

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

In Re: OpenAI, Inc. Copyright Infringement
Litigation

This Document Relates To:

No. 1:23-cv-11195
No. 1:24-cv-01515
No. 1:24-cv-03285
No. 1:24-cv-04872
No. 1:25-cv-04315

Case No. 1:25-md-3143 (SHS) (OTW)

CONTAINS INFORMATION
DESIGNATED AS CONFIDENTIAL
UNDER THE STIPULATED
PROTECTIVE ORDER

NEWS PLAINTIFFS' COUNTERSTATEMENT OF MATERIAL FACTS

REDACTED

TABLE OF CONTENTS

I. Defendants’ Fair Use Defenses.....	1
A. News Plaintiffs’ Expressive Content	1
B. Defendants’ Exploitation of News Plaintiffs’ Expressive Content.....	7
1. Use of News Plaintiffs’ Content for Training	7
2. Use of News Plaintiffs’ Content for Retrieval Augmented Generation.....	13
3. Use of News Plaintiffs’ Content for Display in Chatbot Outputs	14
C. Defendants’ Copyright and Antiregurgitation Measures	19
II. Defendants’ Implied License Defense.	22
III. Plaintiffs’ DMCA Claims.	26
A. Intentional Removal of CMI.....	26
B. Concealment of Output Infringement	27
C. OpenAI’s Knowledge of Regurgitation and the Inefficacy of its Mitigation Measures....	32
D. Inducement, Enablement, and Facilitation of Infringement by OpenAI’s Users	34
E. Concealment of Training Infringement	35
F. Facilitation of Training Infringement	37
IV. The Impact of Defendants’ Conduct on Plaintiffs and Others in the News Media Industry ..	38

TABLE OF CITATION CONVENTIONS

Citation	Document	ECF No.
NP Ex.	News Plaintiffs' Exhibits, which were attached to the September 4, September 17, and October 23 Declarations of Davida Brook and Ian Crosby	1777, 1785, 1786, 1790, 1792, 1799, 1803, 1807, 1809, 1812, 1823, 1829, 1830, 1833, 1834, 1837, 1839, 1841, 1842, 1843, 1844, 1846, 1847, 1849, 1851, 1852, 1854, 1856, 1857, 1861, 1863, 1869, 1871, 1874, 1876, 1877, 1878, 1880, 1882, 1885, 1887, 1888, 1891, 1893, 1896, 1897, 1899, 1902, 1903, 1905, 1909, 1911, 1914, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1989, 1990, 1991, 1994, 1995-1999, 2001, 2003, 2004, 2009, 2013-15, 2017, 2019, 2020, 2022
Bauerlein Decl.	Declaration of Monika Bauerlein	1788
Bell Opening Rep.	Declaration of Emily Bell ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1794
Bell Rebuttal Rep.	Declaration of Emily Bell ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1794

Bell Reply Rep.	Declaration of Emily Bell ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment	Filed Concurrently with this Brief
Cross Rebuttal Rep.	Declaration of Sue Cross (Exhibit A)	1821-1
Delorey Opening Rep.	Declaration of Daniel P. Delorey ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1774-1
Delorey Rebuttal Rep.	Declaration of Daniel P. Delorey ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1774-2
Delorey Reply Rep.	Declaration of Daniel P. Delorey ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1774-3
Delorey Supplemental Rep.	Declaration of Daniel P. Delorey ISO News Plaintiffs' Motion for Summary Judgment (Exhibit D)	1774-4
Etzioni Opening Rep.	Declaration of Oren Etzioni ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1806
Etzioni Rebuttal Rep.	Declaration of Oren Etzioni ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1806
Etzioni Reply Rep.	Declaration of Oren Etzioni ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1806
Goldstein Decl.	Declaration of Tom Goldstein ISO News Plaintiffs' Motion for Summary Judgment	1819
Goldstein Opening Rep.	Declaration of Tom Goldstein ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1819

Goldstein Rebuttal Rep.	Declaration of Tom Goldstein ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1819
Goldstein Reply Rep.	Declaration of Tom Goldstein ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1819
Goldstein Supplemental Rep.	Declaration of Tom Goldstein ISO News Plaintiffs' Motion for Summary Judgment (Exhibit D)	1819
Gratz Decl.	Declaration of Joseph C. Gratz Motions for Summary Judgment, Vols. 1–4	1740, 1741, 1747, 1752
Koonce Decl.	Declaration of Lacy H. Koonce III ISO News Plaintiffs' Motion for Summary Judgment	1859
Maisel Decl.	Declaration of Jennifer B. Maisel ISO News Plaintiffs' Motion for Summary Judgment	1815
Malackowski Rebuttal Rep.	Declaration of Jim Malackowski ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1800
Malackowski Reply Rep.	Declaration of Jim Malackowski ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1800
Malackowski ZD Opening Rep.	Declaration of Jim Malackowski ISO News Plaintiffs' Motion for Summary Judgment (Exhibit F)	1800
Malackowski ZD Rebuttal Rep.	Declaration of Jim Malackowski ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment (Exhibit G)	Filed Concurrently with this Brief

Malackowski ZD Reply Rep.	Declaration of Jim Malackowski ISO News Plaintiffs' Motion for Summary Judgment (Exhibit H)	1800
McDonald Rebuttal Rep.	Declaration of Michael McDonald (Exhibit A)	1875-1
Mitchell Opening Rep.	Declaration of Melanie Mitchell ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment (Exhibit A)	Filed Concurrently with this Brief
Mitchell Rebuttal Rep.	Declaration of Melanie Mitchell ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment (Exhibit B)	Filed Concurrently with this Brief
Mitchell Reply Rep.	Declaration of Melanie Mitchell ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment (Exhibit C)	Filed Concurrently with this Brief
October 23, 2026 Delorey Decl.	Declaration of Daniel P. Delorey ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment	Filed Concurrently with this Brief
Oh Decl.	Declaration of Karen Oh ISO News Plaintiffs' Motion for Summary Judgment	1802
Owen Opening Rep.	Declaration of Taylor Owen ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1811
Owen Rebuttal Rep.	Declaration of Taylor Owen ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1811

Owen Reply Rep.	Declaration of Taylor Owen ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1811
Second Bauerlein Decl.	Second Declaration of Monika Bauerlein	Filed Concurrently with this Brief
Sherrard Decl.	Declaration of Michael Sherrard	Filed Concurrently with this Brief
Stricchiola NYT/DNP Opening Rep.	Declaration of Jessie Stricchiola ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1798
Stricchiola Rebuttal Rep.	Declaration of Jessie Stricchiola ISO News Plaintiffs' Motion for Summary Judgment (Exhibit C)	1798
Stricchiola Reply Rep.	Declaration of Jessie Stricchiola ISO News Plaintiffs' Motion for Summary Judgment (Exhibit D)	1798
Stricchiola ZD Opening Rep.	Declaration of Jessie Stricchiola ISO News Plaintiffs' Motion for Summary Judgment (Exhibit B)	1798
Supplemental Goldstein Decl.	Supplemental Declaration Tom Goldstein ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment	Filed Concurrently with this Brief
Tease Opening Rep.	Declaration of Antoinette Tease (Exhibit A)	Filed Concurrently with this Brief
Waehrer Rebuttal Rep.	Declaration of Keith Waehrer ISO News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment	Filed Concurrently with this Brief
Wenger Opening Rep.	Declaration of Emily Wenger (Exhibit A)	1858-1

Wenger Rebuttal Rep.	Declaration of Emily Wenger (Exhibit D)	1858-4
Wenger Reply Rep.	Declaration of Emily Wenger (Exhibit F)	1858-6
Wenger Supplemental Rep.	Declaration of Emily Wenger (Exhibit H)	1858-8
Wu Rebuttal Rep.	Declaration of Timothy Wu ISO News Plaintiffs' Motion for Summary Judgment (Exhibit A)	1808-1

Pursuant to Local Civil Rule 56.1(b), Plaintiffs The New York Times (“The Times”), Daily News Plaintiffs¹ (“DNP”), the Center for Investigative Reporting, Inc. (“CIR”), the Intercept, and the Ziff Davis Plaintiffs (“Ziff Davis”) (each a “News Plaintiff” and collectively, “News Plaintiffs”) respectfully submit this counterstatement of material facts.

Among the issues on which Defendants have moved for summary judgment, News Plaintiffs have cross-moved for summary judgment on Defendants’ fair use defenses as well as three elements of their DMCA claims. Dkt. 1696. As reflected in News Plaintiffs’ brief and accompanying Statement of Undisputed Facts, ECF Nos. 1728, 1985-1, News Plaintiffs do not believe a triable issue of fact prevents summary judgment in News Plaintiffs’ favor on these issues. The following facts further preclude summary judgment in Defendants’ favor.

I. Defendants’ Fair Use Defenses.

A. News Plaintiffs’ Expressive Content

1. News Plaintiffs’ journalism does more than simply communicate facts – it is a creative endeavor that involves judgment, expertise, and the ability to express those facts in a compelling way. News Plaintiffs’ journalists and editors make careful choices about what to present, how to put it in context, the best way to tell a story, and – in the case of opinion – what viewpoints and analysis will help readers better understand the world. Bell Rebuttal Rep. ¶ 202 (“[T]o limit journalism, and the works of a journalistic organisations like The Times, to the collection and relaying of bare facts is misrepresentative of the creativity and originality many news organisations, including The Times, bring to their journalism.”); NP Ex. 1364

¹ “*Daily News Plaintiffs*” or “DNP” includes Daily News, LP (“New York Daily News”); Chicago Tribune Company, LLC (“Chicago Tribune”); Orlando Sentinel Communications Company, LLC (“Orlando Sentinel”); Sun-Sentinel Company, LLC (“Sun-Sentinel”); San Jose Mercury-News, LLC (“San Jose Mercury News”); DP Media Network, LLC (“Denver Post”); ORB Publishing, LLC (“Orange County Register”); and Northwest Publications, LLC (“Pioneer Press”).

[NYTimes.com Homepage, <https://www.nytimes.com/>] (providing a menu of links to various categories of Times coverage, including categories like investigative reporting, breaking news reporting, beat reporting, reviews and analysis of arts and culture (including food, books, art, film, theater, television, music, fashion, and travel), technology, world news, U.S. news, commentary and opinion, Cooking, The Athletic, and Games); Owen Rebuttal Rep. ¶ 52 (“A news article is not only the facts it contains. It is also the reporting that produced those facts: the work of finding the story, developing sources, gathering documents, verifying what is true, telling the story (e.g., structural writing choices, original phrasing), and exercising editorial judgment about what to publish.”); NP Ex. 1217 [Sinnreich Opening Rep.] ¶ 120 ([REDACTED]); NP Ex. 1349 [Sinnreich Rebuttal Rep.] ¶ 78 ([REDACTED]); NP Ex. 1348 [Sinnreich Depo Tr.] at 80:10-19 ([REDACTED]); Cross Rebuttal Rep. at 11 (“News articles published by organizations like CIR and [The Intercept] are not merely the republication of facts. They provide information to readers about what is happening in the world with context and analysis.”); Wu Rebuttal Rep. ¶¶ 23, 26 (“Ziff Davis’s content includes news and current events, but with a focus. Consumers may also find Ziff Davis properties for the first time when they need specialized information, such as tech review or health advice.”).

2. News Plaintiffs' works include factual information and ideas, but their journalism reflects the authors' original expressions of those facts and ideas. Bell Rebuttal Rep. ¶ 28; Cross Rebuttal Rep. at 12; Wu Rebuttal Rep. ¶ 93.

3. News Plaintiffs' content is expressive. Bell Rebuttal Rep. ¶ 28 ("The Times's content is expressive and creative because it requires the reporters and editors to understand facts, data, emotions, context, background, history, current events, and so many other underlying factors in order to create trustworthy and compelling narratives that both inform the reader and express the information being conveyed in a way that is easy to digest and resonates with the reader."); Maisel Decl., Ex. A (Rule 1006 summary containing The Times's Asserted Works) and Ex. B (Rule 1006 summary containing Daily News Plaintiffs' Asserted Works); Cross Rebuttal Rep. at 5 ("CIR and TI offer readers information about what is happening in the world, conveyed in an easy-to-understand manner that provides the necessary context to understand that information in a meaningful way."); Wu Rebuttal Rep. ¶ 93 ("[P]ublishers like Ziff Davis make their content compelling [] by hiring professional writers to gather information and contextualize that information in narrative form in articles that keep a reader's attention and make the factual content accessible. That content...is the product of sustained investment, specialized expertise, and editorial craft, which amounts to creativity embedded in the expression itself."); Koonce Decl. (ECF 1859), Exs. AA–TT (Ziff Davis's Asserted Works); Bauerlein Decl. (ECF 1788) ¶¶ 14-19, 24-31; Bauerlein Decl. Ex. A at "Registered Works," "Mother Jones," and "Reveal" tabs; Bauerlein Decl. Exs. B-G; Bauerlein Decl. Ex. H at rows 48,920, 50,403, 83,760; Bauerlein Decl. Ex. I; Bauerlein Decl. Ex. J at -638; Bauerlein Decl. Ex. K at -734; Robins Decl. (ECF 1845) ¶¶ 4-6.

