UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

THE SEATTLE TIMES COMPANY and NEWSDAY LLC, Plaintiffs, v. OPENAI, INC., OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO, LLC, OPENAI GLOBAL LLC, OAI CORPORATION, OPENAI HOLDINGS, LLC, OPENAI FOUNDATION, OPENAI GROUP PBC, and MICROSOFT CORPORATION, Defendants.	No. ______________________ **COMPLAINT** **JURY TRIAL DEMANDED**

Plaintiffs The Seattle Times Company ("The Seattle Times") and Newsday LLC ("Newsday") (together, "Plaintiffs"), by their attorneys Klaris Law, PLLC, for their Complaint against Defendants OpenAI, Inc., OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo, LLC, OpenAI Global, LLC, OAI Corporation, OpenAI Holdings, LLC, OpenAI Foundation, OpenAI Group PBC (collectively, "OpenAI"), and Microsoft Corporation ("Microsoft") (OpenAI and Microsoft, together, "Defendants"), allege as follows:

NATURE OF THE ACTION

1. By definition, journalism is created by reporters, editors, photographers, correspondents, videographers, digital producers and the organizations for which they work. Their products—newspapers, journals, and the like—are consumed by the reading public, who in turn defray the expense of independent journalism through subscriptions, advertising and other revenue.

2. While numerous societal and economic forces over recent years have caused cracks in the foundation of independent journalism, many respected news organizations such as Plaintiffs

The Seattle Times and Newsday have transformed and modernized while staying true to their public-service missions proving their resiliency and strength, while continuing to produce high-quality, reliable, award-winning journalism for their readers and acting as a bulwark for the promotion of democracy and a free press.

3. However, in the dark future of journalism being constructed by massive generative AI companies such as OpenAI and its partner Microsoft, this economic cycle may become broken beyond repair. AI products like ChatGPT and CoPilot are touted as producers of content, but in fact they are rapacious consumers, devouring human-authored content and delivering back to the world copies and derivative imitations of that same original content they consumed to achieve their commercial objectives.

4. Like a snake eating its own tail, GenAI that is trained on painstakingly researched, expensive-to-produce content threatens to destroy the very news organizations by competing directly with them through AI-generated substitutive content. If Defendants are allowed to succeed, independent journalism of the kind Plaintiffs produce will struggle to survive.

5. Founded in 1886, The Seattle Times has provided the Pacific Northwest with trusted, award-winning local, regional, and national news coverage, including landmark investigative journalism. Since 1940, Newsday has served Long Island and New York City with highly impactful news reporting.

6. With 30 Pulitzer Prizes between them, the Seattle Times and Newsday have both made extensive investments in independent, local journalism at a time when the economic headwinds in the news industry have continued to strengthen.

7. Over less than three years, Defendants OpenAI and Microsoft have built enormously lucrative generative artificial intelligence ("GenAI") businesses in substantial part by

copying, without permission or compensation, vast quantities of copyrighted journalism, including the original, expressive work of The Seattle Times and Newsday.

8. Upon information and belief, Defendants have for years obtained copies of Plaintiffs' copyrighted content by methodically scraping (or having other parties scrape) Plaintiffs' websites using automated bots, bypassing Plaintiffs' paywalls, and ignoring decades-old norms to obtain the Plaintiffs' crown jewels—copies of their news articles—without paying the customary price. Indeed, without paying any price at all.

9. Upon information and belief, and consistent with the methods OpenAI and Microsoft are alleged to have used against similarly situated news publishers, Defendants have incorporated those copies of Plaintiffs' content, again without permission, into large-scale datasets compiled by Defendants from indiscriminate scrapes of the internet, including but not limited to datasets derived from WebText, WebText2, Common Crawl, and Microsoft's own Bing search index. Defendants then used the content in those datasets to train, fine-tune, and ground the large language models ("LLMs") underlying their GenAI products.

10. Upon information and belief, in doing so, Defendants deliberately remove or alter copyright management information ("CMI") from Plaintiffs' content, such as article titles, author names, and copyright notices.

11. Once trained on Plaintiffs' content, Defendants' LLMs store copies of that content and, upon information and belief, regurgitate it verbatim when prompted.

12. Upon information and belief,

a. Defendants' products have produced output that consists of, or contains, verbatim and paraphrased output from Plaintiffs' copyrighted works;

b. Defendants' products have generated output that competes directly with Plaintiffs' copyrighted content and acts as a substitute for that content in the relevant consumer and licensing markets; and

c. Defendants have made copies of Plaintiffs' content for storage, for retrieval-augmented generation ("RAG"), and otherwise as output answering user queries.

13. Defendants have profited immensely from this conduct. OpenAI's ChatGPT and Microsoft's Copilot and Bing Chat products, all of which derive from OpenAI's LLM models, generate substantial revenue directly attributable to their ability to produce natural, authoritative, well-written responses to user queries—an ability built in large part on ingesting and mimicking, or outright plagiarizing, the creative output of working journalists, including those at The Seattle Times and Newsday, without compensation. Although Microsoft introduced the Bing search engine in 2009, it failed to gain critical usage until integrating with OpenAI's technology in February 2023, crossing 100 million daily active users for the first time in its history. After incorporating OpenAI's LLMs, page visits on Bing immediately rose over 15% in the following six weeks, an increase Microsoft and independent analysts attributed directly to the new AI-powered chat features built on content scraped from sites like Plaintiffs' websites.[1]

14. Defendants' conduct is not protected by the fair use doctrine. Publicly, OpenAI and Microsoft have argued that their unlicensed use of copyrighted news content to train GenAI models is "transformative" because it serves a new purpose distinct from the original journalism. But there is nothing transformative about copying The Seattle Times' and Newsday's journalism, without

---

[1] *See* Tom Warren, *Microsoft Bing Hits 100 Million Active Users in Bid to Grab Share from Google*, The Verge (Mar. 9, 2023, 11:03 AM), *available at* https://www.theverge.com/2023/3/9/23631912/microsoft-bing-100-million-daily-active-users-milestone; Akash Sriram & Chavi Mehta, *OpenAI Tech Gives Microsoft's Bing a Boost in Search Battle with Google*, Reuters (Mar. 22, 2023, 2:34 PM), *available at* https://www.reuters.com/technology/openai-tech-gives-microsofts-bing-boost-search-battle-with-google-2023-03-22/.

payment, to build products that compete directly with Plaintiffs for the same readers, the same advertisers, and the same subscription dollars. Defendants' GenAI products do not merely study or comment on Plaintiffs' journalism—they reproduce, closely paraphrase, and synthesize it into competing informational products that substitute for a visit to newsday.com or seattletimes.com, serving readers the same information Plaintiffs' journalism was created to deliver. Because Defendants' GenAI models compete with, and reproduce exactly or closely mimic, the content used to train them, copying Plaintiffs' works for overlapping purposes is not fair use.

15. Defendants' conduct violates Plaintiffs' rights under the Copyright Act, 17 U.S.C. Section 101 et seq., including as amended by the Digital Millennium Copyright Act ("DMCA"), and the Lanham Act, 15 U.S.C. Section 1125(c), as well as accompanying Washington and New York state laws preventing trademark dilution. Plaintiffs bring this action to recover damages for that conduct and to enjoin its continuation.

JURISDICTION AND VENUE

16. This Court has subject matter jurisdiction over Plaintiffs' claims pursuant to 28 U.S.C. §§ 1331 and 1338, as they arise under the Copyright Act. This Court further has jurisdiction under 28 U.S.C. Section 1367 with respect to the related state-law claims.

17. This Court also has personal jurisdiction over Defendants. Defendants have purposefully availed themselves of the privilege of conducting business in the State of New York, and a substantial portion of the unlawful conduct is alleged to have occurred in New York. OpenAI leases and occupies office space at the Puck Building, 295 Lafayette Street, in Manhattan, and, as of January 2025, employed approximately 450 people in New York City. Microsoft has maintained a continuous New York City presence for well over a decade, including its current offices at 11 Times Square and its prior offices at 1290 Avenue of the Americas. Its New York City footprint also includes 122 Fifth Avenue and 300 Lafayette Street, home to Microsoft Research New York

City, a lab founded in 2012. Microsoft's subsidiary, LinkedIn, occupies roughly 500,000 square feet at the Empire State Building. Upon information and belief, by and through their employees in NYC, Defendants have engaged in the unlawful and infringing conduct at issue in this suit, including the creation and distribution of its products and services containing illicit copies Plaintiffs' copyrighted works.

