The longest verbatim output I identified was 235 words, and the single longest verbatim output constituted 15.06% of its source text.[7]

25. The results of the analysis demonstrate that fewer than 0.000033% of ChatGPT outputs in the 20 million ChatGPT conversation logs contain a verbatim excerpt of Plaintiffs Asserted Works. Overall, verbatim excerpts of News Plaintiffs' Asserted Works appeared in less than 0.000023% of outputs and verbatim excerpts of named Books Plaintiffs' Asserted Works appeared in less than 0.00001% of outputs.[8]

26. **Table 1** below lists the percentage of ChatGPT outputs in the sample that contain verbatim excerpts of Plaintiff-generated text from Asserted Works on a per-plaintiff basis:

[6] Berg-Kirkpatrick Opening ¶ 181.
[7] *Id.*
[8] Berg-Kirkpatrick Opening ¶ 185.

Table 1: Summary of % Outputs Containing Regurgitation of Plaintiffs Works.

Plaintiff	Prevalence of Messages Including Regurgitation (out of 95M user-visible outputs)	
	Number of Instances	Rate of Regurgitation
NYTimes	11	0.000012%
DNP	0	0.0%
CIR	0	0.0%
Intercept	0	0.0%
Ziff Davis	11	0.000012%
George R.R. Martin	8	0.000008%
Jodi Picoult	1	0.000001%
All Other Books Plaintiffs	0	0.0%

27. In responding to my Conversation Regurgitation Study, News Plaintiffs' experts Dr. Goldstein and Mr. Delorey noted that the corpora I used to filter out boilerplate and third-party material contained limited excerpts from some News Plaintiff Asserted Works. To determine the effect (if any) of these inclusions, I checked whether ChatGPT output these quotes during any Merits Sample conversations. I found that, conservatively, at most two additional possible regurgitations of 25-word excerpts of News Plaintiffs' Asserted Works may exist in the 95 million ChatGPT outputs. This raised my upper bound on the prevalence of 25-word regurgitations of News Plaintiffs' content from 0.000023% to 0.000025%, an infinitesimal change with no impact on my overall conclusions.[9]

[9] Berg-Kirkpatrick Reply ¶ 177.

28. These statistics can also be computed as a per-conversation frequency. I identified 25-word verbatim regurgitations of News Plaintiffs' content in 24 outputs out of 20 million conversations, which corresponds to 0.00012%.

29. For the Class Plaintiffs specifically, the alleged regurgitation rate was only 0.000015%, and none of the identified outputs included more than 0.017% of a Class Asserted Work. I identified only 14 instances of verbatim or near-verbatim regurgitation of a small amount of text from their Asserted Works. Of these 14 instances of verbatim or near-verbatim excerpts, twelve were 32 words or less (the two outliers were 67 and 122 words).[10]

30. This figure can also be computed as a per-conversation frequency. I identified 25-word verbatim or near-verbatim regurgitations of Class Plaintiffs' content in 14 outputs out of 20 million conversations, which corresponds to 0.00007%.

31. **Table 2** below identifies the 9 ChatGPT outputs containing verbatim regurgitations associated with Class Plaintiffs' Asserted Works and the percentage of the associated work that appears in the output. **Table 3** below similarly shows the 5 additional ChatGPT outputs which contained a near-verbatim regurgitation of a Class Plaintiffs' Asserted Work, along with the percentage of the associated work that appears in the output.

[10] Berg-Kirkpatrick Opening ¶ 241 Table 33, ¶ 420 Table A.42.

Table 2: Summary of Verbatim Regurgitations (Books)

Author Name	Book Name	Message ID	ChatGPT Message Length (Words)	Regurgitation Length (Words)	% of Message	% of Work
George R.R. Martin	A Clash of Kings	009e310aa620223f20fc405921c8e7d9	293	28	9.56%	0.004%
		5970553a72bd3599dc546f771579c14a	68	28	41.18%	0.004%
		6b31f38635919544a7ae01d1618b8148	121	27	22.31%	0.004%
		4a95c8fb25baef9aae9bad66673a1b20	261	67	25.67%	0.009%
	A Game of Thrones	f1947221559c6632950ec938810ad40c	122	122	100.0%	0.016%
		49624629a355cd3fb04de8379085bceb	134	29	21.64%	0.004%
		749b5e30736202e8be769b959e1abbff	381	29	7.61%	0.004%
	A Dance With Dragons	6d78ca5e1c675c80d3286daaa996e460	214	28	13.08%	0.006%
Jodi Picoult	Nineteen Minutes	3045f9a94f887f277e46fecd6ac5f479	629	32	5.09%	0.005%

