UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

<table>
<tr>
<td>

NICHOLAS A. BASBANES and NICHOLAS NGAGOYEANES (professionally known as Nicholas Gage), individually and on behalf of all others similarly situated,

                    Plaintiffs,

                -against-

MICROSOFT CORPORATION, OPENAI, INC., OPENAI GP, L.L.C., OPENAI HOLDINGS, LLC, OAI CORPORATION, LLC, OPENAI GLOBAL, LLC, OPENAI, L.L.C., and OPENAI OPCO, LLC,

                    Defendants.

</td>
<td>

Civil Action No. 1:24-cv-84

CLASS ACTION COMPLAINT

JURY TRIAL DEMANDED

</td>
</tr>
</table>

Plaintiffs Nicholas A. Basbanes ("Mr. Basbanes") and Nicholas Ngagoyeanes (professionally known as Nicholas Gage, "Mr. Gage"), on behalf of themselves and all others similarly situated (the "Plaintiffs"), for their complaint against Defendants Microsoft Corporation ("Microsoft") and OpenAI, Inc., OpenAI GP, L.L.C., OpenAI Holdings, LLC, OAI Corporation, LLC, OpenAI Global, LLC, OpenAI, L.L.C. and OpenAI OpCo, LLC (collectively "OpenAI" and, with Microsoft, "Defendants"), allege as follows:

INTRODUCTION

1.   Many of our greatest works of nonfiction are crafted by journalists. Using their skills for research, investigation and interviewing, they generate new knowledge and weave together complex and compelling narratives that require years of devotion to perfect.

2.   Plaintiffs Nicholas A. Basbanes and Nicholas Gage are both journalists who used their skills to write some of the most compelling and celebrated works of nonfiction in the past

40 years. The impact of their work on society cannot be overstated.

3. Mr. Basbanes has been recognized by the National Endowment for the Humanities as a Public Scholar and is an internationally renowned authority on the history of books and book culture.

4. Mr. Gage's body of work is so compelling and continuingly relevant that President Ronald Reagan cited it as an inspiration for his summit meetings with the Soviet Union to end the arms race.

5. Yet Defendants, through their willful and flagrant violations of Plaintiffs' copyrights, threaten the very existence of writers because without permission or payment, Defendants copied Plaintiffs' work to build a massive commercial enterprise that is now valued at billions of dollars.

6. This commercial enterprise, commonly known as "AI" or "Artificial Intelligence," relies on "large language models" or "LLMs." These LLMs generate human-like responses to prompts submitted by users, and work by ingesting massive amounts of written material. The higher the quality of the ingested written materials, the higher the quality of the LLMs responses to users' prompts. Thus, to obtain that high-quality written material, Defendants engaged in a massive and deliberate theft of copyrighted works created by writers like Mr. Basbanes and Mr. Gage.

7. Professional writers like Mr. Basbanes and Mr. Gage have limited capital to fund their research. They typically self-fund their projects, which are often supplemented by advances on royalties from prospective publishers, and those advances are then deducted from any future earnings their books may generate. Working thusly on their own resources, they travel wherever they must to locate and secure access to primary research materials that have been previously

unexamined and to conduct essential interviews. If they find it necessary to use certain copyrighted materials in their published works, it is the authors—not their publishers—who pay for the privilege of doing so. Yet, in stark contrast, Defendants, with ready access to billions in capital, simply stole Plaintiffs' copyrighted works to build another billion+ dollar commercial industry. It is outrageous.

8.   It bears emphasis that Defendants, as sophisticated commercial entities, clearly decided upon a deliberate strategy to steal Plaintiffs' copyrighted works to power their massive commercial enterprise. Of course, not paying for the inputs that make their LLMs, and which are thus plainly derivative works, results in an even higher profit margin for Defendants.