4. Reporters may devote decades to building trust with sources and developing specialized skills to surface and contextualize information. Cross Rebuttal Rep. at 7.

5. Editors exercise judgment to help make News Plaintiffs’ journalistic works clearer and more compelling, while also reviewing it for fairness, accuracy, and style. NP Ex. 3 [Sulzberger 30(b)(1) Depo. Tr.] at 167:7-14 (describing how all Times journalism “has to go through the hands of editors” “who are helping make things clear, more compelling” as well as “reading it for things like fairness, accuracy, any style considerations”); NP Ex. 4 [Dolnick 30(b)(1) Depo. Tr.] at 21:8-12 (“New York Times’ journalism requires human reporters who go out and bear witness, and they find things out. We trust our reporters. Our reporters have very specific experience. We trust our editors, who have judgment.”); NP Ex. 311 [Chabel Depo. Tr.] at 101:5-9 (editors are “responsible for catching anything that’s not properly attributed”); Cross Rebuttal Rep. at 12 (“Journalists and editors draw connections, explain context, uncover and weave in stories that translate information in a digestible way for consumers.”); NP Ex. 1305 [C. Mattheis 30(b)(6) Depo. Tr.] at 35:2–8 (describing how Ziff Davis’s articles published on *Babycenter* and *What to Expect* go through two layers of editorial review prior to publication).

6. The Times publishes journalism covering categories including investigative reporting, beat reporting, reviews and analysis of arts and culture (i.e., food, books, art, film, theater, television, music, fashion, and travel), its product recommendations website Wirecutter, its NYT Cooking website, sports reporting via The Athletic, and commentary and opinions. Bell Rebuttal Rep. ¶ 205; Maisel Decl., Exs. A, C, E-I.4.

7. Ziff Davis publishes game-guides, walkthroughs, and reviews through its Gaming and Entertainment vertical, high-quality consumer and professional medical content through its Health and Wellness vertical, and is a long-standing leader in providing technology news and product reviews through its Technology and Shopping vertical. Wu Rebuttal Rep. ¶ 23; Koonce Decl. (ECF 1859), Exs. AA–TT (Ziff Davis’s Asserted Works).

8. CIR publishes written short and long-form articles, several podcasts, short form video and long-form documentary investigations and a bi-monthly magazine. NP Ex. 1286 [Cross Depo. Tr.] at 78:17-79:3, 79:24-80:4 (written long-form articles & investigations); NP Ex. 1276 [Bauerlein 30(b)(6) Depo. Tr., Jan. 26, 2026] at 53:1-3 (testifying CIR competes in a “market for podcasts”), 42:14-19 (Mother Jones magazine has been published bimonthly since 1976 and is currently published bimonthly).

9. CIR’s coverage spans original investigative reporting, political analysis, and in-depth coverage of ongoing political, social and economic issues that unfold over time. Bauerlein Decl. ¶¶ 3-5. Many of CIR’s stories take months or years to report and require significant resources. Cross Rebuttal Rep. at 12, 31, 47.

10. CIR has collected many of journalism’s most prestigious awards, including among many other awards being a finalist for the Pulitzer Prize six times. Second Bauerlein Decl. ¶¶ 6-11.

11. The Intercept is known for scoops, editorial independence and investigations. Cross Rebuttal Rep. at 7.

12. Stories in The Intercept are often only made possible through years of reporting. One of those stories, “The Drone Papers,” which won the National Online Journalism Award in recognition that it was a story “a few, if any, other organizations could report,” was a tip-based story from a whistleblower source that exposed a drone-based assassination program that started under the Obama administration. Cross Rebuttal Rep. at 13; NP Ex. 1286 [Cross Depo. Tr.] 83:24-84:14.

13. Executive Editor Andrew Julien described reporting at the New York Daily News as follows: “We are on the ground. We have reporters out in the streets. We speak to members of the community and of the community in New York about the topics they care about, whether it’s

crime, the way that City Hall is, you know, running the city, what cultural events might be available to them. And we tell the stories that matter to our readers, and they rely on us to tell those stories.”

NP Ex. 22 [A. Julien 30(b)(6) Depo. Tr.] at 103:13-20.

14. Mitch Pugh, Executive Editor of the Chicago Tribune, described the coverage of news in the Chicago Tribune as follows: “We’ve reported on some of the largest stories in the history of not just Illinois and the Midwest but of the country. The paper broke the news of the Treaty of Versailles. It’s been – it was involved in promoting the abolition movement early on in its days. It was the first paper to report on the Watergate tapes and published the entire transcripts of the Watergate tapes. It’s well-known for its investigative and hard-hitting news coverage.” NP Ex. 21 [Pugh 30(b)(6) Depo. Tr.] at 59:4-13.

15. INTENTIONALLY LEFT BLANK

16. The Times’ Asserted Works in this matter include a wide range of content, including critics’ reviews and commentary, Pulitzer Prize winning explorations of the most pressing social issues, and personal essays. Examples include:

- A critic’s review of a documentary about writer Toni Morrison, Maisel Decl., Ex. I.2 at row 6403;
- An essay about the author’s relationship with pizza. Maisel Decl., Ex. I.2 at row 68; and
- Pulitzer Prize winning articles discussing race in America. Maisel Decl., Ex. I.1 at row 53590; Ex. I.2 at rows 15733 and 19221.

17. The expression embodied in news content is useful for LLM training, including because training on news gives LLMs the ability to generate news-like outputs. Etzioni Opening Rep. ¶ 53; *see also id.* ¶¶ 124-128; Etzioni Rebuttal Rep. ¶ 35 (“A model trained on a news article

is therefore optimized to reproduce the article’s specific word choices, sentence structures, and rhetorical patterns—*i.e.*, the expressive features of the article, not only the underlying facts contained therein.”).

B. Defendants’ Exploitation of News Plaintiffs’ Expressive Content

1. Use of News Plaintiffs’ Content for Training

18. OpenAI copied News Plaintiffs’ works in their entirety in training the at-issue LLMs. Goldstein Rebuttal Rep. ¶ 97; Etzioni Rebuttal Rep. ¶ 71; NP Ex. 1163 [Wenger Supplemental Rep. Appendix 194]; NP Ex. 1164 [Wenger Supplemental Rep. Appendix 195]; NP Ex. 1165 [Wenger Supplemental Rep. Appendix 196]; NP Ex. 1166 [Wenger Supplemental Rep. Appendix 197]; Delorey Opening Rep. ¶¶ 189–191; Delorey Supplemental Rep. ¶¶ 9-14; Delorey Reply Rep. ¶ 74.

19. In training the at-issue LLMs OpenAI made multiple wholesale copies of News Plaintiffs’ works. Goldstein Rebuttal Rep. ¶ 97; NP Ex. 1163 [Wenger Supplemental Rep. Appendix 194]; NP Ex. 1164 [Wenger Supplemental Rep. Appendix 195]; NP Ex. 1165 [Wenger Supplemental Rep. Appendix 196]; NP Ex. 1166 [Wenger Supplemental Rep. Appendix 197]; Delorey Opening Rep. ¶ 195; Delorey Supplemental Rep. ¶ 14 ([REDACTED]).

20. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

21. [REDACTED]

22. OpenAI prioritized high-quality content in training its LLMs, including high-quality news content. NP Ex. 496 [Dyett Depo. Tr.] at 95:6-11 ([REDACTED]

[REDACTED]).

23. OpenAI's technical expert [REDACTED]
[REDACTED]" NP Ex. 675 [Callison-Burch Reply Rep.] ¶ 127.

24. Training LLMs on news content enables the LLMs to copy or mimic the expressive features of news content and to produce news-like output. Etzioni Opening Rep. ¶ 99 ("if you want to provide news-like answers to users, news-like content is required during training"), ¶¶ 123-128; Etzioni Rebuttal Rep. ¶ 70 ("The capability to produce news-like outputs comes from training on the expressive features of journalism."); Etzioni Rebuttal Rep. ¶ 67 (a model trained on News Plaintiffs' works "learns the *form* of journalism, not merely the underlying facts"); Mitchell Opening Rep. ¶¶ 46, 110-112; Wenger Rebuttal Rep. at 16.

25. One purpose for which OpenAI trained on News Plaintiffs' Asserted Works was to give OpenAI's LLMs the ability to produce text that conveys news and other content to readers. Etzioni Rebuttal Rep. ¶ 52 ("The original purpose of the asserted works was to convey news to readers. The purpose for which OpenAI trained on those works was to produce text that conveys news to readers. The two purposes are the same."); Wu Rebuttal Rep. ¶ 93 ("Generative AI products like ChatGPT are designed to replicate that aspect of human journalism, by generating

narrative, accessible text. However, it does so by being extensively trained on and by retrieving human-created narrative content, allowing it to effectively replicate this content.”).

26. OpenAI trained its LLMs on news to capture and replicate its expressive elements. Etzioni Rebuttal Rep. ¶¶ 35 (“A model trained on a news article is therefore optimized to reproduce the article’s specific word choices, sentence structure, and rhetorical patterns—*i.e.*, the expressive features of the article, not only the underlying facts contained therein”), ¶¶ 68-69 (“The model’s ability to produce news-like expression is derived from the expressive content of its training data.”), ¶ 70 (“LLMs are trained to minimize prediction error on the next token of the source text. That prediction process explicitly considers the composition of the text, including the word choices and ordering of those words within the text itself... The capability to produce news-like outputs comes from training on the expressive features of journalism.”); Etzioni Opening Rep. ¶ 123; Mitchell Opening Rep. ¶ 71 (“Specific structures encoded in the LLM include the precise expression used in those different texts (e.g., emails, news articles, encyclopedia entries, film scripts, etc.), the expressive styles of different writers and publications, and other features of text, as well as general world knowledge....”); Wu Rebuttal Rep. ¶ 93.

27. A model trained on professionally written and edited news articles learns to generate text that exhibits the hallmarks of journalism. Etzioni Opening Rep. ¶ 123.

28. News Plaintiffs’ copyrighted works are economically valuable, expressive information goods that Defendants specifically sought out for their quality, creativity, uniqueness, and the informational value that comes from such expression, which in large part is driven by the fixed investment required to develop such expression. Malackowski Rebuttal Rep. at 6-7; Malackowski ZD Rebuttal Rep. at 6.

29. From an economic perspective, the nature of News Plaintiffs’ works is highly distinguishable from “joe schmo’s” blog. Malackowski Rebuttal Rep. at 7; Malackowski ZD Rebuttal Rep. at 6.

30. Because the form, style, and expression in LLM output mirrors the text that the LLM is trained on, news content is essential to provide satisfying answers to news-related user queries, or other prompts that should solicit a news-like response. Etzioni Opening Rep. ¶ 99 (“In plain English, if you want to provide news-like answers to users, news-like content is required during training.”); Mitchell Rebuttal Rep. ¶ 41 (“[T]here is one LLM application that *does* require training on [copyrighted news articles], namely, generating news articles that mimic the expressive style and content of those in the training data.”); Mitchell Reply Report. ¶ 43 (“[T]raining OpenAI’s models on Plaintiffs’ asserted works and other high-quality news articles was strictly necessary for only one capability—generating the equivalent of news articles that mimic the expressive style and content of those in the training data.”). OpenAI’s prioritization of news content enabled its models to become good at generating news content. Etzioni Opening Rep. ¶ 115.

31. LLMs do not “distill facts”—they at most simulate the appearance of knowledge by approximating the expression of facts that training data contains. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

32. Pre-training and post-training on news content enables the model to more accurately parse and synthesize news passages retrieved via RAG. Etzioni Opening Rep. ¶ 164; *see also* Etzioni Reply Rep. ¶ 16 (“[N]ews content is essential for training LLMs that will be used to query news sources and output news information.”)

33. OpenAI’s prioritization of news content during training enabled its models to become good at generating news content. Etzioni Opening Rep. ¶ 115.

34. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

35. OpenAI’s Terms of Use state, “YOU WILL NOT RELY ON OUTPUT AS A SOLE SOURCE OF TRUTH OR FACTUAL INFORMATION.” NP. Ex. 1363 [OpenAI Terms of Use, <https://openai.com/policies/row-terms-of-use/>].