Table 3: Summary of Near-Verbatim Regurgitations (Books)

Author Name	Work Name	Message ID	ChatGPT Message Length (Words)	Regurgitation Length (Words)	% of Message	% of Work
Jodi Picoult	Keeping Faith	1d8bd125e8d20b6080d2a5e3330e4622	34	23	67.65%	0.013%
		67852886369efe7ca3d1c5e161aba192	516	23	4.46%	0.013%
	Change of Heart	513c43b6a49481d6b1e1de55522f4def	237	26	10.97%	0.017%
George R.R. Martin	A Storm of Swords	cee4130eac007399fad1d10581514d1d	320	24	7.50%	0.005%
		0c876e0f3e9330f4019439d93cfb8e0a	322	23	7.14%	0.005%

32. The results of the near-verbatim robustness check demonstrated that fewer than 0.000053% of the ChatGPT outputs contained text within a 2-word difference of a verbatim 25-word

sequence of Plaintiff-generated text from an Asserted Work—an increase of only 0.00002% over the verbatim-only figure.[11]

33. Based on my finding that a 25-word verbatim or near-verbatim text regurgitation appeared in approximately one in every 2 million ChatGPT outputs I analyzed, I concluded that most ChatGPT users will never encounter a verbatim output from Plaintiffs' Asserted Works.[12]

IV. CHATGPT DOES NOT OUTPUT THE PLAINTIFFS' ASSERTED WORKS IN TARGETED TESTS

34. To assess whether ChatGPT would output News Plaintiffs' Asserted Works when asked about the subject matter of those works, I conducted a Guided Regurgitation Study. This study tested whether asking ChatGPT about the topics covered in News Plaintiffs' Asserted Works would cause it to regurgitate the text of those works.

35. For this study, I developed prompts closely related to the News Plaintiffs' works, then submitted them to GPT-3.5 turbo, GPT-4, and GPT4o and compared GPT's responses with the News Plaintiffs' Asserted works.

36. I used 710 randomly selected Asserted Works in my analysis, including works from 19 outlets across all 5 News Plaintiff groups. I selected a portion of the Asserted Works because results from Plaintiffs' searches indicated they may have been included in GPT training datasets.

37. I used a set of 21 template prompts for my study, each of which was modified to be specific to each target article by incorporating three descriptions of the target article content at varying levels of detail.

[11] Berg-Kirkpatrick Opening ¶ 188.
[12] Berg-Kirkpatrick Opening ¶ 181.

38. The prompts varied by specificity to induce more or less specific outputs, from general requests for information to targeted asks for the information in a specific work.

39. For the descriptions of target article content, I used the GPT-5 model via the OpenAI API to generate topic descriptions at three levels of specificity—low (a two-to-five-word description of the article's general theme), medium (five to fifteen words, enough to identify the article among several on the same general topic), and high (fifteen words or longer, enough detail to easily find the specific article on the web).[13]

40. I combined those three specificity levels with the 21 prompt templates yielding 63 prompts per target work.

41. Next, I submitted all 63 prompts for each of the 710 target works to the OpenAI batch API, totaling 68,670 prompts. The prompts were tested across three GPT models listed above.[14]