9.   Any concerns Defendants may have had that paying to use copyrighted works would have been cost prohibitive, and possibly preclude the development of this nascent industry, would be absurd. First, even if true, such concerns do not give Defendants license to steal Plaintiffs' copyrighted works. Second, Defendants clearly could have obtained the capital to pay given the extraordinary investments already made and the staggering valuations now associated with these LLMs. Finally, Defendants also could have explored financing alternatives, such as profit sharing or other mechanisms to facilitate their development of this new technology without sapping initial resources. But instead, Defendants just decided to steal. They're no different than any other thief.

10.   Notably, shortly after The New York Times filed suit against these same defendants in this Court, the Defendants publicly acknowledged that copyright owners like Plaintiffs must be compensated for Defendants' use of their work: "We respect the rights of content creators and owners and are committed to working with them to ensure they benefit from

3

AI technology and new revenue models."[1]

11.   Thus, Mr. Basbanes and Mr. Gage seek to represent a class of writers whose copyrighted work has been systematically pilfered by Defendants. Mr. Basbanes and Mr. Gage seek damages for copyright infringement, the lost opportunity to license their works, and for the destruction of the market Defendants have caused and continue to cause to writers. They also seek a permanent injunction to prevent these harms from recurring.

12.   Mr. Basbanes and Mr. Gage complain of Defendants, on personal knowledge as to matters relating to themselves, and on information and belief based on their counsel's reasonable investigation as to all other matters.

JURISDICTION AND VENUE

13.   The Court has jurisdiction over the subject matter of this action under 28 U.S.C. §§ 1331 and 1338(a) because the action arises under the Copyright Act.

14.   The Court also has personal jurisdiction over Defendants because they have purposely availed themselves of the privilege of conducting business in New York.

15.   Microsoft personnel based at its offices in New York assisted with OpenAI's copyright infringement, including exploiting training datasets that relied on copyrighted works.

16.   OpenAI has marketed, sold and distributed, and continues to market, sell and distribute, its infringing products like ChatGPT, ChatGPT Enterprise, ChatGPT Plus, Browse with Bing, and application programming interface tools, within New York and to New York residents and New York based businesses.

17.   Venue is proper in this District under 28 U.S.C. § 1391(b)(2) because a

---

[1] Law360, What To Know About The NYT Suit Against Microsoft, OpenAI, January 2, 2024, available at https://www.law360.com/technology/articles/1781130/what-to-know-about-the-nyt-suit-against-microsoft-openai (last accessed Jan. 5, 2024).

substantial part of the events giving rise to Plaintiffs' claims occurred here.

18. Venue is also proper in this District under 28 U.S.C. § 1400(a) because Defendants or their agents reside or may be found here.

PARTIES

I. Plaintiffs

19. Plaintiff Nicholas A. Basbanes has authored numerous works of nonfiction. He was a reporter, literary editor, and columnist for *The Worcester Telegram & Gazette* and is a resident of North Grafton, Massachusetts. His copyrighted works that Defendants stole required research, investigation and interviews in New York, and these works have been sold to institutions and individuals in New York. In addition, Mr. Basbanes' copyrighted works were published by publishing houses based in New York.

20. Plaintiff Nicholas Gage has authored numerous works of nonfiction. He was an investigative reporter for *The New York Times* and *The Wall Street Journal* and is a resident of North Grafton, Massachusetts. His copyrighted works that Defendants stole required research, investigation and interviews in New York, and these works have been sold to institutions and individuals in New York. In addition, Mr. Gage's copyrighted works were published by publishing houses based in New York.

21. A list of Plaintiffs' copyrighted works infringed by Defendants is set forth on Exhibit A to this Complaint.

II. Defendants

22. Defendant Microsoft Corporation is a Washington corporation with its principal place of business in Redmond, Washington. It is estimated to own 49 percent of Defendant

OpenAI Global, LLC and has invested over $13 billion in OpenAI Global, LLC.[2]

23.   The OpenAI Defendants are a series of interrelated Delaware entities.

24.   Defendant OpenAI, Inc. is a Delaware nonprofit corporation with its principal place of business in San Francisco, California. It was formed in 2015 and owns and controls all other OpenAI entities.