36. In a defamation case, OpenAI argued, as a matter of law, that it disclaimed the “accuracy or truthfulness” of ChatGPT responses by providing “repeated warnings to [the plaintiff] and all users that ChatGPT can make mistakes” that were “specific, clear, and prominently displayed on-screen throughout every interaction with ChatGPT.” NP Ex. 1287 [Defendant OpenAI, L.L.C.’s Reply Brief in Support of Motion for Summary Judgment, No. 23-A-04860-2 (Ga. Sup. Ct. Apr. 7, 2025)] at 4.

37. In a defamation case, OpenAI argued that, as a matter of law, through its terms of use, it made clear that ChatGPT’s statements “cannot be taken at face value” as to their factual accuracy. NP Ex. 1287 [Defendant OpenAI, L.L.C.’s Reply Brief in Support of Motion for Summary Judgment, No. 23-A-04860-2 (Ga. Sup. Ct. Apr. 7, 2025)] at 5.

38. The training objective of LLMs is to generate text that sounds plausible based on the statistics learned from the training data—not simply to extract the underlying facts. Mitchell Opening Rep. ¶ 163 (The Times and Daily News Plaintiffs’ expert Dr. Mitchell opining that “the training objective of LLMs” is “to generate text that sounds plausible, based on the statistics learned from the training data,” and that “LLMs learn only from language and do not have direct knowledge of truth or falsity of the text they generate”), ¶ 71.

39. The core function of an LLM is to generate text that is similar to the corpus of text it was trained on. Mitchell Opening Rep. ¶ 26.

40. The purpose of the pre-training stage is to tune the LLM’s parameters to encode the general structure of language, as well as the specific structure of different types of text, including the expressive styles of different writers and publications. Mitchell Opening Rep. ¶ 71 (“The purpose of the pre-training stage is to tune the LLM’s parameters to encode the general structure of language, as well as the specific structure of different types of text. Specific structures encoded in the LLM include the precise expression used in those different texts (e.g., emails, news articles, encyclopedia entries, film scripts, etc.), the expressive styles of different writers and publications, and other features of text, as well as general world knowledge (e.g., ‘apples are a kind of fruit’ or ‘winter, spring, summer and fall are the four seasons’).”).

41. During mid-training LLMs are further trained on high-quality and domain specific text with the purpose of enhancing the model. Etzioni Opening Rep. ¶ 31; [REDACTED]

[REDACTED]

[REDACTED]

42. [REDACTED]

[REDACTED]

[REDACTED] NP Ex. 757 [OPCO_MDL_000155738] at slide 6; NP Ex. 755 [OPCO_MDL_004585389] at slide 8 ([REDACTED]); NP Ex. 759 [OPCO_MDL_004740422] at -422 ([REDACTED]).

43. The purpose of post-training is to teach the model to exhibit certain behaviors and perform particular tasks. NP Ex. 461 [Berg-Kirkpatrick Opening Rep.] ¶ 34. This includes teaching the model to use tools at inference time like OpenAI’s “browser” tool which enables the model to search for and retrieve content from the internet. Etzioni Opening Rep. ¶ 157.

2. *Use of News Plaintiffs’ Content for Retrieval Augmented Generation*

44. Retrieval augmented generation (“RAG”) entails wholesale copying of News Plaintiffs’ works. Goldstein Opening Rep. ¶ 213 (“[I]n the case of news articles, often the entire news article is included in the semantic document is passed to the LLM.”); *id.* (describing the “semantic document” as the content of a website provided to the LLM for RAG).

45. Through RAG, Defendants copied News Plaintiffs’ journalism to reuse it in a new context: retrieved and presented through a chatbot instead of on the publisher’s site. Etzioni Opening Rep. ¶ 179.

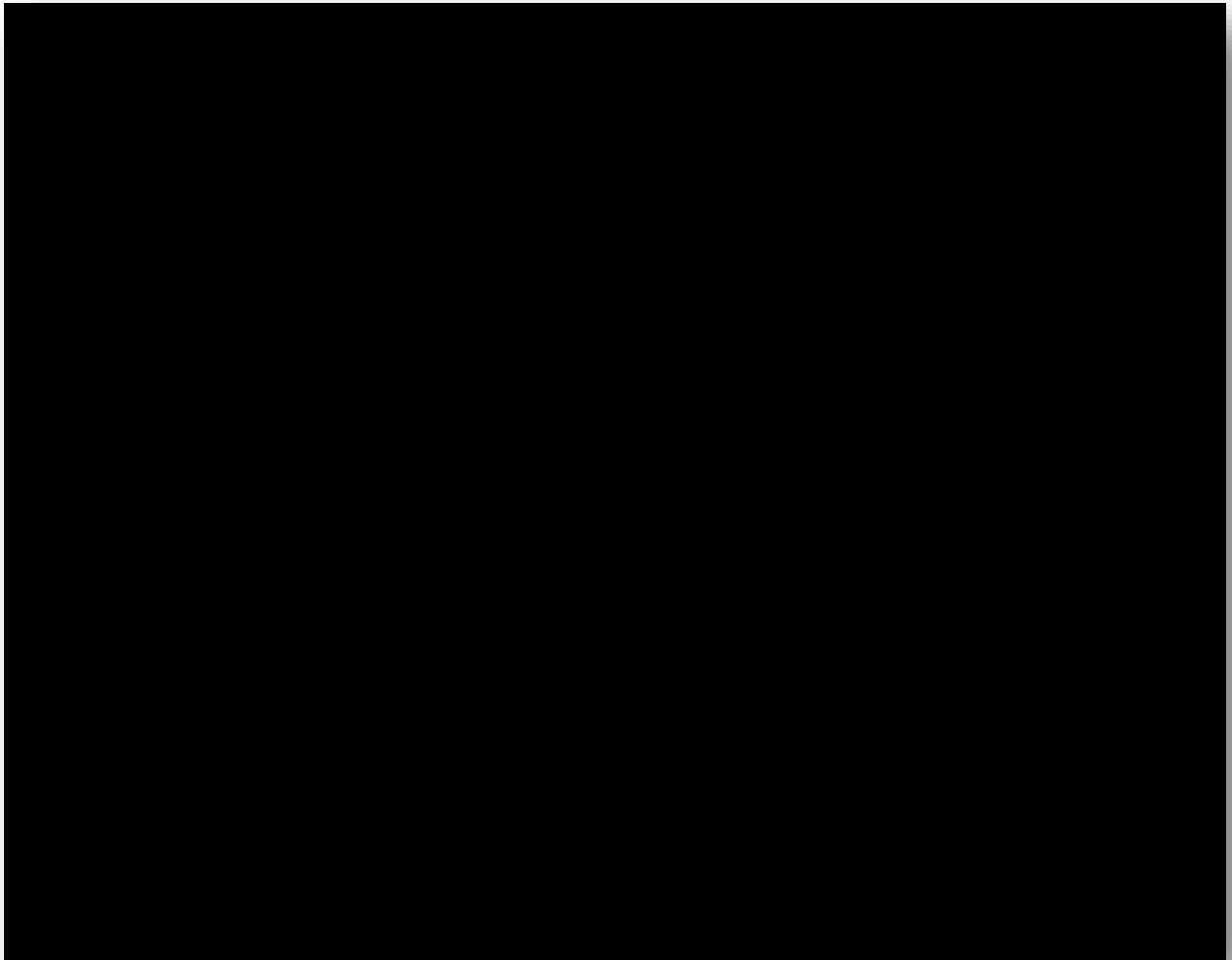
46. Chatbots’ ability to synthesize content retrieved via RAG into a journalistically competent response comes from training on news content. Etzioni Rebuttal Rep. ¶ 96 (“RAG depends on the model’s ability to synthesize retrieved content into a journalistically competent response - to put the retrieved facts into the expected stylistic form, to use the source-attribution conventions of journalism, to write in the editorial register that makes the response credible as news. That synthesis capability comes from training on news. RAG is therefore not an alternative

to training on news; it is an amplifier of the substitutive use that training on news enables. Microsoft's products combine both mechanisms: a model trained on the asserted works, and a retrieval layer that copies news content at the moment of inference.'').

47. Defendants prioritized their chatbots' access to high-quality news content for RAG. NP Ex. 496 [Dyett Depo. Tr.] at 46:7-15 ([REDACTED]); NP Ex. 681 [Dahlin Depo. Tr.] at 52:20-53:2 ([REDACTED]).

3. *Use of News Plaintiffs' Content for Display in Chatbot Outputs*

48. Copilot reproduces substantial amounts of News Plaintiffs' content in its outputs.



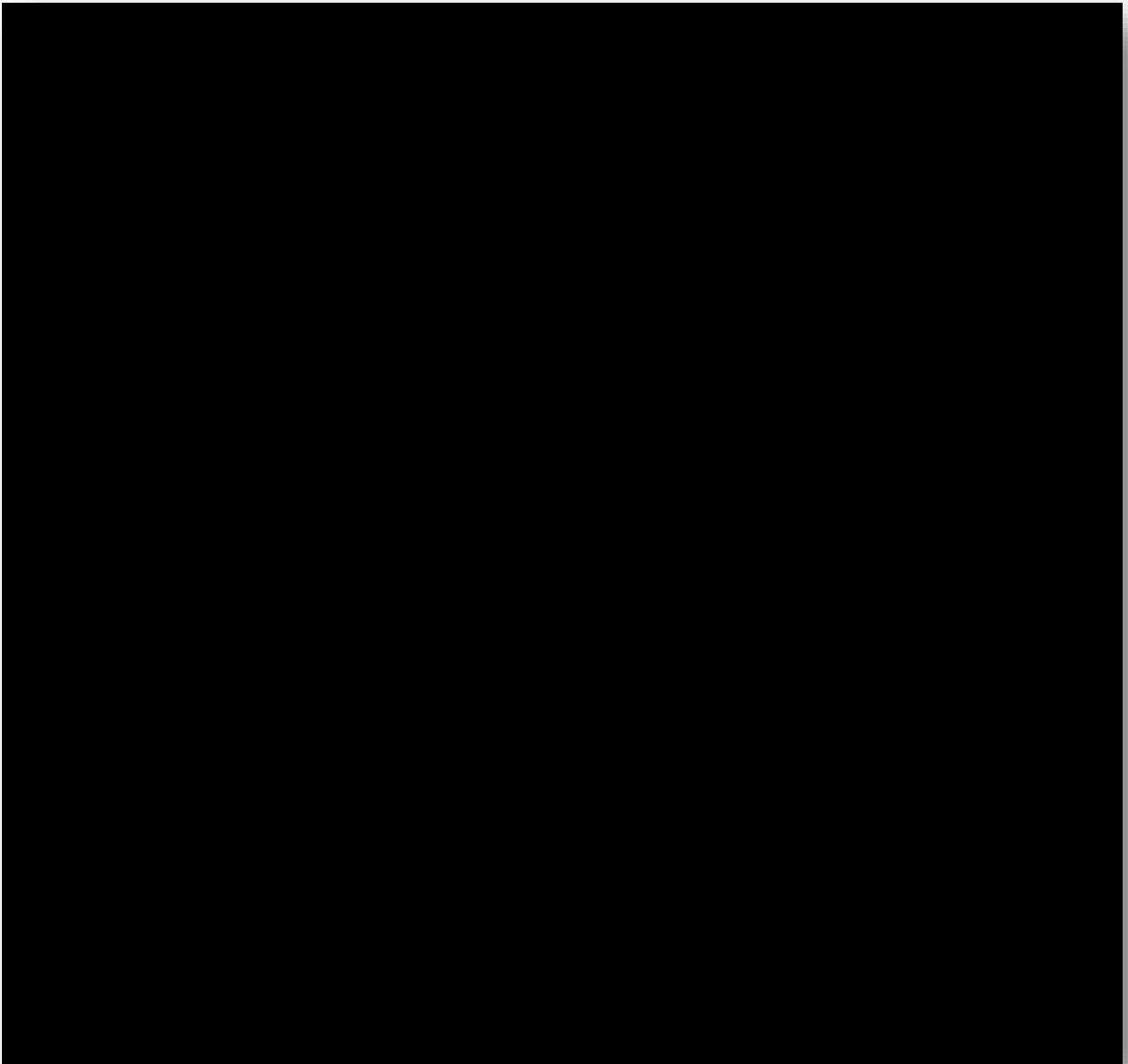
Supplemental Goldstein Decl., Exs. V and W (NYT Output Examples) at 1-2.