18. Venue is proper under 28 U.S.C. Section 1400(a) because Defendants and their agents reside or may be found in this District. Venue is also proper in this District under 28 U.S.C. Section 1391(b)(2) because a substantial part of the events giving rise to Plaintiffs' claims occurred here, including the marketing and promotion, sale and distribution, and licensing of Defendants' GenAI infringing products as alleged herein.

19. Defendants have consented to personal jurisdiction and venue in this District in numerous similar actions, and have declined to contest jurisdiction or venue in, among others, *The New York Times Company v. Microsoft Corp.*, No. 23-cv-11195 (S.D.N.Y.), *Raw Story Media, Inc. v. OpenAI, Inc.*, No. 24-cv-01514 (S.D.N.Y.), *The Intercept Media, Inc. v. OpenAI, Inc.*, No. 24-cv-01515 (S.D.N.Y.), and *Daily News, LP v. Microsoft Corp.*, No. 24-cv-03285 (S.D.N.Y.), all of which have been consolidated in In re OpenAI, Inc. Copyright Infringement Litig., No. 25-md-03143 (S.D.N.Y.).

THE PARTIES

A. Plaintiffs

20. Plaintiff The Seattle Times Company is a Delaware Corporation with its principal place of business in Seattle, Washington. The Seattle Times was founded on May 3, 1886 by Thomas H. Dempsey and Jud R. Andrews. It was purchased by Alden J. Blethen in August 1896 and has remained under Blethen family ownership for five generations since. Published under the

name Seattle Daily Times for many years, its name was shortened to The Seattle Times in 1966, and it has used that name ever since.

21. The Seattle Times has won 11 Pulitzer Prizes since 1950—including three National Reporting prizes, most recently for the Boeing 737 MAX crisis in 2020 and for Breaking News Reporting of the 2015 Oso landslide. The Seattle Times has been a Pulitzer finalist an additional 15 times since 1982, most recently in 2026 for Breaking News coverage of catastrophic flooding in Western Washington.

22. The Seattle Times launched its first web presence in August 1995 and registered the seattletimes.com domain in March 1996. The Seattle Times implemented a metered digital paywall in the spring of 2013. In 2013, readers received 15 free pageviews before being required to pay for further access to The Seattle Times's content. Over the years, the number of free pageviews decreased. Today, readers receive one free pageview before payment is required.

23. The Seattle Times reaches approximately 1.7 million adults monthly in the Seattle-Tacoma designated market area, with seattletimes.com drawing approximately 6 million monthly unique visitors and 29 million monthly pageviews. The Seattle Times has approximately 108,000 digital-only subscribers and approximately 100,000 print subscribers including single-issue purchases.

24. The Seattle Times first registered the trademark THE SEATTLE TIMES with the U.S. Patent and Trademark Office (USPTO) in 2019 (Registration Nos. 5,738,990 (Int'l Class 16, first use 1896), 5,738,991 (Int'l Class 35, first use 1995), 5,738,992 (Int'l Class 41, first use 1995), and 5,738,989 (Int'l Class 9, first use 2010)), with first use in commerce dating back to 1896 (collectively, "The Seattle Times Marks"). The registration certificates corresponding to The Seattle Times Marks are included as **Exhibit 3**. On August 22 and 25, 2025, the U.S. Patent and Trademark Office accepted The Seattle Times' combined Section 8 and Section 15 declarations for

each of these registrations, rendering the mark THE SEATTLE TIMES' incontestable under 15 U.S.C. Section 1065. The mark THE SEATTLE TIMES is known in the minds of the news-media consuming public for high-quality journalism and has a national reputation for excellence.

25. The Seattle Times owns registered copyrights in its published journalism, including the works identified on **Exhibit 1** ("The Seattle Times Registered Works"). Upon information and belief, every one of The Seattle Times Registered Works was copied and ingested by OpenAI and its partner Microsoft for use in their GenAI products.

26. Plaintiff Newsday LLC is a Delaware limited liability company, with a principal place of business at 6 Corporate Center Drive, Melville, New York 11747.

27. Newsday was founded in 1940 and has served Long Island and New York City with award-winning journalism for over eighty-five years. Newsday has won 19 Pulitzer Prizes since 1954, including three Public Service gold medals—journalism's highest honor—for investigations such as the 1954 exposure of race-track scandals and labor racketeering at Roosevelt Raceway, and the 1974 investigation "The Heroin Trail," tracing heroin traffic from Turkish poppy fields to American streets. Newsday has also received recognition for sustained international reporting, including the 2005 International Reporting prize for "Rwanda: The Legacy of Hate." The reporter, Dele Olojede, was the first African-born Pulitzer winner in history. Newsday has continued to receive fame and recognition in the years since Defendants' conduct began, including a George Polk Award and a Peabody Award in 2020 for "Long Island Divided," and National Headliner and New York Press Association investigative-reporting honors in 2026 for "Broken Practice" and "Dangerous Roads." Like The Seattle Times, Newsday has restricted free access to its digital content through a series of increasingly restrictive paywall measures. Since August 2022, Newsday employs a complete hard gate under which no Newsday-produced editorial text content is available to non-subscribers.

28. Newsday is among the ten largest daily circulation U.S. newspapers, with a print circulation of over 84,000 Sunday copies, over 65,000 daily copies, and over 60,000 digital subscribers. Its website receives 51 million average monthly page views and reaches 2.1 million average monthly unique visitors.

29. Newsday first registered the trademark NEWSDAY with the USPTO on March 25, 1997 (Registration No. 2,047,787, Int'l Class 16), with first use in commerce dating back to 1940 ("Newsday Mark"). The registration certificate corresponding to the Newsday Mark is included as **Exhibit 4**. The Newsday Mark is incontestable under 15 U.S.C. Section 1065 and is known in the minds of the news-media consuming public for high-quality journalism. Newsday maintains a strong presence in Long Island/New York City region and has a national reputation for excellence.

30. Newsday owns registered copyrights in its published journalism, including the works identified on Exhibit 2 ("Newsday Registered Works"). Upon information and belief, every one of the Newsday's Registered Works was copied and ingested by OpenAI and its partner Microsoft for use in their GenAI products.

B. OpenAI Defendants

31. OpenAI was formed in December 2015 as a nonprofit "artificial intelligence research company," pledging that its research would be free from financial obligations and that its work and any patents would be open and shared with the world. When it applied for tax-exempt status, it told the IRS that its goal was

> to engage in research activities that advance digital intelligence in the way that is most likely to benefit humanity as a whole, unconstrained by a need to generate financial return. AI technology will help shape the 21st century, and OpenAI wants to help the world build safe Al technology and ensure that AI's benefits are as widely and evenly distributed as possible. To that end, OpenAI

> hopes to build AI as part of a larger community, and wants to openly share its plans and capabilities along the way.[2]

32. Contrary to those representations, OpenAI shed its exclusively nonprofit status shortly after, announcing in March 2019 that it was moving to a for-profit entity operating under OpenAI LP (subsequently renamed OpenAI OpCo, LLC) dedicated to raising capital from investors seeking a return and helping run OpenAI's commercial operations.[3]

33. In the same year, Microsoft invested $1 billion dollars in OpenAI, the first of many such investments. At the time, OpenAI described the purpose of the partnership as "building artificial general intelligence (AGI) with widely distributed economic benefits," and "to develop a hardware and software platform within Microsoft Azure which will scale to AGI."[4]

34. On October 28, 2025, OpenAI completed a complex corporate restructuring, with the OpenAI Foundation (a nonprofit) taking an equity stake in OpenAI's commercial operations that, at the time and by its own public statements, was valued at approximately $130 billion.[5] OpenAI Group PBC, a Delaware public benefit corporation, now operates OpenAI's ChatGPT and other commercial GenAI offerings.[6]

35. As a result of subsequent fundraising, OpenAI is now valued at $852 billion.[7] In June 2026, OpenAI announced that it had filed with the U.S. Securities and Exchange Commission

---

[2] OpenAI IRS Form 1023, dated December 8, 2015; *see also* Thalia Beaty, *Documents show OpenAI's long journey from nonprofit to $157B valued company*, AP News (Oct. 12, 2024, 5:52 PM), *available at* https://apnews.com/article/openai-nonprofit-application-tax-exempt-status-ebf38549cdf0529fc5ec727e2815ed12].

[3] OpenAI, *OpenAI LP*, *available at* https://openai.com/index/openai-lp/ (last accessed Sep. 3, 2026).