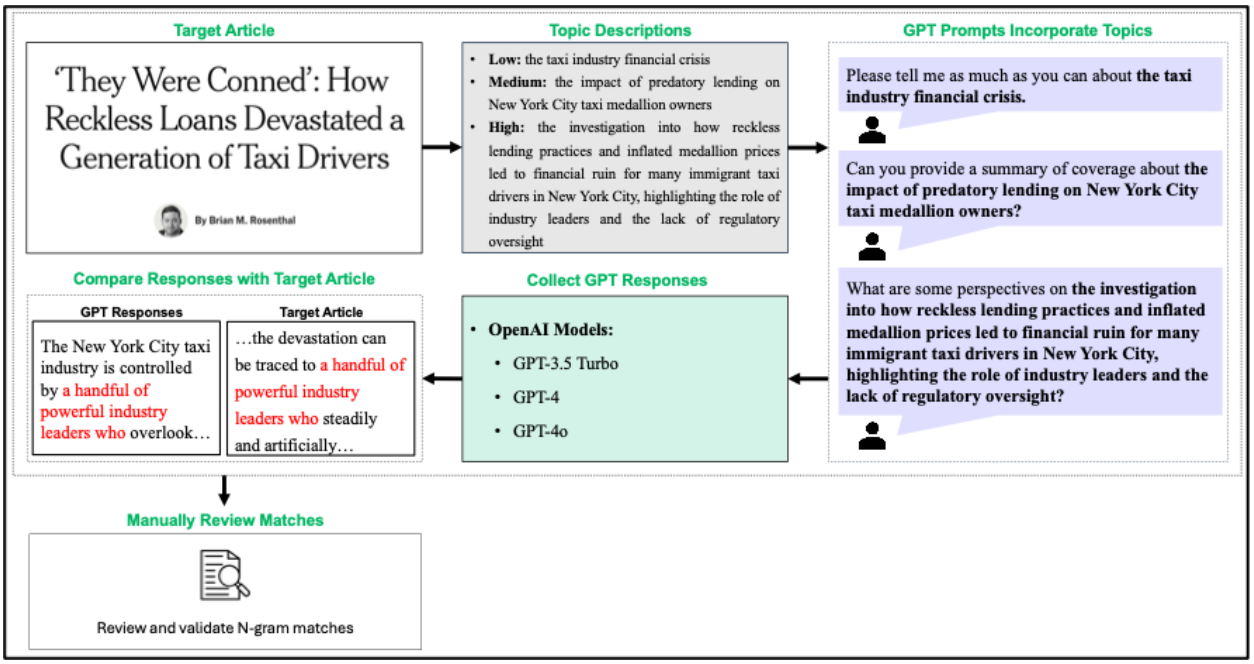
42. Across all 68,670 prompts, I found zero instances of regurgitation across every GPT model tested and every Plaintiff group. I reached this result by comparing the API responses against the News Plaintiffs' Asserted Works for overlapping text of at least 25 words and conducting a manual review of each match to determine whether it met the definition of a regurgitation explained *supra*.[15] ChatGPT generated new text in response to every one of my prompts.[16]

43. **Figure 1** below summarizes the experiment workflow for my **Guided Regurgitation Study**.

---

[13] Berg-Kirkpatrick Opening ¶301.
[14] Berg-Kirkpatrick Opening ¶250.
[15] Berg-Kirkpatrick Opening ¶251.
[16] Berg-Kirkpatrick Opening ¶¶18, 246.

Figure 1: Summary of Workflow (Opening Report Figure 33)



44. The results of my analysis illustrated that the GPT models in my study do not regurgitate the News Plaintiffs' Asserted Works even when presented with prompts closely related to those works.[17]

45. In conducting this analysis, I also found that the models also frequently refused requests that explicitly asked for article text. For prompts that directly requested the content of the targeted work, refusal rates averaged above 90% across several templates.[18]

46. Based on these results, I concluded that GPT models do not regurgitate the News Plaintiffs' Asserted Works, even when directly prompted to do so, and instead frequently refuse to comply with such requests.[19]

[17] Berg-Kirkpatrick Opening ¶ 255.
[18] Berg-Kirkpatrick Opening ¶ 253.
[19] Berg-Kirkpatrick Opening ¶ 255.

V. CHATGPT DOES NOT OUTPUT THE PLAINTIFFS' WORKS IN ADVERSARIAL TESTS

47. I also conducted a separate study to assess whether a technically sophisticated user could induce GPT models to output a selected work using prompts deliberately designed to overcome ChatGPT's ordinary safeguards.

48. I randomly selected a set of representative Asserted Works, including works from all Plaintiffs. In total, I targeted 24 selected works, one randomly selected from each of the 19 outlets across 5 News Plaintiffs and 5 from selected Books Plaintiffs.

49. First, I experimented with adversarially requesting text from one New York Times Asserted Work using the ChatGPT interface. I developed 15 prompts that directly requested article text, supplied portions of article text with a continuation request, and sought to extract article text incrementally through multiple article turns. I found that no attempted method retrieved any percentage of that one work from ChatGPT.