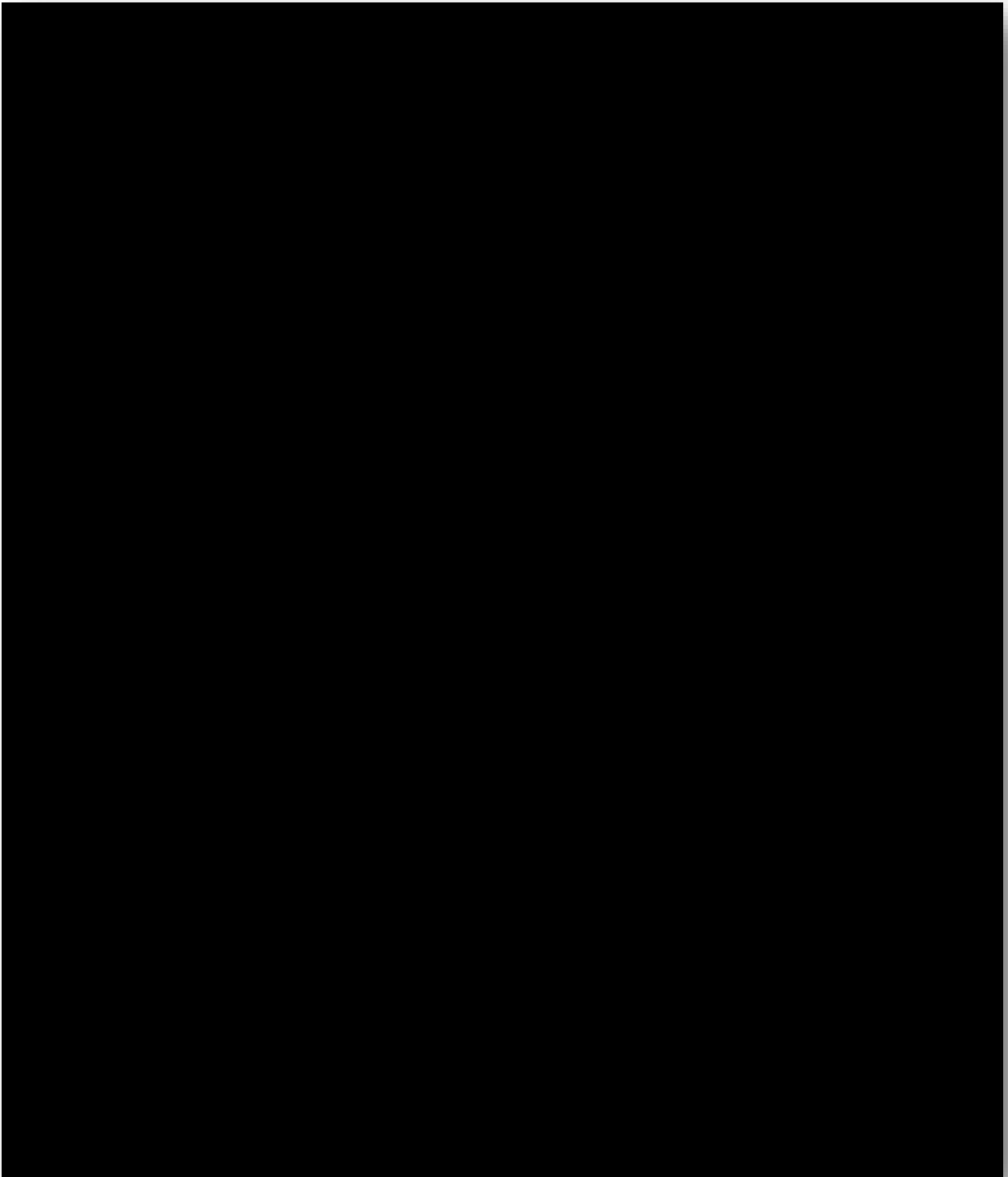
49.

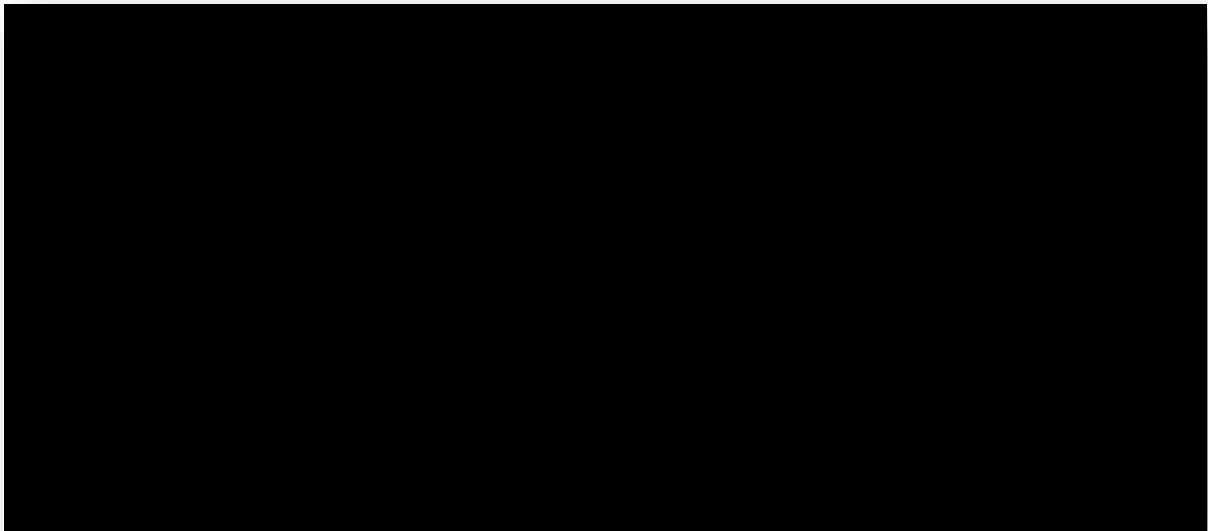
[REDACTED]

[REDACTED]

[REDACTED]:





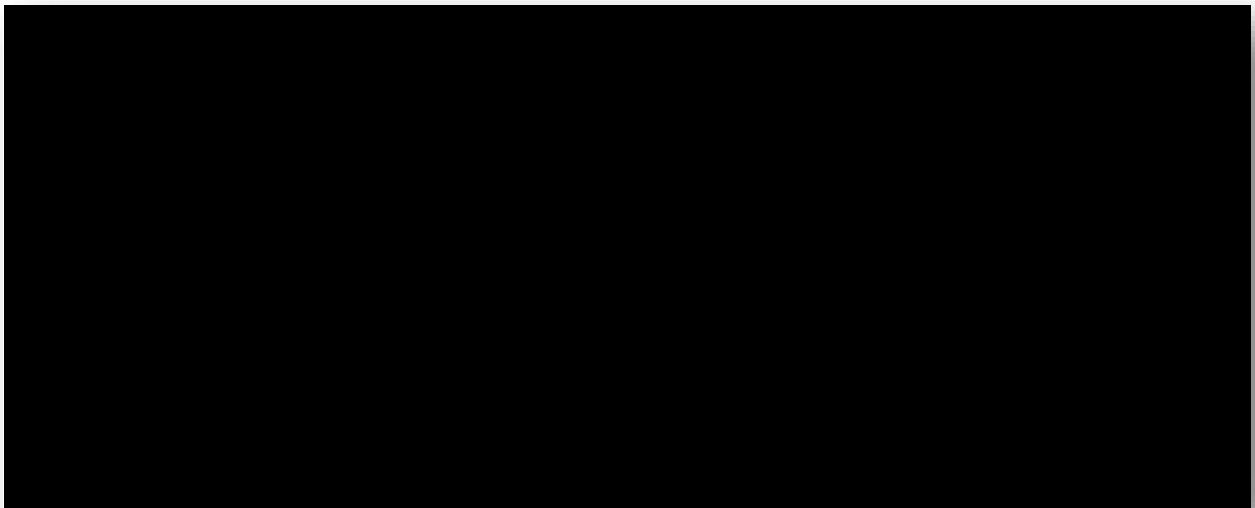


Supplemental Goldstein Decl., Ex. X (CT - Output Examples) at 1-4.

50.



:



Supplemental Goldstein Decl., Ex. Y (DP - Output Examples) at 2.

51. [REDACTED]

[REDACTED]
[REDACTED] Wenger Rebuttal Rep. at 55; NP Ex. 410 (Feamster Opening Rep.) ¶ 327 ([REDACTED]
[REDACTED]).

C. Defendants' Copyright and Antiregurgitation Measures

52.

OpenAI

[REDACTED]. NP Ex. 633 [Paino 30(b)(6) Depo. Tr.] at 132:8-9 [REDACTED]
[REDACTED]"); *id.* at 131:4-9 ("[REDACTED]
[REDACTED]").

53.

Dr. Goldstein reviewed

[REDACTED]
[REDACTED].
Goldstein Rebuttal Rep. ¶ 100. Similarly, Mr. Delorey [REDACTED]
[REDACTED]
[REDACTED]. October 23, 2026 Delorey Decl. ¶¶ 5-7.

54.

[REDACTED].

Table 4: Asserted Works Appearing Multiple Times in Training Data

Model	Works Appearing 2+ Times	Works Appearing 10+ Times
gpt-3.5-turbo	2,262,567	123,927
gpt-4-turbo	2,152,772	11,215
gpt-4	1,949,802	56,883
gpt-3.5	1,108,153	217
gpt-4o	806,255	3,259
gpt-3	746,608	24,653
gpt-2	3,724	31

Asserted Works Appearing Multiple Times in Training Data

Model	Works Appearing 2+ Times	Works Appearing 10+ Times
gpt-3.5-turbo	383,022	5,747
gpt-4-turbo	186,387	930
gpt-4	436,730	22,363
gpt-3.5	6,472	4
gpt-4o	52,880	70
gpt-3	263,309	1,939
gpt-2	265	0

October 23, 2026 Delorey Decl. ¶ 5.

55. Dr. Goldstein found [REDACTED]

[REDACTED]. Goldstein Rebuttal Rep. ¶ 101.

56. Dr. Goldstein found [REDACTED]

[REDACTED]. Goldstein Rebuttal Rep. ¶ 101; Goldstein Rebuttal Rep., Exhibit 2.

57. OpenAI acknowledged [REDACTED]

[REDACTED]. NP Ex. 633 [Paino 30(b)(6) Depo. Tr.] at 124:23-125:4.

58. A model can still memorize and regurgitate text only appearing once in the training data. NP Ex. 635 [Callison-Burch Depo. Tr.] at 268:22-24

59. Memorization and overfitting are different concepts. NP Ex. 633 [Paino 30(b)(6) Depo. Tr.] at 310:17-18; NP Ex. 456 [Radford Depo. Tr.] at 346:19-24.

60. [REDACTED]

[REDACTED]. NP Ex. 1194 [Bird 30(b)(6) Depo. Tr.] at 168:2-13.

61. Microsoft pressed that [REDACTED]

[REDACTED]

[REDACTED] NP Ex. 1308 [MSFT_AI_MDL_000640527] at -527.

62. Microsoft recognized that [REDACTED]

[REDACTED] NP Ex. 1311 [MSFT_AI_MDL_001379616] at -618.

63. [REDACTED]

[REDACTED] NP Ex. 800 [Lafferty Rebuttal Rep.] ¶ 55.

64. [REDACTED]

[REDACTED]

[REDACTED] NP Ex. 1194 [Bird 30(b)(6) Depo. Tr.] at 163:14-164:17.

65. [REDACTED] NP Ex. 657 [Abib 30(b)(6) Depo. Tr.] at 100:19-22.

66. Microsoft's corporate designee, Elbio Abib, [REDACTED]

[REDACTED] NP Ex. 657 [Abib 30(b)(6) Deposition Tr.] at 103:11-20, 107:9-20.

67. Microsoft's corporate designee, Elbio Abib, [REDACTED]
[REDACTED]. NP Ex. 657 [Abib 30(b)(6) Depo. Tr.] at 106:18-107:24.

68. Microsoft considered [REDACTED]
[REDACTED] Bird Decl. ¶¶ 41-42; NP Ex. 1310 [MSFT_AI_MDL_001178782] at 9-10.

II. Defendants' Implied License Defense.

69. When OpenAI announced its ChatGPT-User bot in March 2023, [REDACTED]
[REDACTED] NP Ex. 613 [Gaggara 1/27/26 30(b)(6) Depo] at 287:20-21; *see* NP Ex. 1307 [MSFT_AI_MDL_000170197] at -199 ([REDACTED]
[REDACTED]).