25.   Defendant OpenAI GP, L.L.C. is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. It is wholly owned and controlled by OpenAI, Inc. OpenAI GP, L.L.C. is the vehicle through which OpenAI, Inc. controls OpenAI Global, LLC.

26.   Defendant OpenAI Holdings, LLC is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. The members of OpenAI Holdings, LLC are Defendant OpenAI, Inc. and Aestas LLC, an OpenAI-related limited liability company that is not a defendant here, but which was created to facilitate a capital raise for OpenAI of approximately a half-billion dollars. OpenAI Holdings, LLC was involved in the copyright infringement alleged in this Complaint through its direction, control and ownership of OpenAI OpCo, LLC.

27.   Defendant OAI Corporation, LLC is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. OAI Corporation, LLC's only member is Defendant OpenAI Holdings, LLC and it was involved in the copyright infringement alleged in this Complaint through its direction, control and ownership of OpenAI Global, LLC.

---

[2] Julia Angwin, The OpenAI Coup Is Great for Microsoft. What Does It Mean for Us?, New York Times (Nov. 21, 2023), *available at* https://www.nytimes.com/2023/11/21/opinion/thesam-altman-openai-board-microsoft.html (last accessed Jan. 2, 2024).

28.   Defendant OpenAI Global, LLC is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. OpenAI Global, LLC's members are Defendants Microsoft and OAI Corporation, LLC. OpenAI Global, LLC is controlled by OpenAI GP, L.L.C., which in turn is wholly owned and controlled by Open AI, Inc. OpenAI Global, LLC was involved in the copyright infringement alleged in this Complaint through its direction, control and ownership of OpenAI L.L.C.

29.   Defendant OpenAI L.L.C. is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. OpenAI L.L.C. owns some or all of the services and products provided by OpenAI and has thus monetized the copyrighted works of Plaintiffs without permission and without compensation. The sole member of OpenAI L.L.C. is Defendant OpenAI OpCo, LLC.

30.   Defendant OpenAI OpCo, LLC is a limited liability company formed under the laws of Delaware with its principal place of business in San Francisco, California. OpenAI OpCo, LLC was previously known as OpenAI LP, which was founded in 2019 to be the profit-making arm of OpenAI. OpenAI OpCo, LLC now holds that role.

III.   **Microsoft's New York Based Theft of Plaintiffs' Copyrighted Works**

31.   Microsoft's New York City-based operations directly participated in and supported the copyright infringement alleged in this Complaint by utilizing "a wide variety of topics within theoretical and applied machine learning, including learning from interactive data (e.g., contextual bandits and reinforcement learning), online learning, natural language processing, and topics related to interpretability and fairness of [machine learning] and [artificial

intelligence].”[3]

32.   Microsoft's Azure project, which collaborated heavily with OpenAI in its LLM "training,"[4] does significant work in Microsoft's New York City office.[5]

33.   Microsoft also works in close partnership with OpenAI in connection with the LLMs at issue in this lawsuit. An OpenAI LLM underlies Microsoft's Bing Chat product, offered through its Bing search engine. Microsoft has also integrated OpenAI's LLMs into its sales and marketing software, coding tools, productivity software, and cloud storage services.[6]

34.   Based on its $13 billion investment in OpenAI and for other reasons, Microsoft has "significant rights" in OpenAI. As Microsoft, through its CEO, has explained, "We are below them, above them, around them."[7] Microsoft CEO Satya Nadella also explained, "We have all the IP rights and all the capability. If OpenAI disappeared tomorrow, I don't want any customer of ours to be worried about it quite honestly, because we have all of the rights to continue the innovation. Not just to serve the product, but we can go and just do what we were