50. I repeated each of the 15 prompts on one work from each of the News Plaintiffs' publications and found that my best attempt was, on average, able to extract 2.3% of the contents of an article.[20]

51. Using the same techniques, I also created 11 adversarial prompts requesting text from selected books from Class Plaintiffs' Asserted Works. My best attempt was on average able to extract 5.6 words from each work, or 0.004% of the full book text.[21]

52. Next, I tested prompts on the OpenAI API, which requires technical expertise to configure. I attempted to extract content from the same articles—one from each News Plaintiff—

[20] Berg-Kirkpatrick Opening ¶¶19, 259.
[21] Berg-Kirkpatrick Opening ¶ 260.

while explicitly prompting the model to provide verbatim article content and adjusting other control parameters.

53. Measuring the success of each extraction attempt by calculating the longest sequence of words overlapping between the OpenAI API's output and the target work, I still found that I was able to extract only 3.6% of the contents on my best attempts.[22]

54. Using the same attempts once more on the OpenAI API, but with the selected Class Plaintiffs' Asserted Works, I was able to extract an average of only 14.2 words (comprising on average 0.0046% of the respective works) on my best attempts for each work.[23]

55. Based on the results of this study, I concluded that adversarial techniques designed to extract texts from GPT models are largely ineffective at retrieving the selected Asserted Works (both news articles and books).[24]

56. The adversarial techniques I employed in the Forced Regurgitation Study require a significant investment of time and effort. Notably, I spent over an hour writing 15 prompts adversarially requesting text from a single selected work, which is considerably more time than a casual ChatGPT user would likely spend to obtain a piece of information.

57. The techniques I employed to submit prompts via the OpenAI API also required substantial technical expertise, including writing computer code to submit requests to OpenAI's systems and understanding how to systematically vary model parameters. In my opinion, no rational consumer seeking to access Plaintiffs' Asserted Works would attempt to

[22] Berg-Kirkpatrick Opening ¶ 263.
[23] Berg-Kirkpatrick Opening ¶ 264.
[24] Berg-Kirkpatrick Opening ¶ 265.

procure a book or news article through ChatGPT or the OpenAI API using adversarial techniques.[25]

a. News Plaintiffs' experts Dr. Goldstein and Mr. Delorey's studies corroborate these results.

58. In preparing my Rebuttal Report, I analyzed methodologically similar regurgitation studies conducted by News Plaintiffs' experts Dr. Goldstein and Mr. Delorey. Dr. Goldstein performed these studies on twelve OpenAI models.[26] Mr. Delorey performed his studies on nine OpenAI models.[27]

59. Two of the models both experts prompted as part of their analysis were OpenAI's gpt-4-base and code-davinci-002 models. I understand that gpt-4-base and code-davinci-002 are unaligned base models that were never publicly available nor intended for public use.[28]

60. In these analyses the experts submitted millions of prompts attempting to induce OpenAI's models to output News Plaintiffs' Asserted Works.[29]

61. These experts did not originally produce all the prompts they executed in their analyses. They produced full sets of prompts in June 2026. At their depositions, Dr. Goldstein confirmed that he used approximately 195 million total prompts[30] and Mr. Delorey confirmed that he used approximately 65 million total prompts[31].

62. Despite his efforts, Dr. Goldstein was only able to elicit what he defined as a regurgitation in 0.092% of attempts (179,224 outputs).[32] In his Reply Report, Dr. Goldstein revised this

[25] Berg-Kirkpatrick Opening ¶ 265.
[26] Berg-Kirkpatrick Rebuttal ¶ 53.
[27] Berg-Kirkpatrick Rebuttal ¶ 54.
[28] *See* Declaration of Bright Kellogg, MDL ECF 819-1.
[29] Berg-Kirkpatrick Rebuttal ¶¶53-54.
[30] Goldstein 073126 Dep. 71:15-25.
[31] Delorey Dep. 232:25-233:5
[32] Berg-Kirkpatrick Rebuttal ¶94.

figure down to 176,568 (0.091%).[33] He alleged that these 176,568 outputs contained at least one 16-word excerpt from approximately 87,000 NYT Asserted Works and 28,000 DNP Asserted Works.[34]