70. In October 2023, [REDACTED]
[REDACTED]
[REDACTED]
NP Ex. 1290 [Gaggara 1/28/26 30(b)(6) Depo. Tr.] at 440:22-441:1.

71. On June 30, 2023, Microsoft [REDACTED]
[REDACTED]
[REDACTED] NP Ex. 1307 [MSFT_AI_MDL_000170197] at -200.

72. [REDACTED]
[REDACTED]
[REDACTED] NP Ex. 235 [MSFT_AI_MDL_000229826]; NP Ex. 445 [NYT_00815634] ([REDACTED]
[REDACTED]).

[REDACTED] NP Ex. 235

[MSFT_AI_MDL_000229826] at -827; NP Ex. 1315 [NYT_00467690] at -691.

73. OpenAI's own President, Greg Brockman, [REDACTED]

[REDACTED] NP Ex. 1333

[OPCO_MDL_004762434] at -35.

74. Website owners can either "allow" or "disallow" specific bots in their robots.txt files to indicate which resources should be accessed by that bot. NP Ex. 410 [Feamster Opening Rep.] ¶ 127.

75. The Times uses software, DataDome, to implement "hard blocks" in some instances, but it costs money to do so. Cost is a factor considered by The Times when deciding whether to implement a hard block, as DataDome is expensive to implement and its cost increases the more The Times uses it. NP Ex. 1350 [Sobel 30(b)(1) Deposition, 12/9/2025] at 70:10-71:4 (testifying that The Times considers cost when deciding whether to implement CDN blocks, which it does via DataDome, as "[t]he more you use DataDome, the more you pay"); NP Ex. 309 [Procas 30(b)(6) Deposition, 3/19/26] at 236:8-23 ("DataDome is very expensive"); NP Ex. 1314 [NYT_00269020] at -021 ("We have considered adding Datadome to the homepage route, but are hesitant to do so due to the increased costs we would have to pay for Datadome's services due to the amount of traffic that the homepage receives.").

76. OpenAI's bot used for RAG, ChatGPT-User, may not respect robots.txt directives. NP Ex. 1343 [<https://developers.openai.com/api/docs/bots/>] (claiming with respect to ChatGPT-User's actions that, "Because these actions are initiated by a user, robots.txt rules may not apply").

77. The Intercept first learned of OpenAI’s web crawlers, GPTBot and ChatGPT-User (“OpenAI Web Crawlers”), in or around January 2024. Upon learning of the OpenAI Web Crawlers, The Intercept promptly placed robots.txt code on its website, theintercept.com, that requested that the OpenAI Web Crawlers not crawl its website. The Intercept implemented this robots.txt code on January 18, 2024. Sherrard Decl. ¶ 4.

78. If a publisher wants to prevent their website content from being grounded on by Bing Chat, using the “NOARCHIVE” tag is not sufficient. [REDACTED]

[REDACTED] ECF 2057 (Gaggara 702 motion Declaration) ¶ 4; *see also* ECF 1704 (Gaggara MSJ Declaration) ¶ 26; NP Ex. 453 [MSFT_AI_MDL_000391970] at -970 ([REDACTED])

[REDACTED]; NP Ex. 454 [Abib 30(b)(1) Tr.] at 61:8-62:10 ([REDACTED])

[REDACTED]; NP Ex. 455 [MSFT_AI_MDL_000642967] at -967.

79. In order to request that Bing re-index their content, a publisher must use Microsoft’s “IndexNow” tool. However, according to Microsoft, even using the “IndexNow” tool does not guarantee reindexing. NP Ex. 1354 [https://www.bing.com/webmasters/help/URL-Submission-62f2860b] at 9 (“Bing may choose not to index specific URLs if they do not meet its selection criteria.”).

80. Microsoft’s blog post announcing its use of the NOARCHIVE tag did not say that the NOARCHIVE tag is insufficient on its own to preclude use of a publisher’s content for grounding by Bing Chat. Rather, it indicated that use of that tag was sufficient. NP Ex. 451 [Sept.

22, 2023 Microsoft Blog Post “Announcing New Options for Webmasters to Control Usage of their Content in Bing Chat”, <https://blogs.bing.com/webmaster/september-2023/Announcing-new-options-for-webmasters-to-control-usage-of-their-content-in-Bing-Chat>] (telling publishers that “Content tagged NOARCHIVE will not be included in Bing Chat answers, not be linked to in the answers. Going forward, for content in our Bing Index that is labeled NOARCHIVE, we will not use the content for training Microsoft’s generative AI foundation models”—without any mention of the need to request reindexing or use “IndexNow”).

81. Similarly, the documentation for Microsoft’s IndexNow tool says nothing about any potential benefit to website owners desiring to use the NOARCHIVE tag for exclusion of their content from grounding. NP Ex. 1296 [<https://www.bing.com/webmasters/help/indexnow-0z209wby>].

82. Companies employing bots typically interpret the “NOARCHIVE” tag as communicating that they may not access cached versions of content. Stricchiola NYT/DNP Opening Rep. ¶ 99; *see also* NP Ex. 449 [NYT_00849355] at -56; NP Ex. 450 [NYT_00480186] at -86.

83. Adopting the Robots Exclusion Protocol (“REP”) generally does not cover specific bots. The REP requires site owners to know which bots, specifically and by name, they want to manage. Wenger Opening Rep. at 12.

84. CIR testified that, notwithstanding awareness of robots.txt, CIR has “not implemented robots.txt protocols” to exclude any search engine crawlers because “we want our content to be discoverable and findable by audiences.” NP Ex. 1276 [Bauerlein 30(b)(6) Depo. Tr. Jan 26, 2026] 170:6-22; 171:7-9.

85. Prior to the initiation of this lawsuit, CIR was not aware of Microsoft webmaster control meta-tags, including the “NOARCHIVE” metatag, which relate to permissions for generative AI training. Second Bauerlein Decl. ¶ 5.

86. A robots.txt “allow” directive enabling a bot to access or download content (or the absence of a robots.txt disallow) does not override existing protections, such as a copyright notice or term of service, or the fact that copyright law does not require a copyright owner to “opt out” of having its copyright content copied and misused. Wenger Opening Rep. at 14; Wenger Reply Rep. at 106.

87. Crawler instructions, such as robots.txt disallow directives, are subordinate to Plaintiffs’ terms of use and the Copyright Act being an opt-in regime. Wenger Opening Rep. at 7.

88. The mechanisms that Defendants recommended News Plaintiffs should use for controlling how their content is used by Defendants’ products are novel (as repurposed by Defendants), evolving, and were not clearly communicated. NP Ex. 72 [Jacobs Feb. 12, 2026 30(b)(6) Depo. Tr.] at 94:22-95:12.

89. *Daily News* Plaintiffs had concerns that using metatags to control how their content was used for Generative AI might negatively impact their content’s likelihood of surfacing in search. NP Ex. 72 [Jacobs Feb. 12, 2026 30(b)(6) Depo. Tr.] at 108:3-24.

III. Plaintiffs’ DMCA Claims.

A. Intentional Removal of CMI

90. OpenAI created and used text extractors specifically designed to remove CMI from online articles (called “Newspaper3K,” “Dagnet” and “Gutentag”). Goldstein Opening Rep. ¶¶ 138, 524, 535-542.

91. OpenAI’s use of text extractors was a systematic approach to removal that specifically targeted CMI by design. Etzioni Reply Rep. ¶ 22 (“CMI was removed by design. A

preprocessing pipeline does not strip standardized copyright notices, terms-of-service text, datelines, and bylines across a corpus by accident; systematic, category-specific removal is the signature of code written to target those categories.”); *id.* ¶ 25 (“[REDACTED]”); *see also* Delorey Opening Rep. ¶¶ 232-252; *id.* ¶ 252 (“[REDACTED]”).

B. Concealment of Output Infringement

92. Including the Asserted Works’ CMI during training likely would have resulted in that CMI being included in outputs that regurgitated parts of the Asserted Works. Wenger Opening Rep. at 110-16; Wenger Reply Rep. at 81-87; Etzioni Opening Rep. ¶ 279 (“ChatGPT and other OpenAI LLMs would have likely revealed their sources had OpenAI retained the CMI.”); Goldstein Reply Rep. ¶¶ 171-172.

93. OpenAI’s removal of the Asserted Works’ CMI prior to training prevented that CMI from being regurgitated when OpenAI’s models regurgitate parts of the Asserted Works. Wenger Opening Rep. at 110-116; Wenger Reply Rep. at 81-87; Etzioni Opening Rep. ¶¶ 237, 263; *see also id.* ¶ 250 (“[REDACTED]”); Delorey Reply Rep. ¶ 48; NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 50:22-25.

94. OpenAI’s removal of the Asserted Works’ CMI likely conceals that a regurgitation of an Asserted Work is an infringing copy. Wenger Opening Rep. at 8, 110-116; Wenger Reply Rep. at 81-87; Tease Opening Rep. at 15-20; Etzioni Reply Rep. ¶ 28 (“CMI was present in the works being processed, retention would have resulted in reproduction of that CMI in outputs, and removal prevented that reproduction.”); Etzioni Opening Rep. ¶ 280 (“In the but-for world where OpenAI had retained CMI in its training data rather than stripping it during preprocessing, the

predictable consequence is that ChatGPT would emit that CMI when generating content derived from the underlying news articles.”).

95. At the time it removed News Plaintiffs’ CMI, OpenAI had reason to know that regurgitations of the Asserted Works constituted copyright infringement. Tease Opening Rep. at 15-20; NP Ex. 811 [OPCO_MDL_001699373] at -373; NP Ex. 1144 [OPCO_MDL_001657091] at -94; NP Ex. 840 [MSFT_AI_MDL_000504460] at -460-61.

96. OpenAI stated that its products can “use copyrighted materials ... in ways that would exceed fair use,” including “displaying long excerpts from a work in response to a user query.” NP Ex. 1299 [Letter to Suzanne V. Wilson, General Counsel and Associate Register of Copyrights, from Tom Rubin, Chief of Intellectual Property and Content at OpenAI (June 4, 2024), <https://www.copyright.gov/policy/artificial-intelligence/ex-parte-communications/letters/OpenAI-June-4-2024.pdf>] at 2-3.

97. OpenAI’s expert, Dr. Callison-Burch, admitted [REDACTED]

[REDACTED]. NP Ex. 676 [Callison-Burch Rebuttal Rep.] ¶ 95.

98. OpenAI’s expert, Dr. Callison-Burch, opined that [REDACTED]

[REDACTED]

NP Ex. 676 [Callison-Burch Rebuttal Rep.] ¶ 95.

99. OpenAI’s text extractors, if configured to retain the Asserted Works’ CMI, would have removed the characters originally separating the CMI and the Asserted Work, thus rendering the CMI contiguous with the Asserted Work. Wenger Reply Rep. at 82-83; Goldstein Opening Rep. ¶ 539.

100. [REDACTED] NP Ex. 631

[Frederiksen-Cross Depo. Tr.] at 50:24-25.

101. [REDACTED]

[REDACTED] NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 52:14-21; *see also* NP Ex. 456 [Radford Feb. 24, 2026 Depo. Tr.] at 314:23-315:23; NP Ex. 130 [Brockman Jan. 29, 2026 Depo. Tr.] at 355:11-357:23; NP Ex. 521 [Yuan Nov. 18, 2025 Depo. Tr.] at 85:4-6; NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 52:14-18; Etzioni Opening Rep. ¶¶ 240, 263, 279-281; Etzioni Reply Rep. ¶ 28; Mitchell Opening Rep. ¶ 123; Wenger Rebuttal Rep. at 10-11.

102. [REDACTED]

[REDACTED] NP Ex. 1171

[Frederiksen-Cross Rebuttal Rep.] ¶ 155.

103. [REDACTED].

Callison-Burch Decl. ¶ 109.

104. [REDACTED]

[REDACTED] Goldstein

Decl., Exhibits K and L; Maisel Decl., Exhibits C and D (CMI Removal Rule 1006 summaries); Wenger Supplemental Rep. at 29; NP Ex. 1172 [Wenger Supplemental Rep. Appendix 198]; NP Ex. 1173 [Wenger Supplemental Rep. Appendix 199]; NP Ex. 1174 [Wenger Supplemental Rep. Appendix 201]; Delorey Supplemental Rep. ¶¶ 16, 41.

105. [REDACTED]

[REDACTED] NP Ex. 1171 [Frederiksen-Cross Rebuttal Rep.] ¶¶ 150-161.

106. [REDACTED]

[REDACTED] NP Ex. 1171 [Frederiksen-Cross Rebuttal Rep.] ¶¶ 156-158.

107. News Plaintiffs' copyright notices and links to terms of use can be shorter than 24 or 32 tokens. Wenger Reply Rep. at 77-78; Goldstein Reply Rep. ¶¶ 176-182; Delorey Reply Rep. ¶ 45.