[4] Greg Brockman, *Microsoft Invests in and Partners with OpenAI to Support Us Building Beneficial AGI*, OpenAI (July 22, 2019), *available at* https://openai.com/index/microsoft-invests-in-and-partners-with-openai/ (last accessed Sep. 3, 2026).

[5] Bret Taylor, *Built to Benefit Everyone*, OpenAI (Oct. 28, 2025), *available at* https://openai.com/index/built-to-benefit-everyone/ (last accessed Sep. 3, 2026).

[6] *Id.*; Ashley Capoot, *OpenAI Completes Restructure, Solidifying Microsoft as a Major Shareholder*, CNBC (Oct. 28, 2025, 9:06 AM), *available at* https://www.cnbc.com/2025/10/28/open-ai-for-profit-microsoft.html (last accessed Sep. 3, 2026).

[7] OpenAI, *OpenAI Raises $122 Billion to Accelerate the Next Phase of AI* (Mar. 31, 2026), *available at* https://openai.com/index/accelerating-the-next-phase-ai/ (last accessed Sep. 3, 2026).

for an upcoming initial public offering, which is expected to fetch 100s of millions of dollars for its founders, employees, and investors.[8]

36. Defendant OpenAI, Inc. is a Delaware nonprofit corporation, formed in December 2015, with a principal place of business at 3180 18th Street, San Francisco, California. Until the October 2025 restructuring, OpenAI, Inc. indirectly owned and controlled all other OpenAI entities and has been directly involved in perpetrating the conduct alleged herein.

37. Defendant OpenAI OpCo, LLC (formerly OpenAI LP) is a Delaware limited liability company with a principal place of business at 3180 18th Street, San Francisco, California, and was directly involved in the design, development, and commercialization of OpenAI's GPT-based products.

38. Defendant OpenAI GP, LLC is a Delaware limited liability company with a principal place of business at 3180 18th Street, San Francisco, California, and has served as the general partner/manager directing OpenAI OpCo, LLC and OpenAI Global, LLC.

39. Defendant OpenAI, LLC is a Delaware limited liability company with a principal place of business at 3180 18th Street, San Francisco, California, formed in September 2020, and owns, sells, licenses, and monetizes OpenAI offerings including ChatGPT, ChatGPT Enterprise, and OpenAI's API tools.

40. Defendant OpenAI Global, LLC is a Delaware limited liability company formed in December 2022 with a principal place of business at 3180 18th Street, San Francisco, California.

41. Defendant OAI Corporation and Defendant OpenAI Holdings, LLC are Delaware limited liability companies with principal places of business at 3180 18th Street, San Francisco,

[8] Matt O'Brien, *OpenAI Files Confidential SEC Paperwork for IPO, Opening the Door to a Wall Street Debut*, The Seattle Times (June 8, 2026, 2:56 PM), *available at* https://www.seattletimes.com/business/openai-files-confidential-sec-paperwork-for-ipo-opening-the-door-to-a-wall-street-debut/.

California, forming part of the holding-company structure through which OpenAI, Inc. and, following the October 2025 restructuring, the OpenAI Foundation exercise ownership and control over OpenAI's operating entities.

42. Defendant OpenAI Foundation is, upon information and belief, a Delaware foundation with a principal office in 1455 3rd Street, San Francisco, California, and controls Defendant OpenAI Group PBC.

43. Defendant OpenAI Group PBC is, upon information and belief, a Delaware public benefit corporation with a principal place of business in San Francisco, California, and is the current corporate home for OpenAI's ChatGPT and related commercial GenAI operations.[9]

44. OpenAI's GPT-based LLMs and the products built on them have evolved substantially since ChatGPT's public launch in November 2022. On information and belief, and as relevant to this action, the OpenAI models at issue include GPT-1, GPT-2, GPT-3, GPT-3.5, GPT-3.5 Turbo, GPT-4, GPT-4 Turbo, GPT-4o, GPT-4o mini, o1, o1-pro, o1-mini, GPT-4.5, o3, o3-pro, o3-mini, o4-mini, GPT-4.1, GPT-4.1 mini, GPT-OSS, GPT-5, GPT-5 mini, GPT-5 nano, and GPT-5.1, together with any successor models developed using similar methods and trained in a similar manner. OpenAI's latest models include GPT-6 Astra, GPT-5-Codex, GPT-5.1, GPT-5.2, GPT-5.2-Codex, GPT-5.3, GPT-5.3-Codex, GPT-5.4, GPT-5.4 mini, GPT-5.4 nano, GPT-5.5, GPT-5.6 Sol, GPT-5.6 Terra, GPT-5.6 Luna, and GPT-5.6 Cyber.

45. OpenAI offers these models to the public and to businesses through, among other offerings, its ChatGPT consumer products (Free, Go, Plus, Pro, and Business tiers), ChatGPT Business, ChatGPT Enterprise, ChatGPT Edu, ChatGPT Search, ChatGPT Work, ChatGPT Sites, Codex, an application programming interface ("API") platform, and Custom GPTs, as well as, at

[9] OpenAI, *Our Structure*, *available at* https://openai.com/our-structure/ (last accessed Aug. 2026).

various times during the relevant period, ChatGPT Canvas, ChatGPT Agent, and the ChatGPT Atlas browser.

46. Each of the OpenAI entities identified in this complaint has had a role in developing, marketing, and releasing OpenAI's LLM-based products and services.

C. Defendant Microsoft

47. Defendant Microsoft Corporation is a Washington corporation with its principal place of business and headquarters in Redmond, Washington.

48. Since at least 2019, Microsoft has committed to invest a total of approximately $13 billion in OpenAI, of which at least $11.6 billion had been funded as of September 30, 2025.[10] In October 2025, Microsoft and OpenAI announced a "new definitive agreement" pursuant to which Microsoft "holds an investment in OpenAI Group PBC valued at approximately $135 billion, representing roughly 27 percent on an as-converted diluted basis."[11]

49. Microsoft obtained, and upon information and belief continues to obtain, direct access to OpenAI's GenAI models—models trained on and incorporating unauthorized copies of Plaintiffs' content—and uses those models to generate content for users of its own products and services, including Copilot and Bing Chat/Bing search.

50. Upon information and belief, Microsoft's Copilot and Bing Chat/Bing search products continue to incorporate OpenAI's GenAI models trained on Plaintiffs' content. Microsoft's products also used Plaintiffs' content to build the Microsoft Bing search index, which Microsoft accesses and makes copies of to generate output to users of its OpenAI's LLM-based

---

[10] Microsoft Corp., Quarterly Report (Form 10-Q) (Sept. 30, 2025), *available at* https://www.sec.gov/Archives/edgar/data/789019/000119312525256321/msft-20250930.htm (last accessed Sept. 3, 2026).

[11] Microsoft Corp., *The Next Chapter of the Microsoft-OpenAI Partnership*, Microsoft Blog (Oct. 28, 2025), *available at* https://blogs.microsoft.com/blog/2025/10/28/the-next-chapter-of-the-microsoft-openai-partnership/ (last accessed Sept. 3, 2026).

products. It is not surprising, then, that in a 2023 interview, Microsoft's CEO Satya Nadella stated that if "OpenAI disappeared tomorrow," Microsoft could "continue the innovation" on its own because "we have the data, we have everything," underscoring Microsoft's direct access to and control over the underlying copied content and the resulting models.[12]

FACTUAL ALLEGATIONS

A. Plaintiffs' Journalism and Business Model

51. Producing excellent journalism today is more difficult, more expensive, and more essential to society and our democracy, than ever. Over the past two decades, the traditional advertising and circulation-based business models that once supported quality journalism have come under sustained pressure, forcing the closure or contraction of newspapers across the country. Northwestern University's Medill School of Journalism has produced a "State of Local News Report" which most recently reported that the U.S. is losing more than 2 newspapers per week. The Seattle Times' own newsroom contracted significantly during 2000-2010, a period marked by technical disruption, the emergence of Craigslist decimating traditional classified revenues, and significant declines in print advertising revenue, which produced sizable staffing cuts and outsourcing of key business functions.