63. I found that, on average, generations that Dr. Goldstein claimed contained "regurgitations" featured 4.4% of the "regurgitated" Asserted Work. For each Asserted Work that Dr. Goldstein alleged was regurgitated in his analysis, I computed the percentage of the work that was output by ChatGPT by dividing the total number of words in each output's matched 16-grams (as identified by Dr. Goldstein) by the total number of words in the corresponding article text.[35]

64. Mr. Delorey was able to elicit what he defined as a regurgitation[36] in 2.5% of attempts (1,652,219 outputs).[37] In his Reply Report, Mr. Delorey revised this figure down to 1,451,711 outputs (2.2%).[38] He claimed that these outputs contained, collectively, excerpts from 357,310 Ziff Davis Asserted Works (16%). He also determined that the outputs contained 10-word excerpts from 172,952 Ziff Davis Asserted Works.[39]

---

[33] Goldstein Reply Table 4.

[34] Berg-Kirkpatrick Rebuttal ¶93.

[35] Berg-Kirkpatrick Rebuttal Report ¶120.

[36] Mr. Delorey's definition of "regurgitation" required much less matching text than Dr. Goldstein's. Dr. Goldstein required that the model output and Asserted Work share at least one 16-word verbatim sequence. Mr. Delorey's definition classified outputs with as little as four words in common with an Asserted Work as "regurgitations."

[37] Delorey Reply ¶ 134, Berg-Kirkpatrick Rebuttal Table 16.

[38] Delorey Reply ¶128.

[39] These statistics were referenced but not precisely presented in Mr. Delorey's Reply Report. I computed them from his produced data. Mr. Delorey's updated "probe_reproduction_results.parquet" file produced in his Ninth Supplemental Production (6/22/2026) contains "regurgitations" related to 357,310 Ziff Davis Asserted Works. "Regurgiations" related to 172,952 Ziff Davis Asserted Works remain after applying the first and second "Tightenings" described in ¶77 of Mr. Delorey's Reply Report, which filters to "regurgitations" with a "full 10 word run of contiguous verbatim text[.]"

65. In his Reply Report, Mr. Delorey computed "how much of an article comes back when it does reproduce" and determined that, in instances that he claimed constituted "regurgitations," GPT models reproduced between 1.48% and 4.27% of the article (depending on the model).[40] He did not provide his statistics for each model, only the upper and lower bounds.

66. I thus compute the following statistics: between Dr. Goldstein and Mr. Delorey's analyses the experts claimed at least 10-word "regurgitations" of approximately 288,000 New York Times, Daily News Plaintiffs, and Ziff Davis Asserted Works (87,000 NYT, 28,000 DNP, and 173,000 ZD), out of the roughly 11 million total works these Plaintiffs are asserting in this case. This represents 2.6% of all Asserted Works.

67. Furthermore, in the rare instances when Dr. Goldstein and Mr. Delorey alleged that GPT models "regurgitated" Plaintiffs Asserted Works, both experts on average extracted less than 4.4% of the Asserted Work (under 4.4% for Dr. Goldstein and a per-model upper bound of 4.3% for Mr. Delorey).

b. Class Plaintiff's expert Dr. Shan's study corroborates these results.

68. In preparing my Reply Report, I analyzed a methodologically similar forced regurgitation study conducted by Class Plaintiffs' expert Dr. Shan, in which he submitted approximately 5.3 million prompts attempting to induce OpenAI's models to output Class Plaintiffs' Asserted Works.[41]

[40] Delorey Reply ¶101.
[41] Berg-Kirkpatrick Reply ¶¶322, 341.

69. Despite his voluminous attempts, Dr. Shan was only able to elicit what he defined as a regurgitation in 0.53% attempts, sourced from only 40 of the 212 books asserted by the Class Plaintiffs.

70. Dr. Shan only generated three excerpts that exceeded 300 contiguous words; each of these instances came from a George R.R. Martin book.[42] As shown in **Table 4** below, each of these alleged reproductions account for less than 1% of the content of the corresponding Asserted Work. The longest span of regurgitated text he identified was a 1,899 word sequence from A Game of Thrones, representing 0.62% of that 305,507-word book.