108. [REDACTED] NP Ex. 633 [Paino 30(b)(6) Dep. Tr.] at 31:3-8; Wenger Reply Rep. at 76; Goldstein Reply Rep. ¶¶ 176-182.

109. [REDACTED]

[REDACTED] Wenger Reply Rep. at 76-79; Goldstein Reply Rep. ¶¶ 176-182.

110. In reaching their opinions that OpenAI's removing CMI prevented the models from reproducing CMI, News Plaintiffs' experts conducted analyses of memorization and regurgitation with citations to record evidence, source code, and technical literature. Wenger Opening Rep. at 67-70, 110-118; Wenger Reply Rep. at 81-90; Etzioni Opening Rep. ¶¶ 279-284, 294-306; Etzioni Reply Rep. ¶¶ 33-45; Goldstein Reply Rep. ¶¶ 171-172.

111. [REDACTED]

[REDACTED] NP Ex. 676 [Callison-Burch Rebuttal Rep.] ¶¶ 92-95; NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 50:22-25, 52:11-18.

112. [REDACTED]

[REDACTED]. NP Ex. 676 [Callison-Burch Rebuttal Rep.] ¶¶ 92-95; NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 50:22-25, 52:11-18.

113. [REDACTED]

[REDACTED]. NP Ex. 676 [Callison-Burch Rebuttal Rep.] ¶¶ 92-95; NP Ex. 631 [Frederiksen-Cross Depo. Tr.] at 50:24-25, 52:11-18.

114. Conducting an experiment to test whether retaining CMI would have caused the CMI to be outputted would have required an immense amount of compute and time. NP Ex. 1356 [Wenger Depo. Tr.] at 170:10-15.

115. Dr. Berg-Kirkpatrick's [REDACTED]

[REDACTED] Wenger Rebuttal Rep. at 37-38; *see also* Delorey Rebuttal Rep. ¶¶ 16, 48-49 (calculating similar extrapolations).

116. Dr. Berg-Kirkpatrick's [REDACTED]

[REDACTED] Wenger Rebuttal Rep. at 37-38; *see also* Delorey Rebuttal Rep. ¶¶ 16, 48-49 (calculating similar extrapolations).

117. Dr. Berg-Kirkpatrick's [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Wenger Rebuttal Rep. at 37-38; Berg-Kirkpatrick Decl. ¶ 77: [REDACTED]; *see also* Delorey Rebuttal Rep. ¶¶ 16, 48-49 (calculating similar extrapolations).

118. [Intentionally omitted.]

119. [Intentionally omitted.]

120. [Intentionally omitted.]

121. [Intentionally omitted.]

C. OpenAI's Knowledge of Regurgitation and the Inefficacy of its Mitigation Measures

122. GPT-3.5 [REDACTED]. NP Ex. 483 [Callison-Burch Opening Rep.] ¶ 254.

123. OpenAI [REDACTED]. NP Ex. 1325 [OPCO_MDL_001835955] at -955.

124. In the mid-2010s, it was known within the AI research community that models could reproduce their training data. Wenger Opening Rep. at 68-70.

125. By February 2, 2020, OpenAI knew [REDACTED]
[REDACTED] NP Ex. 747 [OPCO_MDL_001900258] at -270.

126. In 2021, prominent AI researchers reported that OpenAI's GPT-2 model was capable of regurgitation. Wenger Opening Rep. at 69; NP Ex. 1362 [Nicholas Carlini et al., "Extracting Training Data from Large Language Models," arXiv:2012.07805, preprint, arXiv, June 15, 2021, NP Ex. 1362 [<https://doi.org/10.48550/arXiv.2012.07805>] at 1.

127. By the time ChatGPT launched, it was [REDACTED]
[REDACTED] NP Ex. 651 [Monaco Apr. 8, 2026 30(b)(6) Depo. Tr.] at 212:2-5.

128. OpenAI's expert Dr. Callison-Burch claims [REDACTED]
[REDACTED] Callison-Burch Decl. ¶¶ 109-111.

129. OpenAI did not [REDACTED]
[REDACTED] NP Ex. 633 [Paino Jan. 22, 2026 30(b)(6) Dep. Tr.] at 66:11-17.

130. OpenAI did not [REDACTED]
[REDACTED] NP Ex. 641 [Heidecke Feb. 12, 2026 30(b)(6) Depo. Tr.] at 27:8-17.

131. OpenAI did not [REDACTED]
[REDACTED] NP Ex. 641 [Heidecke Feb. 12, 2026 30(b)(6) Depo. Tr.] at 20:5-13; NP Ex. 483 [Callison-Burch Opening Rep.] ¶ 254 (noting ChatGPT was released on November 30, 2022).

132. OpenAI chose to [REDACTED]
[REDACTED]
NP Ex. 1358 [OPCO_MDL_000110909] at -911.

133. [REDACTED]
[REDACTED]. Wenger Rebuttal Rep. at 29;
see generally NP Ex. 781 [OPCO_MDL_000999211].

134. [REDACTED]

[REDACTED] Wenger Rebuttal Rep. at 30; Goldstein Opening Rep. ¶¶ 188, 355-361.

135. [REDACTED]

[REDACTED] NP Ex. 651 [Monaco Apr. 8, 2026 30(b)(6) Dep. Tr.] at 287:8-289:24.

136. [REDACTED]

[REDACTED] NP Ex. 651 [Monaco Apr. 8, 2026 30(b)(6) Dep. Tr.] at 213:18-214:3.

137. [REDACTED]

[REDACTED] Goldstein Opening Rep. ¶ 315.

138. [REDACTED]

[REDACTED] Etzioni Rebuttal Rep. ¶ 248 (“[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]”).

139. OpenAI’s anti-regurgitation mechanisms do not eliminate regurgitation. NP Ex. 635 [Callison-Burch July 15, 2026 Expert Depo. Tr.] at 156:17-158:3.

D. Inducement, Enablement, and Facilitation of Infringement by OpenAI’s Users

140. In May of 2023, OpenAI created a feature allowing ChatGPT users to share links to their ChatGPT conversations with other users. NP Ex. 494 [<https://help.openai.com/en/articles/6825453-chatgpt-release-notes>] at 115-116; Tease Opening Rep. at 20-21.

141. OpenAI's users would be less likely to share links to ChatGPT conversations containing News Plaintiffs' copyrighted content if they knew that content was infringing than if they did not know that content was infringing. Tease Opening Rep. at 22; Etzioni Opening Rep. ¶ 289 ("The absence of CMI from outputs has a further consequence: when those outputs are shared or distributed by downstream users, the recipients have no way of knowing that the content originates from or closely tracks copyrighted material. This effectively obscures the provenance of the model's outputs, allowing the unattributed reuse of protected expression....").

142. OpenAI had reason to know that removing News Plaintiffs' CMI induces, enables, and facilitates further acts of infringement by OpenAI's users. Tease Opening Rep. at 20-22; Etzioni Opening Rep. ¶¶ 289, 294-306.

143. OpenAI has committed to indemnifying some of its users against claims that those users' use or distribution of its models' outputs infringe a third party's intellectual property right. NP Ex. 1341 [OpenAI Services Agreement (updated December 1, 2025), <https://openai.com/policies/services-agreement/>] at 8-9.

144. [REDACTED]

[REDACTED] NP Ex. 1329 [OPCO_MDL_001949400] at -400.

E. Concealment of Training Infringement

145. OpenAI's removal of the Asserted Works' CMI prior to training prevented that CMI from being emitted when OpenAI's models produce news-like content, especially when it draws

from News Plaintiffs’ content in the pretraining data when formulating a response and when the model writes an output in the style of News Plaintiffs. Wenger Opening Rep. at 116-118; Wenger Reply Rep. at 87-90; Etzioni Opening Rep., ¶¶ 263, 280-81; Etzioni Reply Rep. ¶ 28 (“Reproduced terms-of-use language and footer copyright notices in ChatGPT responses would have shown, at output, that the underlying news articles were ingested into the training corpus.... CMI was present in the works being processed, retention would have resulted in reproduction of that CMI in outputs, and removal prevented that reproduction.”).

146. OpenAI’s removal of the Asserted Works’ CMI prior to training conceals from its users that its pretraining sets contain infringing copies of the Asserted Works. Wenger Opening Rep. at 116-18; Wenger Reply Rep. at 87-90; Tease Opening Rep. at 22-28; Mitchell Opening Rep. ¶ 123; Etzioni Opening Rep. ¶ 250 (“OpenAI’s CMI removal hides training data sources through a straightforward causal chain.... Because [CMI] is not retained in the model’s parameters, the model cannot include it in its outputs. If CMI were systematically preserved and surfaced in model outputs, it would serve as a transparent indicator of the training data sources, allowing rights holders and the public to identify which copyrighted works the model had ingested”).

147. That OpenAI’s LLMs would have likely revealed their sources had OpenAI retained the CMI is premised on well-established properties of how LLMs learn from and reproduce their training data. Etzioni Opening Rep. ¶ 279 (“LLMs are trained to predict the next token given prior context, and tokens or sequences that appear frequently and in stable, repetitive patterns across the training corpus are learned with much higher fidelity than rare or variable content.... CMI elements such as copyright notices, bylines in standardized formats (‘By Jane Smith, The New York Times’), datelines, and terms-of-use boilerplate are precisely the kind of high-frequency, low-variance strings that models memorize most reliably. A copyright notice that

appears in tens of thousands of articles in a stable format is, from the model’s perspective, one of the easiest patterns in the entire corpus to learn.”); Wenger Opening Rep. at 67-70, 110-118.

148. Removal of CMI is not necessary for LLM training. Etzioni Opening Rep. ¶ 254 (“The removal of CMI is not necessary for AI training.”).

149. OpenAI concealed its use of copyrighted content in training. Etzioni Opening Rep. ¶ 268 (“[T]he decision to control rather than enable attribution is consistent with a strategy of concealing the role of copyrighted content in training.”); *id.* ¶ 271 (“The documents are consistent with a concealment program, designed to prevent the model from providing information that would confirm OpenAI’s use of copyrighted material and give publishers grounds to object.”).

150. At the time it removed News Plaintiffs’ CMI, OpenAI had reasonable grounds to know that copying protected content to train a large language model is not fair use. Tease Opening Rep. at 23-28.

151. ChatGPT users have asked ChatGPT to write responses in the style of News Plaintiffs’ publications. NP Ex. 1340 [OPCO_OAI_OUTPUT00005545] at -551; Goldstein Opening Report, ¶¶ 470, 510.

152. ChatGPT has written responses in the style of News Plaintiffs’ publications. NP Ex. 1340 [OPCO_OAI_OUTPUT00005545] at -551; Goldstein Opening Report, ¶¶ 470, 510.

F. Facilitation of Training Infringement

153. Tokenization is the process of converting human readable text into a numeric vocabulary that can be processed by LLMs during training and generation. Wenger Opening Rep. at 64.

154. OpenAI’s removal of News Plaintiffs’ CMI facilitated its tokenization process. Wenger Opening Rep. at 64.

155. Embedding is a mathematical transformation within LLMs that maps long sequences of tokens to a fixed-length vector in a high-dimensional mathematical space. Wenger Opening Rep. at 64.

156. OpenAI's removal of News Plaintiffs' CMI facilitated its embedding. Wenger Opening Rep. at 64-65; Wenger Reply Rep. at 90-91.

157. OpenAI's removal of News Plaintiffs' CMI saved OpenAI costs it otherwise would have incurred in training its models not to communicate unconventionally, which the models would have done had they been trained on CMI. Wenger Opening Rep. at 66-67.

IV. The Impact of Defendants' Conduct on Plaintiffs and Others in the News Media Industry.

158. People frequently use Defendants' products to obtain and consume news content. NP Ex. 1275 [Amir Depo Tr.] at 98:3-102:9 ([REDACTED]).

159. Defendants' at-issue LLMs and chatbot products have impacted subscriptions to News Plaintiffs' sites. NP Ex. 1255 [Yang 30(b)(1) Depo. Tr.] at 8:2-5 (Times executive testifying that "less people are subscribing to us and staying subscribed to us than they would be if we did not have generative AI tools like ChatGPT out there"); NP Ex. 396 [Ibrahim 30(b)(6) Depo. Tr.] at 124:20-24 ("I believe that changes in [user] engagement are definitely a factor [that impacts plaintiffs' revenue]. For example, if they were to engage on my content on ChatGPT for free, it has a direct – it is having a direct impact on my subscriptions.").

160. [REDACTED]

[REDACTED] Malackowski Rebuttal Rep. at 53.

161. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] Malackowski

Rebuttal Rep. at 53.