52. For more than a century, The Seattle Times has remained under family ownership, a commitment that has required significant sacrifice to sustain a sizable newsroom, invest in the technology needed to distribute the paper's coverage, and honor an enduring pledge to deliver quality journalism that drives change and impact in the Puget Sound region. Since 2011, The Seattle Times has sold 7.5 acres of real estate for $80.9 million and reinvested those proceeds in

[12] Kevin Okemwa, *Microsoft CEO says "It wouldn't matter if OpenAI disappeared tomorrow. We have the data, IP rights, and all the capability."*, Yahoo!Tech (Mar. 21, 2024), *available at* https://tech.yahoo.com/ai/articles/microsoft-ceo-says-wouldnt-matter-134433669.html (last accessed Sept. 3, 2026).

the company's ongoing commitment to modernizing without sacrificing the high-quality journalism its readers have come to expect. Since its founding, The Seattle Times has invested substantial resources in original reporting and investigative journalism, and, since 2013, has pioneered philanthropically supported "impact journalism" initiatives, including Education Lab, Climate Lab, the Mental Health Project, Traffic Lab, and Project Homeless. The Seattle Times retains full and exclusive editorial control over all content produced under those initiatives.

53. The Seattle Times, like other professional news publishers, depends on subscription, advertising, and licensing revenue to fund its journalism. The purpose of The Seattle Times's digital paywall was to deliver more value to print subscribers willing to pay for credible, rigorously reported, fact-based information regardless of how it was delivered. Consistent with that strategy, The Seattle Times restricts free access to its content through the paywall and, under its Terms of Service, requires third parties to obtain its express written consent before reproducing, distributing, or creating derivative works from its content for commercial purposes.

54. The Seattle Times' current Terms of Service expressly and specifically prohibit the use of its content to train or ground any form of artificial intelligence. They provide that a user "may not use or incorporate any Content available through the Sites and Services for any form of artificial intelligence ("AI"), including in any generative or other form of AI for training or grounding purposes," and separately prohibit "the use of webcrawler, spidering, script, site search/retrieval applications or other manual or automated device, tool, process or other means to retrieve, scrape, text or data mine information or content . . . other than as expressly authorized."[13] The Seattle Times' Terms of Service further require compliance with "exclusionary protocols (e.g.,

[13] The Seattle Times, *User Notices of The Seattle Times Company*, (July 9, 2026), *available at* https://www.seattletimes.com/notices/terms.html (last accessed Sept. 3, 2026), at Section 2(b).

Robots.txt, Automated Content Access Protocol (ACAP), etc.)," and prohibit any use of its content to create products that could "generate . . . a substitute for a subscription to the Sites or Services."[14]

55. The Seattle Times has also aggressively worked to protect its intellectual property by maintaining a diligent copyright registration program, registering its trademarks, and engaging in carefully modulated enforcement efforts. To further protect its rights, it has often foregone near-term, lucrative distribution opportunities to protect the core value of its properties and its customer relationships, including declining to distribute its content through programs such as Facebook Instant Articles, Apple News+, and Google Showcase.

56. As of 2025, Newsday was the sixth-largest newspaper in the United States by print circulation, and it continues to invest in original reporting across its portfolio of print, digital, video, audio, and live event products. Newsday's website, newsday.com, receives approximately 51 million average monthly page views and reaches approximately 2.1 million average monthly unique visitors, with roughly 80% of that digital traffic arriving from mobile devices.

57. Newsday has progressively tightened free access to its content over the past several years, moving from a metered paywall introduced in January 2019, to a complete hard gate as of August 2022, reflecting increased pressure on subscription and advertising revenue.

58. Newsday's Terms of Service, effective October 12, 2023, expressly prohibits using its content "for the development of any software program, including, but not limited to, training a machine learning or artificial intelligence (AI) system," and separately prohibits the use of "robots, spiders, scripts, service, software or any manual or automatic device, tool, or process designed to data mine or scrape the Content, data or information from the Services, or otherwise use, access, or collect the Content, data or information from the Services using automated means."

[14] *Id*. at sections 2(c)-(d).

59. Newsday's own robots.txt file instructs both OpenAI and Common Crawl not to scan or crawl its website, reflecting a publisher-side technical measure restricting the same category of automated access at issue in this action.[15]

60. Plaintiffs maintain rigorous, published editorial standards governing accuracy and attribution that stand in direct contrast to the conduct alleged against Defendants. Their corrections policies require prompt correction of errors and republication of corrections across every platform where the error appeared. Newsday's editorial ethics policy imposes a zero-tolerance plagiarism standard, treating unattributed use of as few as seven to ten consecutive words from an outside source as plagiarism, and requires clear sourcing attribution in every published story. The Seattle Times maintains similarly rigorous standards, and decides whether and how to address an error based on reader respect, journalistic clarity, and common sense.

61. Newsday's internal Artificial Intelligence Usage Policy likewise prohibits any employee from using generative AI to create original content, images, or video for publication under the Newsday brand, and requires human oversight and managerial approval before any AI-assisted work is published. Upon information and belief, Defendants' GenAI products observe no comparable standards when reproducing, paraphrasing, or attributing Plaintiffs' journalism.

B. Defendants' GenAI Products and How They Are Built

62. Generative AI products like Defendants' are built around a large language model, or "LLM," which is software that generates text by predicting which words are statistically likely to come next given the text that precedes them. During training, portions of the training text are masked, and the model attempts to predict the missing portions. Each time it does so, its internal parameters—numeric values that collectively encode what the model has "learned" from the text—

---

[15] Ben Welsh, *Who Blocks OpenAI, Google AI and Common Crawl?*, Palewire, *available at* https://palewi.re/docs/news-homepages/openai-gptbot-robotstxt.html (last accessed Sept. 3, 2026).

are adjusted to make its predictions more accurate. Throughout this process, OpenAI chooses what content to process and which subsets of that content to emphasize by processing them more frequently. Carrying out this process necessarily requires making and retaining copies of the training material in computer memory, feeding those copies through the model repeatedly, and tuning parameters against them, to the point that the model can largely reproduce the training text verbatim.

63. It is well-documented that Defendants' LLMs "memorize" the content that they are trained on and can "regurgitate" it on command. With only minimal prompting, an LLM can output substantial verbatim or near-verbatim excerpts of specific documents it was trained on. That an LLM can reproduce a work on demand shows that a retrievable copy of that work persists in the model's parameters long after training is complete, notwithstanding Defendants' characterization of those parameters as merely reflecting abstract statistical relationships.

64. Defendants' GenAI products are also "grounded" through retrieval-augmented generation ("RAG"), a technique that operates independently of the training process. When a user submits a query, the system searches a stored index or the live internet for information responsive to that query, appends the retrieved text to the user's prompt, and directs the model to generate its answer from that combined material. Because RAG performs its copying after the model has been trained (and often in real time), when a user submits a query, the response can extensively reproduce or closely track the language of a specific work even when that work was never part of the dataset used to train the underlying model in the first place. Moreover, when the GenAI product uses a copyrighted work in RAG, it has the potential to produce myriad responses that are in the style of the material it was trained on, including the Plaintiffs' copyrighted works. That is, of course, the whole point of Defendants' GenAI products—they are "answer engines" whose purpose is to regurgitate, summarize, and mimic the expressive voice of the news reporting and other textual

works they depend on.[16] OpenAI's commercial objectives are driven by keeping users *on* its platform. With copies of the world's content on hand, users have no need to leave their ChatGPT interface to visit seattletimes.com or newsday.com for answers to their news-and-information related queries.

65. ChatGPT model testing confirms that Defendants' GenAI products do, in fact, reproduce Plaintiffs' copyrighted journalism verbatim. Upon information and belief, in many instances, Defendants' models reproduced substantial passages of The Seattle Times and Newsday articles verbatim even when supplied with nothing more than the article's headline, publication date, and URL—confirming that the reproduced text was recalled from the model's training rather than copied from text a user had supplied to it.

66. As just one of many examples, displayed below is a copy of an article from The Seattle Times Pulitzer-winning series covering the 2019 Boeing 737 Max crash and the resulting federal investigation. The Defendants' model regurgitated 88 consecutive words verbatim (pictured in yellow highlights below) with a simple prompt that included the headline and URL, showing that readers determined to evade the Plaintiffs' paywalls need only ask ChatGPT for help.

[16] *See* Ben Wood, *How to Evolve Your Organic Approach for the Rise of Answer Engines*, Search Engine Land (Oct. 9, 2024), *available at* https://searchengineland.com/evolve-organic-approach-answer-engines-447352 (last accessed Sept. 3, 2026).