_____

[3] Microsoft, Machine Learning & AI | NYC, *available at* https://www.microsoft.com/enus/research/theme/machine-learning-ai-nyc/ (last accessed Jan. 2, 2024).
[4] Jennifer Langston, Microsoft announces new supercomputer, lays out vision for future AI work, Microsoft (May 19, 2020), *available at* https://news.microsoft.com/source/features/ai/openaiazure-supercomputer/ (last accessed Jan. 2, 2024).
[5] John Roach, Microsoft responsible machine learning capabilities build trust in AI systems, developers say, Microsoft (May 19, 2020), *available at* https://news.microsoft.com/source/features/ai/azure-responsible-machine-learning/ (last accessed Jan. 2, 2024).
[6] Jordan Novet, Microsoft's $13 billion bet on OpenAI carries huge potential along with plenty of uncertainty, CNBC (Apr. 9, 2023), *available at* https://www.cnbc.com/2023/04/08/microsoftscomplex-bet-on-openai-brings-potential-and-uncertainty.html (last accessed Jan. 2, 2024).
[7] Intelligencer Staff, Satya Nadella on Hiring the Most Powerful Man in AI When OpenAI threw Sam Altman overboard, Microsoft's CEO saw an opportunity, Intelligencer (Nov. 21, 2023), *available at* https://nymag.com/intelligencer/2023/11/on-with-kara-swisher-satya-nadella-onhiring-sam-altman.html (last accessed Jan. 2, 2024).

doing in partnership ourselves. We have the people, we have the compute, we have the data, we

have everything."[8]

<div align="center">

FACTUAL ALLEGATIONS

</div>

35.   OpenAI has kept secret the materials it took without compensation or permission

to "train" its LLMs. As a result, Plaintiffs cannot perfectly discern the contents of their works

that were stolen by Defendants. Thus, Plaintiffs make the specific allegations of infringement

based on what is publicly known about OpenAI's training practices and what is known about the

contents, uses, and availability of the pirate book repositories such as LibGen, Bibliotik, and Z-

Library.

I.   Mr. Nicholas Basbanes

36.   Mr. Basbanes is an internationally renowned and celebrated author. He obtained a

master's degree in journalism from Pennsylvania State University and, after receiving his

commission as an officer in the United States Navy, completed a course of study at the Defense

Information School. He made two tours to Vietnam as a Public Affairs Officer aboard the aircraft

carrier *USS Oriskany* (CV-34) in 1969 and 1970, and was awarded a Navy Achievement Medal

by the Secretary of the Navy for "professional achievement in the superior performance of his

duties" during combat operations on Yankee Station in the Gulf of Tonkin with Task Force 77.

37.   After leaving the United States Navy in 1971, Mr. Basbanes became a reporter for

The Evening Gazette in Worcester, Massachusetts, winning a first prize for investigative

reporting for large circulation newspapers in 1974 from the Associated Press Managing Editors

---

[8] Intelligencer Staff, Satya Nadella on Hiring the Most Powerful Man in AI When OpenAI threw
Sam Altman overboard, Microsoft's CEO saw an opportunity, Intelligencer (Nov. 21, 2023),
available at https://nymag.com/intelligencer/2023/11/on-with-kara-swisher-satya-nadella-
onhiring-sam-altman.html (last accessed Jan. 2, 2024).

Association of New England. In 1978 he was named the books editor of *The Worcester Sunday Telegram* and literary columnist for *The Evening Gazette*.

38.   After leaving *The Worcester Telegram & Gazette* in 1991 to become a full-time author, Mr. Basbanes devoted his professional life and financial resources to craft ten works of heavily researched and thoroughly sourced nonfiction, using the research, investigative and interviewing skills he developed as a professional journalist.