71. Each of these excerpts are available through ordinary internet searches or publicly accessible websites, including Amazon kindle samples for the books.[43] Moreover, the results of my review demonstrate that public availability is not limited to isolated examples but is present among some of the most notable passages underlying Dr. Shan's findings.[44]

Table 4: Share of Words From the Three Longest Reproductions Alleged by Dr. Shan

Asserted Work Title	Longest Block of Words Allegedly Reproduced from Asserted Work	Word Count of Asserted Work	Share of Allegedly Reproduced Words
A Game of Thrones	1,899 words	305,507 words	0.62%
A Clash of Kings	1,458 words	324,551 words	0.45%
A Storm of Swords	304 words	432,066 words	0.07%

[42] Berg-Kirkpatrick Reply ¶¶322, 341.
[43] Berg-Kirkpatrick Reply ¶¶277-279, Figure 6
[44] Berg-Kirkpatrick Reply ¶352.

72. In total, the 211 Class Plaintiffs' Asserted Works comprise 25,262,536 words.[45] The 39,267 words Dr. Shan alleges he reproduced therefore represent ~0.16% of the total words contained in those Asserted Works.[46]

73. The fact that Dr. Shan was able to identify only a vanishingly small fraction of allegedly reproduced text from the Asserted Works despite conducting approximately 5.3 million regurgitation prompting attempts supported my opinion that his adversarial prompting attacks are, on the whole, largely unsuccessful. Dr. Shan's results corroborated my conclusion that adversarial methods are ineffective at inducing GPT models to reproduce Plaintiffs' Asserted Works.[47]

VI. CHATGPT OUTPUTS TYPICALLY INCLUDE ATTRIBUTION INFORMATION

74. In preparing my Rebuttal Report, I was also asked to review all instances of "regurgitation" claimed by all experts in conversation records produced in this case and determine whether the source of the "regurgitated" text was clear to the user in the identified ChatGPT outputs.[48]

75. I analyzed all 68 instances of regurgitation identified by myself, Dr. Shan, and Dr. Goldstein in our Opening Reports.[49] The regurgitations were identified in the 20M (Merits Sample), 78M, and 10M conversation datasets, *i.e.* from a pool of 108M total conversations.[50]

---

[45] Although the class list contains 212 Asserted Works, I received OCR text files for only 211 of them. Nevertheless, the omission of a single Asserted Work does not materially impact my analysis or opinions of Dr. Shan's regurgitation prompting analysis.

[46] Berg-Kirkpatrick Reply ¶324.

[47] Berg-Kirkpatrick Reply ¶322.

[48] Berg-Kirkpatrick Rebuttal ¶¶251-252.

[49] Plaintiff experts have continued to identify purported "regurgitations" in their rebuttal, reply, and supplemental reports. Plaintiffs' experts have not argued that the "regurgitations" identified in these experts' subsequent reports are characteristically different from those in their opening reports.

[50] Berg-Kirkpatrick Rebuttal ¶254.

76. For my analysis, I considered three ways the source of "regurgitated" text could be made clear to the user: (1) if the user requested text from a specific source, and ChatGPT provided the text; (2) if ChatGPT identified the source it quoted to the user; and (3) if ChatGPT provided a link to the source it is quoting to the user.

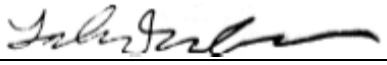
77. I found that in 63 of the 68 identified instances—92.7%—the source of the allegedly regurgitated text would have been obvious to the user. In the majority of these cases, the user requested the specific article by name, provided its URL, or otherwise identified the work in the conversation before ChatGPT produced the matched output. In other instances, the user generically requested quotes related to a topic and ChatGPT identified the specific source in its output. Therefore, I concluded that, in all but five total identified instances by both parties, any "regurgitation" could not have been "concealed" because the context of the conversation clearly identified the source of the "regurgitated" material.[51]

78. Based on this analysis, I concluded that even in the rare instances (68 total across 108 million produced conversations and 375 million ChatGPT outputs) where ChatGPT produced text matching a portion of an Asserted Work, the output did not function to conceal the identity of the work's source or author from the user. To the contrary, in nearly all identified cases, the user was already aware of the work's identity because the user had supplied that information to ChatGPT or ChatGPT named the source in its response.

[51] Berg-Kirkpatrick Rebuttal ¶253.

I declare under penalty of perjury that the foregoing is true and correct, and that I have personal knowledge of these facts and am competent to testify thereto if called upon to do so.

Dated: September 3, 2026
San Diego, California

By: 
Taylor Berg-Kirkpatrick