EVERYTHING SUPPLIED TO THE MODEL	TEXT THE MODEL RETURNED
Seattle Times FBI joining criminal investigation into certification of Boeing 737 MAX 2019-03-20 https://www.seattletimes.com/business/boeing-aerospace/fbi-joining-criminal-investigation-into-certification-of-boeing-737-max/	The FBI has joined the criminal investigation into the certification of the Boeing 737 MAX, lending its considerable resources to an inquiry already being conducted by U.S. Department of Transportation agents, according to people familiar with the matter. The federal grand jury investigation, based in Washington, D.C., is looking into the certification process that approved the safety of the new Boeing plane, two of which have crashed since October. The FBI's Seattle field office lies in proximity to Boeing's 737 manufacturing plant in Renton, Wash. The investigation, which is being overseen by the U.S. Justice Department's criminal division and carried out by the Transportation Department's inspector general, began in response to information obtained after a Lion Air 737 MAX 8 crashed shortly after takeoff from Jakarta on Oct. 29, killing 189 people. The probe is focused on whether Boeing provided incomplete or misleading information about the 737 MAX's stall-prevention system, known as MCAS, to U.S. air-safety regulators and customers, according to

67. The following examples highlighting passages from Newsday's articles in red further demonstrate that Defendants' models retain infringing copies of Plaintiffs' content.

Original Newsday Article Text	Defendants' Model Output
Up to $5 billion of the FEMA money could be used to help local governments — especially Puerto Rico's central government and the island's local governments — remain functional as they endure unsustainable cash shortfalls in the aftermath of Maria, which has choked off revenues and strained resources.	Up to $5 billion of the FEMA money could be used to help local governments … remain functional as they endure unsustainable cash shortfalls in the aftermath of Maria, which has choked off revenues and strained resources.
Last month, the Island's overall employment rose to 1.357 million jobs, from 1.352 million a year earlier. The department uses year-over-year comparisons because local data aren't adjusted to account for seasonal fluctuations in employment.	The number of government jobs on Long Island in December was 230,000, up 1,400 from a year earlier. The department uses year-over-year comparisons because local data aren't adjusted to account for seasonal fluctuations in employment.
Spade created a popular line of sleek handbags in the early 1990s. Kate Spade New York, the company she and her husband, Andy Spade, founded, has more than 140 retail shops and outlet stores across the United States and more than 175 shops internationally.	Kate Spade, the fashion designer who created a line of sleek handbags in the early 1990s that … Spade's company, Kate Spade New York, has more than 140 retail shops and outlet stores across the United States and more than 175 shops internationally, according to its website.
The Labor Department uses year-over-year comparisons because local data aren't adjusted to account for seasonal fluctuations.	The state Labor Department uses year-over-year comparisons because local data aren't adjusted to account for seasonal fluctuations in employment.

1. The Origins of Defendants' Training Datasets

68. OpenAI's GPT models are a family of LLMs first introduced as GPT-1 in 2018, followed by GPT-2 in 2019, GPT-3 in 2020, GPT-3.5 in 2022, GPT-3.5 Turbo and GPT-4 in 2023, with substantially larger and more capable successor models released regularly since, including GPT-4 Turbo, GPT-4o, GPT-4o mini, o1, o1-preview, o1-pro, o1-mini, GPT-4.5, o3, o3-pro, o3-mini, o4-mini, GPT-4.1, GPT-4.1 mini, GPT-4.1 nano, gpt-oss-120b, gpt-oss-20b,, GPT-5, GPT-5 mini, GPT-5 nano, GPT-5-Codex, GPT-5.1, GPT-5.2, GPT-5.2-Codex, GPT-5.3, GPT-5.3-Codex, GPT-5.4, GPT-5.4 mini, GPT-5.4 nano, GPT-5.5, GPT-5.6 Sol, GPT-5.6 Terra, GPT-5.6 Luna, GPT-5.6 Cyber, and GPT-6 Astra. OpenAI's ever-expanding list of products and features includes all "Instant," "Thinking," "mini," "nano," "Pro," and "Codex" variants and successive versions, OpenAI's Daybreak Red and Daybreak Blue models, and any successor models developed using similar methods and trained in a similar manner.

69. OpenAI disclosed meaningful information about the training data underlying GPT-1 and GPT-2, but beginning with GPT-4, OpenAI has refused to disclose the contents, size, or provenance of the datasets used to train its models, stating only that its technical report "contains no further details about the architecture (including model size), hardware, training compute, dataset construction, [or] training method."[17] The dataset OpenAI used to train GPT-2, called "WebText," consists of the text of content scraped from the internet using 45 million links posted to the Reddit social media platform that received at least three "karma" (i.e., users' approval).[18] OpenAI released publicly the top 1000 domains of websites used in the WebText training dataset for GPT-2 and a

[17] OpenAI, GPT-4 Technical Report, arXiv:2303.08774, (Mar. 2023), *available at* https://arxiv.org/abs/2303.08774.
[18] OpenAI, GPT-2 Model Card, *available at* github.com/openai/gpt-2 (last accessed Sept. 3, 2026)).

sample selection of the WebText training dataset itself. The Seattle Times website (seattletimes.com) was among the top 1000 domains used (ranked as #221).

70. WebText's successor dataset, "WebText2," was created using the same Reddit-based methodology but scraped over a longer period of time; OpenAI itself describes WebText2 as "a high-quality" dataset that is "an expanded version of the WebText dataset . . . collected by scraping links over a longer period of time."[19] WebText made up 22% of the training corpus for GPT-3 despite constituting less than 4% of the total number of tokens in that mix, reflecting OpenAI's own view that higher-quality sources—including, upon information and belief, major metropolitan newspapers—were sampled more frequently than lower-quality sources during training.[20]

71. Upon information and belief, OpenAI's LLMs were also trained on a dataset Common Crawl, a large-scale scrape of the internet maintained by a third-party nonprofit organization.[21] Common Crawl's web scraper captures the entire HTML of the webpages it crawls, including a substantial amount of paywalled content, and saves the results in the WARC file format, which preserves a copy of each scraped webpage. Upon information and belief, OpenAI used and retained copies of Common Crawl's WARC files for its training purposes.

72. Upon information and belief, Microsoft provided OpenAI with a copy of the Bing search index, which is Microsoft's own version of Common Crawl, and created bespoke web crawlers for OpenAI's infringing use.

73. Defendants have repeatedly copied Plaintiffs' copyrighted content, without any license and without compensation, in the course of assembling the training datasets their GenAI

---

[19] Tom B. Brown et al., *Language Models Are Few-Shot Learners*, arXiv:2005.14165 (2020), *available at* https://arxiv.org/abs/2005.14165.
[20] *Id*. at Table 2.2.
[21] Common Crawl, Frequently Asked Questions, commoncrawl.org/faq (last accessed Sept. 3, 2026).

products rely on. In training the GPT models, Microsoft and OpenAI collaborated to design and build a complex supercomputing system specifically to house copies of the training data. Upon information and belief, Plaintiffs' works were copied and ingested multiple times for the purpose of training successive generations of Defendants' GPT models.

74. Defendants' products continue to grow and iterate on the backs of Plaintiffs' copyrighted content. OpenAI's latest plans and releases include the introduction of advertising as part of OpenAI's free-to-the-public products, ChatGPT Free and Go. This past August, fewer than 200 days after launching ads in the U.S., OpenAI announced that its advertising business had reached a $1 billion annualized revenue run rate—a milestone the company touted as proof of a "diversified business model" ahead of its anticipated IPO.[22]

75. The negative flywheel effect will cause further harm to publishers—in our modern "eyeball" economy, where advertisers compete for attention across a wide range of media, Defendants products draw users to their platforms, spit out copies of Plaintiffs' content, and, in doing so, siphon advertising dollars away from Plaintiffs.

2. Web Crawlers, Scraping, and Robots.txt

76. OpenAI operates three primary automated web crawlers. "GPTBot" is used to gather content that, OpenAI publicly states, "may be used in training our generative AI foundation models." "OAI-SearchBot" is used to surface content in ChatGPT's search features.[23] ChatGPT-User visits web pages when a user makes an inquiry that may require live information.

77. A website operator can instruct any bot, including OpenAI's three main bots, not to crawl a site by adding a disallow directive to the site's robots.txt file. A robots.txt file is a decades-

---

[22] *See* Ashley Capoot, *OpenAI's Ad Business Hits $1 Billion Annualized Revenue Run Rate*, CNBC (Aug. 31, 2026), *available at* https://www.cnbc.com/2026/08/31/open-ai-chatgpt-ads-revenue.html (last accessed Sept. 3, 2026).
[23] OpenAI, *Overview of OpenAI Crawlers*, developers.openai.com/api/docs/bots (last accessed Sept. 3, 2026).

old, voluntary technical convention following the Robots Exclusion Protocol ("REP") that many website owners use to signal that they do not want an unwanted bot to access their site.