39.   To help support his research efforts, Mr. Basbanes continued to write his weekly book column, each featuring an exclusive interview with a major author, which he distributed through his own entity, Literary Features Syndicate, to more than thirty newspapers, including the *Toldeo Blade, Quincy Patriot Ledger, Minneapolis Star Tribune, Allentown Morning Call, Eugene Register Guard, Salt Lake Tribune, Knoxville News Sentinel, Des Moines Register, Gainesville Sun, New London Day, Columbus Dispatch, Grand Rapids Press* and *Wichita Eagle.* Recordings he made of these exclusive interviews—along with the interviews he conducted of individuals for his books—are now in the Cushing Memorial Library & Archives of Texas A&M University, where they have been digitized and catalogued for scholarly use. The archive of primary research materials assembled by Mr. Basbanes in support of his work over a period of forty years, when acquired by Texas A&M University in 2015, filled 365 packing boxes with documents, transcriptions, drafts, field notebooks, photographic negatives, and the like, all acquired by Mr. Basbanes in pursuit of his literary activities, and at his expense and initiative.[9]

40.   His first book, *A Gentle Madness: Bibliophiles, Bibliomanes, and the Eternal*

---

[9] *See Histoire de la Bibliophile,* "Basbanes Collection Acquired by Cushing Library at Texas A&M," *available at* https://histoire-bibliophilie.blogspot.com/2015/12/basbanes-collection-acquired-by-cushing.html (last accessed on Jan. 2, 2024). *See also* "Audio Interviews from the Nicholas A. Basbanes Collection," *available at* https://library.tamu.edu/research/digital_collections (last accessed on Jan. 2, 2024).

Passion for Books, was named a *New York Times* notable book of the year and was a finalist for the National Book Critics Circle Award in nonfiction for 1995. The *Wall Street Journal* recognized it as one of the most influential works about book collecting published in the twentieth century. *A Gentle Madness* is included in the collections of over 1,400 research libraries worldwide, according to a compilation of the WorldCat database of library collections. It is also listed as required or suggested reading in dozens of educational institutions in history of the book courses. In the words of one prominent reviewer: "No other writer has traced the history of the book so thoroughly or engagingly."[10]

41.   Mr. Basbanes devoted seven years of his professional life to research and write *A Gentle Madness*. He conducted approximately 150 interviews. He travelled across Europe and the United States entirely at his own expense to conduct research, in one instance spending two weeks attending the trial in the United States District Court for the Southern District of Iowa of the notorious book thief Stephen C. Blumberg, who stole more than 20,000 books valued at approximately $20 million over a twenty-year period. Mr. Basbanes' seven-hour interview with Mr. Blumberg was a centerpiece of a chapter that remains the definitive and frequently cited account of these crimes. And, in stark contrast to Defendants, Mr. Basbanes paid for access to and the rights for copyrighted material used in this manuscript.

42.   Mr. Basbanes' other works likewise required enormous investments of time, travel and funds. His second book, *Patience & Fortitude: A Roving Chronicle of Book People, Book Places, and Book Culture*, took approximately eight years to complete. It required approximately 100 interviews. It involved extensive travel across the United States and Europe.

---

[10] Andre Bernard, "Fear of Book Assasination [sic] Haunts Bibliophile's Musings," The New York Observer, December 15, 2013, *available at* https://observer.com/2003/12/fear-of-book-assasination-haunts-bibliophiles-musings/ (last accessed on December 30, 2023).

And, again, in stark contrast to Defendants, Mr. Basbanes paid for access to and the rights for copyrighted material used in his manuscript. The publication of this book prompted the esteemed biographer and historian David McCullough to proclaim that "Nicholas Basbanes has become the leading authority of books about books, and this, his latest, is a jewel."[11]

43.   Mr. Basbanes' nonfiction work entitled *On Paper: The Everything of Its Two-Thousand-Year History* was published in 2013 after six years of intensive effort. Mr. Basbanes' extensive travels for this project included a three-week research trip to the Hunan and Sichuan provinces in Western China, and a week in Japan, the purpose, in both instances, to acquire first-hand knowledge of ancient papermaking skills unique to those countries, and still practiced by a handful of artisans today. And, again, in stark contrast to Defendants, Mr. Basbanes paid for access to and the rights for copyrighted material used in this manuscript.

44.   In 2014, *On Paper* was named one of three recipients of the Carnegie Medal for Excellence in Nonfiction awarded annually by the American Library Association. Foreign language editions of this book resulted in invitations for Mr. Basbanes to give keynote addresses at two international conferences, one in Seoul, South Korea, the other in Mexico City.