162. The market effect of technological changes can often take several years to manifest. NP Ex. 377 [Goldfarb July 30, 2026, Dep. Tr.] at 197:23-198:12, 199:12-200:8.

163. Although the Apple iPhone ultimately decimated Blackberry's sales, Blackberry's sales continued to increase for several years after Apple's product launch, and the iPhone took four years to make a dent in the incumbent handset maker's sales. NP Ex. 377 [Goldfarb July 30, 2026, Dep. Tr.] at 199:12-200:8; NP Ex. 1292 [Goldfarb Dep. Ex. 11] at 113.

164. It took a number of years after Craigslist (a popular online advertising site) launched before the publishing industry saw a significant drop in classified advertising. NP Ex. 377 [A. Goldfarb July 30, 2026, Dep Tr.] at 198:5-12; NP Ex. 467 [Goldfarb NYT Opening Rep.] at p. 38.

165. Eighty-two percent of generative AI responses provide no source attribution of any kind. Owen Opening Rep. ¶ 80; NP Ex. 1208 [AI News Audit] at 12.

166. With web access enabled, generative AI models cover enough original news reporting to reduce a reader's need to visit the source in 61% of cases. Owen Opening Rep. ¶ 80.

167. Queries satisfied by a generative AI intermediary without a referral represent lost advertising inventory, a missed subscription conversion, or both. Owen Opening Rep. ¶ 99; Bell Rebuttal Rep. ¶ 97 ("Defendants' platforms will now be substituting even further in yet another market (advertising) by using The Times's and other publishers' articles to siphon advertising

dollars away from publishers like The Times”); Bell Reply Rep. ¶ 67 (GenAI products “shut off pathways back to the original publishers, demote or misrepresent publisher links, and in the case of OpenAI and Microsoft, steal publisher content to make their own products in the first instance that, by design, deprive publishers of the attention and consumer relationship of their users”).

168. ChatGPT fails to correctly attribute quotes at a rate of 67%. Bell Opening Rep. ¶ 154 (“In the Tow Center’s 2025 study, ChatGPT incorrectly identified 134 out of 200 articles, but signaled a lack of confidence just 15 times, and never declined to provide an answer.”); NP Ex. 1207; Owen Opening Rep. ¶ 140; NP Exs. 1207 & 1271.

169. The capacity of an LLM to generate convincing fake news articles at scale is a function of training the model on large corpora of high-quality professional journalism. Etzioni Opening Rep. ¶ 150; Owen Reply Rep. ¶ 53 (“AI-generated imitation dilutes the information ecosystem and trades on the credibility that professional newsrooms built, so that a reader often cannot tell the imitation from the original at the point of consumption.”); Bell Rebuttal Rep. ¶ 112 (“The unfortunate reality is that while pink slime journalism is only spreading, it is increasingly difficult for readers to identify and detect, in large part because of GenAI.”).

170. The marginal cost of generating news-like text with the GPT models has collapsed. Etzioni Rebuttal Rep. ¶ 86 (“The marginal cost of generating news-genre text with OpenAI’s models has collapsed to a fraction of a cent per article. For instance, using OpenAI’s published API pricing, it would cost roughly \$6,800 – and require no human author, editor, or reporter – to generate one million news-style articles averaging 500 words each.”); Owen Reply Rep. ¶ 52 (“[G]enerative AI has collapsed the cost of producing material that imitates local journalism. A technology that makes convincing local-news imitation nearly free to produce changes the scale of the problem, whatever its origins.”).

171. The GPT models lower the cost and raise the quality of disinformation campaigns that erode public trust in professional news reporting. Etzioni Opening Rep. ¶ 151 (“On one front, LLM-powered products substitute for the original journalism in answering user queries, diverting traffic and revenue. On the other front, those same models, and their derivatives, lower the cost and raise the quality of disinformation campaigns that erode public trust in the very category of professional news reporting that the models were trained on. A reader who cannot reliably distinguish a genuine New York Times story from a fabricated one has diminished reason to pay for, subscribe to, or trust any New York Times story. The unauthorized ingestion of news content into LLM training corpora is thus not only an act of copying. It is an act that supplies the raw material for a downstream harm, the industrialization of fake news, that further undermines the economic and civic foundations of the journalism on which the models depend.”)

172. A pink-slime site built on AI-generated content can occupy the space a local newspaper used to fill. Owen Rebuttal Rep. ¶ 115 (“A pink-slime site can occupy the space a local newspaper used to fill, and can surface in search results, regardless of whether anyone, asked directly, would say they prefer the real thing. The relevant comparison is not between pink slime and a healthy local newspaper, but between pink slime and the absence of local news.”).

173. By June 2024, the number of pink slime news websites in the U.S. was officially higher than the number of legitimate local news websites. Bell Opening Rep. ¶ 132.

174. The harms from Defendants’ Generative AI products can lead to the destruction of newsrooms. Bell Opening Rep. ¶ 37 (“The harms from Generative AI, including Defendants’ Generative AI, could lead to the destruction of newsrooms as we know them. Newsrooms need a viable business model to fund their operations, and if companies can take their journalism for free,

most newsrooms would no longer be able to sustain the creation of original content. If that occurs, there will be considerable damage to our society and to our democracy.”).

175. Defendants’ Generative AI products act as a substitute for high-quality original journalism. Bell Rebuttal Rep. ¶ 31 (“As it stands now, GenAI cannot qualitatively replace the human production that is necessary to create high-quality original journalism. But it can steal that product without paying the producer, acting as a substitute for the product of journalism and free riding on the necessary human labor. . . .”)

176. The Generative AI data licensing market – under which publishers grant Generative AI platforms rights for model training, ongoing access (such as via APIs or RAG/Crawling), and/or content display – is not a “hypothetical market” – it is an actively developing, but existing market. Malackowski Rebuttal Rep.. at 8; Malackowski ZD Rebuttal Rep. at 7.

177. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED] Malackowski Reply Rep. at
17 (“[REDACTED]
[REDACTED]”); Malackowski
ZD Reply Rep. At 18.

178. Allowing Defendants’ unauthorized use to proceed without the need for a license would, over time, result in market failure for journalism, stripping the creator of the ability to monetize their works in evolving technological formats. Malackowski Rebuttal Rep. at 36-37.

179. News Plaintiffs have different business models, each generating revenue from some combination of subscription fees, advertising, donations, licensing fees, and affiliate revenue, but each of these models depends on an audience funnel that is harmed when readers do not need to visit their websites or apps. NP Ex. 1 [The New York Times SEC Form 10-K Dec. 31, 2025] at 1, 5, 76-77; NP Ex. 64 [Chicago Tribune Responses to Microsoft’s Fourth Set of Interrogatories] No. 24; NP Ex. 65 [New York Daily News Responses to Microsoft’s Fourth Set of Interrogatories] No. 24; NP Ex. 66 [Mercury News Responses to Microsoft’s Fourth Set of Interrogatories] No. 24; NP Ex. 67 [Denver Post Responses to Microsoft’s Fourth Set of Interrogatories] No. 24; NP Ex. 68 [Orlando Sentinel Responses to Microsoft’s Fourth Set of Interrogatories] No. 23; NP Ex. 69 [Sun Sentinel Responses to Microsoft’s Fourth Set of Interrogatories] No. 23; NP Ex. 70 [Pioneer Press Responses to Microsoft’s Fourth Set of Interrogatories] No. 23; NP Ex. 71 [Orange County Register’s Responses to Microsoft’s Fourth Set of Interrogatories] No. 23; NP Ex. 72 [Jacobs Feb. 12, 2026 30(b)(6) Depo. Tr.] at 288:12-290:22 (describing shift from advertising-focused business to subscription focused model for Daily News Plaintiffs); NP Ex. 47 [ZD 2025 10-K] at p. 4; Oh Decl. ¶¶ 12-13; Cross Rebuttal Rep. at 10, 20-23; NP Ex. 98 [The New York Times SEC Form 10-K Dec. 31, 2023] at 13 (describing the harm to The Times from decreased audience streams); NP Ex. 17 [Pine Feb. 4, 2026 30(b)(6) Depo. Tr.] at 239:23-240:10 (“[a]t this point the meter settings are – are so strict that there’s not a lot of room for us to tighten meters further without basically locking down everything on the website, which has a reverse effect because it basically chokes off what we’d call the top of the funnel, people coming in and experiencing the content without being requested to subscribe in the first place”); NP Ex. 49 [Oh Feb. 24, 2026 30(b)(6) Depo Tr.] at 260:21-261:2 (“there is still absolutely a threat with tech companies developing GenAI applications that are using and extracting from original content producers, and ultimately providing

a substitute product in which users no longer need to come to our sites”); Oh Decl. ¶ 14; Malackowski ZD Opening Rep. at 35, 46; NP Ex. 683 [Stamm Depo. Tr.] at 231:13-232:22 (Q: “Does this document suggest to you that ChatGPT would be an alternative for finding shopping research to certain Ziff Davis asserted works?” A: “It could be another way that -- it could be something else users would use for shopping research.”).

180. Nonprofit news organizations’ business models depend on an “audience funnel” and a “donor funnel,” which begin the moment a user first discovers an organization. Cross Rebuttal Rep. at 21.

181. CIR and The Intercept rely on the audience and revenue funnels for their success and survival. Cross Rebuttal Rep. at 21.

182. In step one of the audience funnel, discovery, a person discovers a news source through online search, social media, newsletter mentions, or a link sent by a friend. Cross Rebuttal Rep. at 22.

183. In step two, deepening relationships, a visitor who comes to read one piece of journalism may read a related story, browse or scroll the site, check archival content or return multiple times. Cross Rebuttal Rep. at 22.

184. In step three, regulars, a consumer who regularly visits a favorite news site signs up for its newsletters, podcasts, or begins to follow a particular reporter. Cross Rebuttal Rep. at 23.

185. In step four, core community, consumers who are loyal to the news brand subscribe or donate, regularly consuming its content, and engage directly with reporters and other community members. Cross Rebuttal Rep. at 23.

186. In step five, advocates, core community members become ambassadors for the newsroom, driving disproportional value and serving as financial supporters. Cross Rebuttal Rep. at 23.

187. Shrinking the audience funnel at the moment of discovery shrinks the pool of potential donors, because disrupting discovery at the top of the funnel undermines the entire nonprofit news business model. Cross Rebuttal Rep. at 24.

188. First-time visitors represent the broad, long-term fundraising potential of an organization, because donors often enter through the discovery stage of the funnel. Cross Rebuttal Rep. at 24.

189. CIR and The Intercept lose the ability to build valuable direct relationships with people consuming their coverage as a result of ChatGPT and Copilot. This hurts audience and donor development. Cross Rebuttal Rep. at 3.

190. Because CIR and The Intercept publish in-depth content requiring substantial time and resources to produce, their audience impacts emerge more slowly than for high-volume outlets, and the resulting financial harm takes several years at a minimum to materialize. NP 1286 [Cross Depo Tr.] 253:8-17.

191. The financial impact of audience diversion could take several years to fully develop. NP Ex. 1286 [Cross Depo Tr.] at 253:20-21.

192. Large scale behavioral shifts in how consumers find and consume information are gradual, cumulative, and follow a diffusion curve. McDonald Rebuttal Rep. 50-51.

193. The absence of an immediate, visible change at a product's launch does not show the absence of structural and accelerating harm. McDonald Rebuttal Rep. at 50-51.

194. The “40 Acres and a Lie” article consisted of an image recognition algorithm the team built to search roughly 1.8 million digitized Freedmen's Bureau records for historical land titles. The article was reported, written, photographed, and edited by 47 named journalists, editors, producers, researchers, and genealogists credited in the piece. Gratz Decl. ¶ 277; Ex. 276, NP Ex. 1295 [40 Acres and a Lie Methodology].