78. OpenAI did not disclose any bots used to perform earlier scrapes, and did not even begin to attempt to support REP disallows until August 2023—years after GPT-1, GPT-2, and GPT-3 had already been trained on a broad, largely indiscriminate scraping of the internet and a substantial number of the world's copyrighted works.[24]

3. Outputs: Search, RAG, and Regurgitation

79. Defendants' GenAI products produce infringing copies of Plaintiffs' content during RAG and in resulting output to users. When Defendants' GenAI products use a Seattle Times or Newsday article in RAG or for output, the resulting text does not add any new purpose or character to the underlying journalism. Rather, the use in RAG or output merely repackages the original reporting into a competing form. Because a majority of ChatGPT users are unlikely to click through to read a full article once they have received a ChatGPT version of it, such use directly substitutes for, rather than complements, Plaintiffs copyrighted works.[25]

4. CMI Removal

80. The Seattle Times and Newsday convey CMI in connection with their published content. The Seattle Times conveys a copyright notice on each page of its website—at the bottom of each article—which links to the copyright notice page stating that all content "is owned by The Seattle Times, its licensors, and/or its content providers" and "may not be copied or retransmitted via any means." The Seattle Times also conveys a copyright notice in its Terms of Service, stating,

---

[24] Ariel Bogle, New York Times, *CNN and Australia's ABC Block OpenAI's GPTBot Web Crawler From Accessing Content*, The Guardian (Aug. 24, 2023, 8:31 PM EDT), *available at* https://www.theguardian.com/technology/2023/aug/25/new-york-times-cnn-and-abc-block-openais-gptbot-web-crawler-from-scraping-content (last accessed Sept. 3, 2026).

[25] Pew Rsch. Ctr., *Google Users Are Less Likely to Click on Links When an AI Summary Appears in the Results*, (July 22, 2025), *available at* https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/ (last accessed Sept. 3, 2026).

"All Site design, graphics, text, text selections, arrangements, and all software are copyrighted by Seattle Times Company or its licensors. ALL RIGHTS RESERVED."[26]

81. Newsday similarly includes a copyright notice on each page of its website, including at the bottom of each article, as well as in its Terms of Service, stating: "All information, content, services and software displayed on, transmitted through, or used in connection with the site … including for example news articles, reviews, directories, guides, text, photographs, images, illustrations, audio clips, video, html, source and object code, trademarks, logos, and the like … is owned by Newsday, and/or its affiliated companies, licensors and suppliers."[27]

82. In building the dataset, OpenAI has admitted to using two content-extraction algorithms, "Dragnet" and "Newspaper," to separate an article's main article content from the surrounding webpage.[28] Upon information and belief, OpenAI's application of these same or substantially similar extraction methods to Plaintiffs' content—as OpenAI is alleged to have done to the websites of other news publishers, in related litigation—removed CMI, including copyright notices, author, title, and terms-of-use information, from copies of Plaintiffs' works incorporated into OpenAI's training datasets.

83. OpenAI's removal of CMI was intentional and was undertaken with knowledge, or reasonable grounds to know, that doing so would induce, enable, facilitate, or conceal infringement by Defendants and by end-users of Defendants' GenAI products. Upon information and belief, OpenAI did not disclose to news media publishers, including Plaintiffs, that it had surreptitiously copied hundreds of thousands of their articles and essentially made them available to the public

---

[26] Seattle Times, Copyright, https://www.seattletimes.com/notices/copyright.html (last visited Sept. 3, 2026); Seattle Times, Terms of Service § 35 (Copyright Notice), https://www.seattletimes.com/notices/terms.html (last visited Sept. 3, 2026).
[27] Newsday, Terms of Service, https://www.newsday.com/services/terms-of-service-qymn88cv (last visited Sept. 3, 2026).
[28] Alec Radford et al., *Language Models Are Unsupervised Multitask Learners*, OpenAI (2019), *available at* https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

without credit, compensation, or even a blue link to encourage users to click-through to Plaintiffs' websites.

84. Upon information and belief, Microsoft likewise created copies of Plaintiffs' works with CMI removed in connection with training and operating its Copilot and Bing Chat products.

85. Upon information and belief, Microsoft and OpenAI shared with each other copies of Plaintiffs' works with CMI removed in connection with their collaborative development of generative AI products, and distributed copies of Plaintiffs' works with CMI removed to end users. OpenAI, in particular, distributes those copies by purporting to assign users all rights in the output under its applicable terms of service.

C. Hallucination and False Attribution

86. At the same time that Defendants' GenAI products are copying, reproducing, and paraphrasing Plaintiffs' content without consent or compensation, Defendants' products are also causing Plaintiffs commercial and competitive injury when they "hallucinate" content that the product misattributes to Plaintiffs that is not Plaintiffs' content, or falsely attributes Plaintiffs' content to another party. In such instances (or any variation thereof), Plaintiffs and their trusted brands are tarnished by the misinformation.

D. Defendants' Continuing Copying and Ongoing Infringement

87. Upon information and belief, Defendants' GenAI products have made, and continue to make, new copies of Plaintiffs' current content in connection with RAG-based search functionality. First, Defendants conduct broad web crawls to create new indices of the internet that can be used in RAG, and, upon information and belief, these crawls retrieve and store Plaintiffs' journalistic content. Second, Defendants' GenAI products retrieve and reproduce Plaintiffs' reporting in response to user queries about current events through direct scrapes of Plaintiffs' websites, without Plaintiffs' authorization and without preserving CMI.

E. Willfulness and Notice

88. Defendants' infringement is willful. Defendants have been repeatedly placed on notice, through widely publicized litigation against OpenAI and Microsoft by other news publishers beginning in December 2023, that their conduct in continuing to acquire, train on, and reproduce copyrighted journalism without permission infringes copyright and violates the DMCA.[29] Defendants have barreled ahead nonetheless, continuing to create infringing models and products based on and incorporating Plaintiffs' copyrighted works.

89. OpenAI has entered into licensing agreements with over a dozen other news and media organizations, including, on information and belief, the Associated Press, NewsCorp, Axios, Axel Springer, The Atlantic, Financial Times, Dotdash Meredith, and Vox Media, demonstrating OpenAI's own recognition that a license is required to use copyrighted journalism in its GenAI products—a license Defendants never sought from, or entered into with, Plaintiffs.[30] According to public reports, the disclosed terms of just three of these agreements reflect that OpenAI has paid more than $300 million for the right to use news organizations content for its products, with the terms of the remaining agreements undisclosed.

F. Harm to Plaintiffs

90. Defendants' conduct has caused, and continues to cause substantial and multi-faceted harm to Plaintiffs' specific business models. By generating complete, text-based responses that answer a reader's question directly in various stylistic voices that includes, reliance on Plaintiffs' content, without requiring the reader to visit seattletimes.com or newsday.com,

---

[29] *N.Y. Times Co. v. Microsoft Corp. & OpenAI, Inc.*, No. 23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023); *Ziff Davis, Inc. v. OpenAI, Inc.*, No. 25-cv-04315 (S.D.N.Y. filed May 22, 2025); *U.S. News & World Report, L.P. v. OpenAI, Inc.*, No. 25-cv-09912 (S.D.N.Y. filed Nov. 26, 2025); *Ctr. for Investigative Reporting, Inc. v. OpenAI, Inc.*, No. 24-cv-04872 (S.D.N.Y. filed June 27, 2024).

[30] Rebecca Bellan, *The New York Times and Amazon Ink AI Licensing Deal*, TechCrunch (May 29, 2025), *available at* https://techcrunch.com/2025/05/29/the-new-york-times-and-amazon-ink-ai-licensing-deal/ (last accessed Sept. 3, 2026).

Defendants' products reduce the number of readers who click through to Plaintiffs' own websites, directly reducing the referral traffic on which Plaintiffs' digital advertising revenue depends.[31] According to industry data from December 2024 to December 2025, search referral traffic to mid-sized regional and metro daily publishers, like Newsday and The Seattle Times, declined by approximately 47%, a materially steeper decline than the approximately 22% experienced by larger national publishers, because mid-sized publishers have historically depended more heavily on incidental search traffic that AI-generated answers now keep for themselves.[32]

91. Also, because Plaintiffs' revenue model depends on converting readers into paying digital subscribers, and because a reader who can obtain a satisfactory answer to a news question from ChatGPT or Copilot has less incentive ever to subscribe, Defendants' conduct suppresses new subscription growth and accelerates subscriber churn. The harm will only worsen as advertisers convert their dollars to ChatGPT's Free and Go products and away from traditional advertising platforms like those of Plaintiffs.