45.   Mr. Basbanes' most recent work, *Cross of Snow: A Life of Henry Wadsworth Longfellow*, was published in 2020 after seven years of research and writing. It was named one of the best nonfiction books of the year by the *Christian Science Monitor*,[12] and one of the Books

---

[11] This comment was made in a public talk Mr. McCullough, a two-time winner of the Pulitzer Prize, gave shortly after publication of *Patience & Fortitude* at the Boston Public Library, and which he modified and allowed to be used as a testimonial in the release of the paperback edition a year later.

[12] The Christian Science Monitor, "The Best Nonfiction Books of 2020 Offer Wisdom and Insight," December 9, 2020, *available at* https://www.csmonitor.com/Books/2020/1209/The-best-nonfiction-books-of-2020-offer-wisdom-and-insight (last accessed on December 30, 2023).

of the Year by *Times Literary Supplement*.[13] It also received Top Honors in Nonfiction for the

Massachusetts Book Award given by the Massachusetts Center for the Book.[14] *The New York*

Times recognized that Mr. Basbanes, in crafting this manuscript, "is a painstaking researcher, the

kind who turns every page, as Robert Caro would say, and he has benefited from access to lots of

material previously unavailable."[15] *The Worcester Telegram & Gazette* noted that it represented

"a great deal of new biographical research and reporting."[16] *The Boston Globe* commented,

succinctly: "His research is prodigious."[17] In a lead review for *The Wall Street Journal,* the noted

Longfellow scholar Cristoph Irmscher declared *Cross of Snow* to be "the biography Longfellow

himself would have most liked to read."[18]

46.   In 2016, Mr. Basbanes was selected by the National Endowment for the

Humanities as part of its first group of nonfiction writers formally designated as Public Scholars.

47.   He is now researching and writing, with his own resources, a book under contract

[13] TLS, "Books of the Year 2020, Sixty-five Writers Make Their Selections From Around the World," November 13, 2020, *available at* https://www.the-tls.co.uk/articles/books-of-the-year-2020/ (last accessed on December 30, 2023).

[14] Massachusetts Center for the Book, 21st Annual Massachusetts Book Awards, *available at* https://myemail.constantcontact.com/Massachusetts-Book-Awards-Announcement-and-Call-for-Submissions.html?soid=1118683784419&aid=z-5RN-3yWZo, (last accessed on December 30, 2023).

[15] The New York Times, "Henry Wadsworth Longfellow: America's No. 1 Literary Celebrity," June 4, 2020, *available at* https://www.nytimes.com/2020/06/04/books/review/cross-of-snow-a-life-of-henry-wadsworth-longfellow-nicholas-a-basbanes.html (last accessed on December 30, 2023).

[16] The Worcester Telegram & Gazette, "Delving Deep Into Longfellow's Life," June 6, 2020, *available at* https://www.telegram.com/story/entertainment/local/2020/06/06/grafton-author-nicholas-basbanes-forges-biographers-bond-with-longfellow/113779660/ (last accessed on December 30, 2023).

[17] The Boston Globe, "Reconsidering Longfellow in 'Cross of Snow,'" May 28, 2020, *available at* https://www.bostonglobe.com/2020/05/28/arts/reconsidering-longfellow-cross-snow/ (last accessed on December 30, 2023).

[18] The Wall Street Journal, "'Cross of Snow' Review: Our Poet of Loneliness," *available at* https://www.wsj.com/articles/cross-of-snow-review-our-poet-of-loneliness-11590156990 (last accessed on Jan. 2, 2024).

with Yale University Press to be titled *Before Paper: The Hunt for the World's Earliest Writings,* a prequel to his 2013 book, *On Paper.*

48.   Mr. Basbanes other copyrighted works, which are listed on Exhibit A to this Complaint, likewise required enormous investments of his time and resources to conduct interviews, pore over archival material, and ultimately generate the knowledge that anyone can readily buy for a reasonable price. His works have now been published in five countries and translated into four languages.