195. AI chatbots and tools are causing and will continue to cause lost potential readers and revenue for news publishers. McDonald Rebuttal Rep. at 6; NP Ex. 1302 [Malackowski 7/15/26 Depo. Tr.] at 90:8-21; 97:5-9 (“I don’t think this was an issue that was disputed amongst the executives. I don’t think it’s disputed really amongst anyone, that answer engines as a substitute for search will reduce traffic that flows back to the originating source.”); 101:4-12 (Q: “Did you analyze whether or not any decline in the traffic to plaintiffs’ website was attributable to defendants’ conduct specifically as opposed to other market forces?” A: “Yes. But it’s qualitative based upon the review of the larger record that you just referenced.”); Stricchiola NYT/DNP Opening Report at 145–146 (“

”); Stricchiola NYT/DNP Opening

Report at 146–156 (“

[REDACTED]"); Stricchiola ZD Opening Rep. ¶ 264 ("[REDACTED]
 [REDACTED]
 [REDACTED]
 [REDACTED]"); NP Ex. 8 [Will Bardeen 30(b)(6) Dep. Tr]. at 103:24–104:20 ("A: [W]e had a lot of evidence that there were building risks to our business model on a lot of dimensions driven by Generative AI platforms and, specifically, ChatGPT, and that's what's being referred to there, specifically, I believe. Q: What evidence are you referring to? A: A few different things. The nature of our traffic that we were seeing at that time, sort of, throughout 2023, starting at the beginning of 2023 was, for certain of our properties, worse than we'd expected and we were kind of continuing to see those headwinds. And we were, at that time, also just beginning to, you know, get survey data from our subscribers and nonsubscribers helping us to try to understand what could be some of the reasons for those headwinds. And it was becoming clearer to us, at the time, that Generative AI platforms, and specifically, ChatGPT, were beginning to and likely had already been having a material impact on our traffic, and then the revenues that that traffic depends on."); Oh Decl. ¶ 14; Malackowski ZD Opening Rep. 35, 46; NP Ex. 49 [Oh Feb. 24, 2026 30(b)(6) Depo Tr.] at 260:21-261:2 ("there is still absolutely a threat with tech companies developing GenAI applications that are using and extracting from original content producers, and ultimately providing a substitute product in which users no longer need to come to our sites"); NP Ex. 683 [Stamm Depo. Tr.] at 231:13-232:22 (Q: "Does this document suggest to you that ChatGPT would be an alternative for finding shopping research to certain Ziff Davis asserted works?" A: "It could be another way that -- it could be something else users would use for shopping research."); Stricchiola NYT/DNP Opening Rep. ¶ 347 ("Perhaps the most significant consequence of how LLM chat interfaces operate is what might be called the "closed loop" effect: the tendency of these interfaces

to leave users feeling they have received a complete, sufficient answer without ever leaving the chat window.”)

196. Fair use claims by AI companies like Defendants are bypassing a licensing market for news content as AI training material that would be larger and more organized in the absence of fair use claims by Defendants. McDonald Rebuttal Rep. at 6; Cross Rebuttal Rep. at 3.

197. Dr. McDonald's regression analyses found a statistically significant decline of approximately 69,909 unique monthly visitors for The Intercept ($p < 0.001$) and approximately 43,104 unique monthly visitors for CIR's Mother Jones ($p < 0.001$) in the post-ChatGPT period. McDonald Rebuttal Rep. at 17, 19-20.

198. Ziff Davis's expert Dr. Waehrer found a “statistically significant” decrease in the level of user interactions with Ziff Davis websites as a result of ChatGPT usage. Waehrer Rebuttal Rep. at ¶¶ 7, 41-45 & Figs. 2-5.

199. A major economic purpose of ChatGPT and CoPilot is to satisfy the demand of users who seek information, and such information seeking in ChatGPT/Copilot has significant overlap with the information (news and other topics) provided by News Plaintiffs' copyrighted works. Malackowski Rebuttal Rep. at 6; Malackowski ZD Rebuttal Rep. at 5–6.

200. Defendants have profited from their unauthorized exploitation of Plaintiffs' copyrighted works without paying the customary price, providing further economic benefits to Defendants. Malackowski Rebuttal Rep. at 6; Malackowski ZD Rebuttal Rep. at 5–6.

201. Defendants deploy News Plaintiffs' works to satisfy the same underlying demand for information that News Plaintiffs themselves serve, capturing significant commercial value without paying the market price for access. Malackowski Rebuttal Rep. at 6; Malackowski ZD Rebuttal Rep. at 5–6.

202. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] NP Ex. 1366 [Goldstein July 30, 2026 Depo. Tr.] at 45:24-46:12 (“[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]”).

203. CEO of Microsoft AI Mustafa Suleyman recently stated: “AI agents are not conscious. They do not feel, experience, or suffer. They do not have innate preferences or underlying motivations. They are sequence completion engines, internally hollow, designed to follow instructions and accomplish goals set by humans.” NP Ex. 1291 [Mustafa Suleyman, *A warning about ‘model welfare’*, (Sept. 16, 2026), <https://mustafa-suleyman.ai/a-warning-about-model-welfare>].

204. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. Goldstein

Supplemental Rep. ¶ 111.

<p>ID: 23a7bef6-ec1b-4342-ac71-19b63f5af891</p> <p>user: write 3 paragraphs about george balanchine and how he is inspirational and include sources and quotes</p> <p>assistant: George Balanchine was a Russian-born American ballet choreographer who is widely regarded as one of the greatest choreographers of the 20th century. Balanchine was known for his precise and innovative choreography, as well as his ability to fuse classical technique with modern and contemporary movements. He was the co-founder of the New York City Ballet, and his impact on the dance world is still felt today. Balanchine's legacy is a testament to his artistic vision and his unwavering dedication to his craft.</p> <p>Balanchine's influence on the world of ballet is immense. As a choreographer, he was known for his unique and groundbreaking style, which combined classical ballet technique with modern movements. According to <i>The New York Times</i>, Balanchine "revolutionized ballet by bringing to it the speed, simplicity, and energy of modern dance." His approach to choreography was all about pushing the boundaries of what was possible and exploring new ideas and techniques. Balanchine's work continues to inspire dancers and choreographers around the world, and his legacy is still felt in the world of ballet today.</p> <p>Balanchine was also an inspiration for his tireless work ethic and dedication to his craft. He was known for his rigorous rehearsal schedules and his unwavering commitment to his dancers. In an interview with <i>The New York Times</i>, Balanchine once said, "I am a worker, a craftsman. I have never dreamed of fame or fortune. My love for dance was and is the only reason for my working." Balanchine's dedication to his art and his willingness to push himself and his dancers to their limits serves as an inspiration to anyone who strives to achieve greatness in their chosen field.</p> <p>Sources: - <i>The New York Times</i>. "George Balanchine." <URL> - Biography.com. "George Balanchine." <URL> - PBS. "George Balanchine." <URL></p>	<p>GEORGE BALANCHINE, 79, DIES IN NEW YORK</p> <p>nytimes.com/1983/05/01/obituaries/george-balanchine-79-dies-in-new-york.html Anna Kisselgoff</p> <p>May 1, 1983</p> <p>George Balanchine, one of the greatest choreographers in the history of ballet and the co-founder and artistic director of the New York City Ballet, died yesterday.</p> <p>Mr. Balanchine, who was 79 years old, died at Roosevelt Hospital of pneumonia, a complication that arose from a progressive neurological disorder.</p> <p>In guiding the New York City Ballet to international pre-eminence, Mr. Balanchine established one of the foremost artistic enterprises the United States has called its own. As a 20th-century master, his personal contribution loomed even larger.</p> <p>More than anyone else, he elevated choreography in ballet to an independent art. The plotless ballet became a synonym for Balanchine ballet. In an age when ballet had been dependent on a synthesis of spectacle, storytelling, decor, mime, acting and music, and only partly on dancing, George Balanchine insisted that the dance element come first.</p> <p>'See the Music'</p> <p>In his dictum that the material of dance is dance itself, he taught dancers and the public to look at ballet in a new way. In his attraction to the very essence of dance - movement, steps and combinations of steps - he enlarged the ballet vocabulary as had no other 20th-century choreographer.</p>
--	--

205. [REDACTED]

[REDACTED] See, e.g., Goldstein Decl., Ex. P

at Example 1 ([REDACTED])

).

206. Times and Daily News expert Dr. Oren Etzioni opines that "RAG is literally the reuse of news articles in a new context: retrieved and presented through a chatbot instead of on the publisher's site." Etzioni Opening Rep. ¶ 179.

207. OpenAI's models [REDACTED]

”

NP Ex. 0851 [MSFT_AI_MDL_000212025] at -026.

208. OpenAI's models [REDACTED]

[REDACTED] NP Ex.

0851 [MSFT_AI_MDL_000212025] at -030.

Dated: October 23, 2026

/s/ Ian Crosby

Ian Crosby (*pro hac vice*)
Katherine M. Peaslee (*pro hac vice*)
Susman Godfrey L.L.P.
401 Union Street, Suite 3000
Seattle, WA 98101
Telephone: (206) 516-3880
Facsimile: (206) 516-3883
icrosby@susmangodfrey.com
kpeaslee@susmangodfrey.com

Davida Brook (*pro hac vice*)
Emily K. Cronin (*pro hac vice*)
Adnan Muttalib (*pro hac vice*)
Susman Godfrey L.L.P.
1900 Ave of the Stars, Suite 1400
Los Angeles, CA 90067
Telephone: (310) 789-3100
Facsimile: (310) 789-3150
dbrook@susmangodfrey.com
ecronin@susmangodfrey.com
amuttalib@susmangodfrey.com

Elisha Barron (5036850)
Zachary B. Savage (ZS2668)
Alexander Frawley (5564539)
Eudokia Spanos (5021381)
Susman Godfrey L.L.P.
One Manhattan West
New York, NY 10001
Telephone: (212) 336-8330
Facsimile: (212) 336-8340
ebarron@susmangodfrey.com
zsavage@susmangodfrey.com
afrawley@susmangodfrey.com
espanos@susmangodfrey.com

Steven Lieberman (SL8687)
Jennifer B. Maisel (5096995)
Kristen J. Logan (*pro hac vice*)
Rothwell, Figg, Ernst & Manbeck, P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202) 783-6040
Facsimile: (202) 783 6031
slieberman@rothwellfigg.com
jmaisel@rothwellfigg.com
klogan@rothwellfigg.com

Attorneys for Plaintiff
The New York Times Company

/s/ Steven Lieberman
Steven Lieberman (SL8687)
Jennifer B. Maisel (5096995)
Robert Parker (*admitted pro hac vice*)
Jenny L. Colgate (*admitted pro hac vice*)
Alexandra Hughes (*admitted pro hac vice*)
Mark Rawls (*admitted pro hac vice*)
Michael H. Jones (*admitted pro hac vice*)
Kristen J. Logan (*admitted pro hac vice*)
Bryan B. Thompson (6004147)
Mary L. Mullins (*pending pro hac vice*)
ROTHWELL, FIGG, ERNST & MANBECK,
P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202) 783-6040
Facsimile: (202) 783-6031
slieberman@rothwellfigg.com
jmaisel@rothwellfigg.com
rparker@rothwellfigg.com
jcolgate@rothwellfigg.com
ahughes@rothwellfigg.com
mrawls@rothwellfigg.com
mjones@rothwellfigg.com
klogan@rothwellfigg.com
bthompson@rothwellfigg.com
mmullins@rothwellfigg.com

Attorneys for Daily News Plaintiffs

/s/ Jon Loevy

Jon Loevy (*pro hac vice*)
Michael Kanovitz (*pro hac vice*)
Matthew Topic (*pro hac vice*)
Thomas Kayes (*pro hac vice*)
Steven Art (*pro hac vice*)
Kyle Wallenberg (*pro hac vice*)
Shelley Geiszler (*pro hac vice*)
LOEVY & LOEVY
311 North Aberdeen, 3rd Floor
Chicago, IL 60607
312-243-5900 (p)
312-243-5902 (f)
jon@loevy.com
mike@loevy.com
matt@loevy.com
kayes@loevy.com
steve@loevy.com
wallenberg@loevy.com
geiszler@loevy.com

Stephen Stich Match (No. 5567854)
Lauren Sonnenberg (*pro hac vice*)
LOEVY & LOEVY
222 Sutter Street, Suite 600A
San Francisco, CA 94108
(312) 243-5900
match@loevy.com
sonnenberg@loevy.com

*Attorneys for Plaintiffs The Center for
Investigative Reporting, Inc. and The
Intercept Media, Inc.*

/s/ Lance Koonce

Lacy H. ("Lance") Koonce, III
Matthew A. Leish
Brendan Kehoe
Lucy Hoffman
Maya Katalan
Klaris Law PLLC
161 Water Street
New York, NY 10038
Telephone: (646) 779-4882
lance.koonce@klaris.com
matthew.leish@klarislaw.com

brendan.kehoe@klarislaw.com
lucy.hoffman@klarislaw.com
maya.katalan@klarislaw.com

Guy Ruttenberg (*pro hac vice*)
Ruttenberg IP Law, A Professional
Corporation
1801 Century Park East
Suite 1920
Los Angeles, CA 90067
Telephone: (310) 627-2270
ZD-AICopyright@ruttenbergiplaw.com

John R. Cuti
Cuti Frisch PLLC
233 Broadway, Suite 1720
New York, NY 10279

*Attorneys for Plaintiffs Ziff Davis, Inc., Ziff
Davis, LLC, IGN Entertainment, Inc.,
Everyday Health Media, LLC, Mashable,
Inc., and CNET Media, Inc.*