92. A well-established market exists for licensing professional journalism for use in training and operating GenAI systems, as evidenced by Defendants' own licensing agreements with other publishers. Countless news and digital media agreements have been reported by third-party sources, forming a market that is already valued in the billions. There is no justification for Defendants' wholesale taking and use of Plaintiffs' content without payment or permission. Defendants' unauthorized use of Plaintiffs' content deprived Plaintiffs of the opportunity to license its content on fair terms and to obtain just compensation.

---

[31] Pew Rsch. Ctr., supra note 25; Sara Guaglione, *In Graphic Detail: AI Platforms Are Driving More Traffic — But Not Enough To Offset "Zero-Click" Search*, Digiday (July 10, 2025), *available at* https://digiday.com/media/in-graphic-detail-ai-platforms-are-driving-more-traffic-but-not-enough-to-offset-zero-click-search/ (last accessed Sept. 3, 2026).

[32] *See* Danny Goodwin, *Small Publisher Search Traffic Fell 60% Over Two Years: Data*, Search Engine Land (Mar. 19, 2026), *available at* https://searchengineland.com/small-publisher-search-traffic-drops-data-471974 (last accessed Sept. 3, 2026).

COUNT I -- COPYRIGHT INFRINGEMENT (17 U.S.C. Section 501)

93. Plaintiffs incorporate by reference the preceding paragraphs as though fully set forth herein.

94. Plaintiffs are the owners of the exclusive rights under 17 U.S.C. Section 106 in their respective Registered Works.

95. By reproducing, storing, and processing copies of Plaintiffs' works to build training datasets and to train, fine-tune, and ground Defendants' GenAI models, and by disseminating generative output containing copies and derivatives of Plaintiffs' works, Defendants have directly infringed Plaintiffs' exclusive rights in their copyrighted works.

96. On information and belief, Defendants' infringement was and continues to be willful and undertaken with full knowledge of Plaintiffs' rights.

97. As a direct and proximate result of Defendants' infringement, Plaintiffs have suffered and will continue to suffer substantial and irreparable injury, for which there is no adequate remedy at law, entitling Plaintiffs to injunctive relief in addition to damages.

COUNT II -- VICARIOUS COPYRIGHT INFRINGEMENT (Against Microsoft, OpenAI Inc., OpenAI GP, OpenAI LP, OAI Corporation, OpenAIHoldings LLC, and OpenAI Global LLC)

98. Plaintiffs incorporate by reference the preceding paragraphs as though fully set forth herein.

99. Microsoft directed, controlled, and profited from the infringement by the OpenAI defendants. Microsoft had the right and ability to control the processes and computing infrastructure used by the Defendants to reproduce Plaintiffs' works. Microsoft provided the massive computing platform used to store, process, and reproduce the training datasets containing Plaintiffs' works into its own commercial product offerings, including Copilot and Bing Chat.

100. Defendants OpenAI Inc., OpenAI GP, LLC, OAI Corporation, OpenAI Holdings LLC, and Microsoft directed, controlled, and profited from the infringing conduct perpetrated by the other OpenAI entities, including the reproduction and distribution of Plaintiffs' works.

101. Defendants OpenAI Global LLC and OpenAI LP directed, controlled, and profited from the infringement perpetrated by Defendants OpenAI OpCo LLC and Open AI, LLC, including the reproduction and distribution of Plaintiffs' works.

102. Defendants Open AI Inc., OpenAI LP, OAI Corporation, Open AI Holdings, LLC, OpenAI Global LLC, and Microsoft are vicariously liable for copyright infringement.

COUNT III -- VIOLATION OF THE DMCA, 17 U.S.C. Section 1202(b)(1) (REMOVAL OF COPYRIGHT MANAGEMENT INFORMATION)

103. Plaintiffs incorporate by reference the preceding paragraphs as though fully set forth herein.

104. Plaintiffs' works conveyed CMI, including copyright notices, author and title information, and terms-of-use information.

105. Without Plaintiffs' authority, Defendants removed CMI from Plaintiffs' works in the course of building training datasets, training and operating their GenAI models, and generating output containing copies or derivatives of Plaintiffs' works.

106. Defendants knew that Plaintiffs' works contained CMI, and knowingly removed such CMI.

107. Defendants removed such CMI knowing, or having reasonable grounds to know, that doing so would induce, enable, facilitate, or conceal infringement of Plaintiffs' copyrights.

108. Plaintiffs have been injured by Defendants' removal of CMI and are entitled to statutory damages, actual damages, restitution of profits, and attorneys' fees and costs.

COUNT IV -- VIOLATION OF THE DMCA, 17 U.S.C. Section 1202(b)(3)
(DISTRIBUTION OF WORKS WITH CMI REMOVED)

109. Plaintiffs incorporate by reference the preceding paragraphs as though fully set forth herein.

110. Defendants distributed, and continue to distribute, to users of its products, copies of Plaintiffs' works, and generative output derived from Plaintiffs' works, knowing that CMI had been removed from those works without Plaintiffs' authority, and knowing or having reasonable grounds to know that such distribution would induce, enable, facilitate, or conceal infringement.

111. On information and belief, OpenAI and Microsoft each shared copies of Plaintiffs' works, with CMI removed, with the other in connection with the collaborative development of their respective GenAI products.

COUNT V -- TRADEMARK DILUTION (15 U.S.C. Section 1125(c))

112. Plaintiffs incorporate by reference the preceding paragraphs as though fully set forth herein.

113. Plaintiffs are the respective owners of the following federally registered word marks (together, the "Plaintiffs' Marks"): (a) THE SEATTLE TIMES, U.S. Registration Nos. 5,738,989 (Int'l Class 9), 5,738,990 (Int'l Class 16), 5,738,991 (Int'l Class 35), and 5,738,992 (Int'l Class 41), each registered on the Principal Register on April 30, 2019, and each now incontestable under 15 U.S.C. Section 1065; and (b) NEWSDAY, U.S. Registration No. 2,047,787 (Int'l Class 16), registered on the Principal Register on March 25, 1997, renewed, and incontestable under 15 U.S.C. Section 1065.

114. Plaintiffs' Marks are famous within the meaning of 15 U.S.C. Section 1125(c)(2)(A) and became famous before Defendants' first use in commerce of their diluting GenAI products. Plaintiffs' Marks are widely recognized by the general consuming public of the

U.S., and by the news-media consuming public in particular, as designating the source of Plaintiffs' journalism. Plaintiffs Marks are famous under the non-exclusive statutory factors under 15 U.S.C. Section 1125(c)(2)(A):

115. The duration, extent, and geographic reach of advertising and publicity of Plaintiffs' Marks: they have been used continuously and prominently since at least 1896 (THE SEATTLE TIMES) and 1940 (NEWSDAY).

116. The amount, volume, and geographic extent of sales of goods and services offered under Plaintiffs' Marks: the publications operating under Plaintiffs' Marks have reached a combined audience of several million readers monthly.

117. The extent of actual recognition of Plaintiffs' Marks: this is reflected in a combined 30 Pulitzer Prizes and extensive independent press recognition of Plaintiffs' journalism.

118. Whether Plaintiffs' Marks are registered on the Principal Register: each of Plaintiffs' Marks is.

119. Defendants' GenAI products reproduce, output, and associate Plaintiffs' Marks with content generated by Defendants' own models, including content that is not Plaintiffs' authentic journalism, thereby creating an association between Plaintiffs' Marks and Defendants' commercial GenAI products that is likely to impair the distinctiveness of Plaintiffs' Marks in violation of 15 U.S.C. Section 1125(c)(2)(B). Defendants' conduct dilutes Plaintiffs' Marks by blurring because it whittles away the exclusive, source-identifying association the public has long held between Plaintiffs' Marks and Plaintiffs' journalism, associating those Marks instead with Defendants' AI-generated substitute content over which Plaintiffs exercise no editorial control.

120. Separately and independently, Defendants' conduct dilutes Plaintiffs' Marks by tarnishment in violation of 15 U.S.C. Section 1125(c)(2)(C). As alleged above, Defendants' GenAI products "hallucinate" content that they misattribute to Plaintiffs, or falsely attribute Plaintiffs'

actual content to another source, associating Plaintiffs' Marks with inaccurate, fabricated, or substandard content that Plaintiffs did not create, did not review, and would not publish under their own rigorous editorial and accuracy standards. This association harms the reputation of Plaintiffs' Marks by connecting them, in the minds of consumers, with unreliable or false information—the direct opposite of the accuracy, integrity, and trustworthiness for which Plaintiffs' Marks are known.