49.   Mr. Basbanes is the sole author of and beneficial owner of the registered copyrights identified in Exhibit A to this Complaint, all of which OpenAI used without compensation or permission to "train" its LLMs and power its massive commercial enterprise.

II.   Mr. Nicholas Gage

50.   Mr. Gage is among America's greatest and most acclaimed investigative reporters. He has a master's degree from the Columbia University Graduate School of Journalism and his work for *The Wall Street Journal* and *The New York Times* impacted the course of key events in American and world history.

51.   Mr. Gage's investigations were instrumental in exposing past corruption of Vice President Spiro Agnew, leading to his resignation. Mr. Gage was also the first reporter to hear any of President Richard Nixon's secret tapes that were uncovered in connection with the Watergate Scandal. His deeply sourced and exclusive writings for *The New York Times* and *The Wall Street Journal* set the standard for investigative inquiries into the activities of organized crime in the United States.

52.   Using his journalistic skills, Mr. Gage went on to become one of the most influential nonfiction authors of the twentieth century. His 1983 best-selling autobiographical

memoir *Eleni*, which described the experiences of his family during the Second World War and Greek Civil War, was ultimately adapted into a feature film starring John Malkovich and Kate Nelligan. Mr. Gage describes in the book how the communists executed his mother for helping Mr. Gage and his sisters escape the occupied village of Lia in northern Greece. President Ronald Reagan cited *Eleni* as an inspiration for his summit meetings with the Soviet Union to end the arms race.

53.   A finalist for the National Book Critics Award in the United States, *Eleni* has been translated into 15 languages, published in 32 countries, and was the recipient of numerous other prestigious awards, including a first prize by the Royal Society of Literature of Great Britain.

54.   Mr. Gage's coverage of the mafia during his time as a reporter also led to two other bestsellers: *The Mafia Is Not An Equal Opportunity Employer* and *Mafia, U.S.A.*

55.   Mr. Gage was also an executive producer of the acclaimed film *The Godfather Part III*, co-writing an early draft of the script with Mario Puzo. The film was ultimately nominated for seven Golden Globe Awards and seven Academy Awards.

56.   Mr. Gage's other nonfiction books include: *A Place for Us* (1989)*,* a sequel to *Eleni,* which describes the immigrant experience of Mr. Gage and his five sisters to the United States following their escape from civil-war-torn Greece, and the establishment of new lives as naturalized American citizens, and *Greek Fire: The Story of Maria Callas and Aristotle Onassis* (2000), which was based on thoroughly researched and impeccably sourced material including previously unexamined materials.

57.   Mr. Gage is the sole author of and beneficial owner of the registered copyrights identified in Exhibit A to this Complaint, all of which OpenAI used without compensation or

permission to "train" its LLMs and power its massive commercial enterprise.

III.   Generative AI and Large Language Models

58.   The terms "artificial intelligence" or "AI" refer generally to computer systems designed to imitate human cognitive functions.

59.   The terms "generative artificial intelligence" or "generative AI" refer specifically to systems that are capable of generating "new" content in response to user inputs called "prompts."

60.   For example, the user of a generative AI system capable of generating images from text prompts might input the prompt, "A writer working at her desk." The system would then attempt to construct the prompted image. Similarly, the user of a generative AI system capable of generating text from text prompts might input the prompt, "Tell me a story about a writer working at her desk." The system would then attempt to generate the prompted text.

61.   Recent generative AI systems designed to recognize input text and generate output text are built on "large language models" or "LLMs."

62.   LLMs use predictive algorithms that are designed to detect statistical patterns in the text datasets on which they are "trained" and, on the basis of these patterns, generate responses to user prompts. "Training" an LLM refers to the process by which the parameters that define an LLM's behavior are adjusted through the LLM's ingestion and analysis of large "training" datasets.