121. Upon information and belief, Defendants' dilutive uses of Plaintiffs' Marks began, and have continued, after Defendants were placed on actual and constructive notice of Plaintiffs' rights, including through the widely publicized litigation identified in paragraph 83 above, rendering Defendants' dilutive conduct willful within the meaning of 15 U.S.C. Section 1125(c)(5).

122. As a direct and proximate result of Defendants' dilution of Plaintiffs' Marks, Plaintiffs have suffered, and will continue to suffer, injury to the distinctiveness and reputation of Plaintiffs' Marks for which there is no adequate remedy at law, entitling Plaintiffs to injunctive relief under 15 U.S.C. Section 1125(c)(1), and, because Defendants' dilutive use was willful, to the additional remedies available under 15 U.S.C. Sections 1125(c)(5) and 1117(a), including an accounting of Defendants' profits, damages, costs, and, in exceptional cases, reasonable attorneys' fees.

COUNT VI -- TRADEMARK DILUTION UNDER RCW 19.77.160
(ON BEHALF OF THE SEATTLE TIMES)

123. The Seattle Times incorporates by reference the preceding paragraphs as though fully set forth herein.

124. The Seattle Times owns The Seattle Times Marks.

125. The Seattle Times Marks are famous in the state of Washington within the meaning of RCW 19.77.160 and became famous before Defendants' first use in commerce of their diluting

GenAI products. The Seattle Times Marks are widely recognized by the general consuming public of Washington state, and by the news-media consuming public of Washington in particular, as designating the source of Plaintiffs' journalism. The Seattle Times Marks are famous under the non-exclusive statutory factors under RCW 19.77.160:

126. The duration and extent of the mark in connection with the goods or services with which the mark is used: The Seattle Times Marks have been used continuously and prominently since at least 1896 and has reached an audience of over 1 million readers monthly.

127. The geographical extent of the trading area in which the mark is used: the publications operating under The Seattle Times Marks have been used continuously and prominently throughout Washington state.

128. The degree of recognition of the mark in the trading areas and channels of trade in this state used by the mark's owner and the person against whom the injunction is sought: this is reflected in the 11 Pulitzer Prizes and extensive independent press recognition of The Seattle Times journalism.

129. Whether the mark is the subject of state registration in this state or United States registration: The Seattle Times Marks were registered on the Principal Register on April 30, 2019.

130. Defendants' GenAI unauthorized commercial products, available in Washington state, reproduce, output, and associate The Seattle Times Marks with content generated by Defendants' own models, including content that is not The Seattle Times' authentic journalism, thereby creating an association between The Seattle Times Marks and Defendants' commercial GenAI products that is likely to impair the distinctiveness of The Seattle Times Marks in violation of RCW 19.77.160.

131. Separately and independently, Defendants' conduct dilutes The Seattle Times Marks by tarnishment as Defendants' GenAI products "hallucinate" content that they misattribute

to The Seattle Times, or falsely attribute Plaintiff's actual content to another source, associating The Seattle Times Marks with inaccurate, fabricated, or substandard content that Plaintiff did not create, did not review, and would not publish under their own rigorous editorial and accuracy standards.

132. Upon information and belief, Defendants' dilutive uses of The Seattle Times Marks began, and have continued, after Defendants were placed on actual and constructive notice of The Seattle Times' rights, including through the widely publicized litigation identified in paragraph 83 above, rendering Defendants' dilutive conduct in bad faith within the meaning of RCW 19.77.150.

133. As a direct and proximate result of Defendants' dilution of The Seattle Times Marks, The Seattle Times has suffered, and will continue to suffer, injury to the distinctiveness and reputation of its marks for which there is no adequate remedy at law, entitling The Seattle Times to injunctive relief under RCW 19.77.150, and, because Defendants' dilutive use was in bad faith, to the additional remedies available under RCW 19.77.150, including up to three times profits and damages and reasonable attorneys' fees.

COUNT VII – TRADEMARK DILUTION UNDER NY Gen. Bus. L. § 360-L
(ON BEHALF OF NEWSDAY)

134. Newsday incorporates by reference the preceding paragraphs as though fully set forth herein.

135. Newsday owns the Newsday's Mark.

136. The Newsday Mark is distinctive through acquiring secondary meaning within the meaning of NY Gen. Bus. L. § 360-L and became distinctive before Defendants' first use in commerce of their diluting GenAI products. Newsday's Mark is widely recognized by the general consuming public of New York, and by the news-media consuming public in particular, as

designating the source of Plaintiffs' journalism. Newsday's Mark acquired secondary meaning in the minds of the consuming public:

137. The Newsday Mark has been used continuously and prominently since at least 1940.

138. The publications operating under the Newsday Mark have reached a combined audience of several million readers monthly.

139. The extent of actual recognition of the Newsday Mark: this is reflected in the 19 Pulitzer Prizes and extensive independent press recognition of Newsday's journalism.

140. The Newsday Mark was registered on the Principal Register on March 25, 1997.

141. Defendants' GenAI unauthorized commercial products reproduce, output, and associate the Newsday Mark with content generated by Defendants' own models, including content that is not Plaintiff's authentic journalism, thereby creating an association between the Newsday Mark and Defendants' commercial GenAI products that is likely to impair the distinctiveness of the Newsday Mark in violation of NY Gen. Bus. L. § 360-L.

142. Separately and independently, Defendants' conduct dilutes the Newsday Mark by tarnishment as Defendants' GenAI products "hallucinate" content that they misattribute to Newsday, or falsely attribute Plaintiff's actual content to another source, associating the Newsday Mark with inaccurate, fabricated, or substandard content that Newsday did not create, did not review, and would not publish under their own rigorous editorial and accuracy standards.

143. Upon information and belief, Defendants' dilutive uses of the Newsday Mark began, and have continued, after Defendants were placed on actual and constructive notice of Newsday's rights, including through the widely publicized litigation identified in paragraph 83 above, rendering Defendants' dilutive conduct in bad faith within the meaning of NY Gen. Bus. L. § 360-L.

144. As a direct and proximate result of Defendants' dilution of the Newsday Mark, Newsday has suffered, and will continue to suffer, injury to the distinctiveness and reputation of its marks for which there is no adequate remedy at law, entitling Newsday to injunctive relief under NY Gen. Bus. L. § 360-L, and, because Defendants' dilutive use was with knowledge and in bad faith, to the additional remedies available under NY Gen. Bus. L. § 360-L, including up to three times profits and damages and reasonable attorneys' fees.

PRAYER FOR RELIEF

WHEREFORE, Plaintiffs respectfully request that this Court enter judgment against Defendants as follows:

(a) Awarding Plaintiffs statutory damages, or actual damages and Defendants' profits, as elected by Plaintiffs pursuant to 17 U.S.C. Section 504;

(b) Awarding Plaintiffs statutory damages for each violation of 17 U.S.C. Section 1202, or actual damages and Defendants' profits, as elected by Plaintiffs;

(c) Permanently enjoining Defendants from further unlawful, unfair, and infringing conduct as alleged herein;

(d) Ordering the impoundment and/or destruction, pursuant to 17 U.S.C. Section 503, of all copies of Plaintiffs' works, and all LLMs and training datasets incorporating Plaintiffs' works or derivatives thereof, in Defendants' possession, custody, or control;

(e) Enjoining Defendants from further dilutive uses of Plaintiffs' Marks pursuant to 15 U.S.C. Section 1125(c)(1), RCW 19.77.160, and NY Gen. Bus. L. § 360-L, and awarding Plaintiffs an accounting of Defendants' profits, damages, and, in light of Defendants' bad faith and willful conduct, treble damages and reasonable attorneys' fees pursuant to 15 U.S.C. Sections 1117(a), 1125(c)(5), RCW 19.77.160, and NY Gen. Bus. L. § 360-L;

(f) Awarding Plaintiffs their costs, expenses, and reasonable attorneys' fees as permitted by law; and

(g) Granting such other and further relief as the Court deems just and proper.

DEMAND FOR JURY TRIAL

Plaintiffs hereby demand a trial by jury on all issues so triable in accordance with Federal Rule of Civil Procedure 38(b).

DATED: September 4, 2026

Respectfully submitted,

/s/ Lacy H. Koonce, III______

Lacy H. ("Lance") Koonce, III
Brendan T. Kehoe
KLARIS LAW PLLC
161 Water Street
New York, NY 10038
Telephone: (646) 779-4882
lance.koonce@klaris.com
brendan.kehoe@klarislaw.com

Attorneys for Plaintiffs The Seattle Times Company and Newsday LLC