63.   Once "trained," the LLM analyzes the relationships among words in an input prompt and generates a response that is an approximation of similar relationships among words in the LLM's "training" data. In this way, LLMs can be capable of generating sentences, paragraphs, and even complete texts, from short pieces to book length treatments.

64.   "Training" an LLM requires inputting large numbers of parameters in the model

and then supplying the LLM with large amounts of text for the LLM to ingest—the more text,

the better. That is, in part, the large in large language model.

65.   As the U.S. Patent and Trademark Office has observed, LLM "training" "almost

by definition involve[s] the reproduction of entire works or substantial portions thereof."[19]

66.   "Training" in this context is therefore a technical-sounding euphemism for

copying and using large volumes of written material, including material that belongs to

Plaintiffs.

67.   Moreover, in some form and to some degree currently unknowable to the public,

Defendants' LLMs have "memorized" or stored their "training" data (even if in a "translated"

form), such that the data (at least in part) can be accessed, recalled, and reproduced by the LLM

at will.[20]

68.   The quality of the LLM (that is, its capacity to generate human-seeming responses

to prompts) is dependent on the quality of the datasets used to "train" the LLM.

69.   Professionally authored, edited, and published books—such as those authored by

Plaintiffs here—are an especially important source of LLM "training" data.

70.   As one group of AI researchers (not affiliated with Defendants) has observed,

"[b]ooks are a rich source of both fine-grained information, how a character, an object or a scene

---

[19] U.S. Patent & Trademark Office, *Public Views on Artificial Intelligence and Intellectual Property Policy* 29 (2020), *available at* https://www.uspto.gov/sites/default/files/documents/USPTO_AI-Report_2020-10-07.pdf (last accessed Jan. 2, 2024).
[20] *See* Jason Koebler, *Google Researchers' Attack Prompts ChatGPT to Reveal Its Training Data*, 404 Media (Nov. 29, 2023), *available at* https://www.404media.co/google-researchersattack-convinces-chatgpt-to-reveal-its-training-data/ (last accessed Jan. 2, 2024); Kent K. Chang *et al.*, *Speak, Memory: An Archaeology of Books Known to ChatGPT/GPT-4* (2023), *available at* https://arxiv.org/pdf/2305.00118v1.pdf (last accessed Jan. 2, 2024).

looks like, as well as high-level semantics, what someone is thinking, feeling and how these states evolve through a story."[21]

71.   In other words, Plaintiffs' copyrighted books are the high-quality materials Defendants want and need to build their new and massive commercial enterprise. Instead of paying for those high-quality materials, like a general contractor might buy insulation, nails and screws to build a house, Defendants simply stole them.

IV.   Defendants Willful Infringement of Plaintiffs' Copyrights

72.   Defendant OpenAI Inc. was founded in 2015 as a nonprofit organization with the self-professed goal of researching and developing AI tools "unconstrained by a need to generate financial return."[22]

73.   OpenAI's nonprofit status was short lived. In 2019, OpenAI, Inc. relaunched itself (specifically, through Defendant OpenAI GP, L.L.C. and Defendant OpenAI OpCo, LLC)[23] as a for-profit enterprise.

74.   Investments quickly poured in. Microsoft, one of the world's largest technology companies, invested $1 billion in 2019, an estimated $2 billion in 2021, and a staggering $10 billion in 2023, for a total investment of $13 billion so far.

75.   Industry observers currently value OpenAI at up to $80 billion.[24]

---

[21] Yukun Zhu *et al.*, *Aligning Books and Movies: Towards Story-like Visual Explanations by Watching Movies and Reading Books* 1 (2015), *available at* https://arxiv.org/pdf/1506.06724.pdf (last accessed Jan. 2, 2024).
[22] OpenAI, *Introducing OpenAI* (Dec. 11, 2015), https://openai.com/blog/introducing-openai (last accessed Jan. 2, 2024).
[23] Defendant OpenAI OpCo LLC was then known as OpenAI LP.
[24] Kate Clark, *Thrive Capital to Lead Purchase of OpenAI Employee Shares at $80 Billion-Plus Valuation*, The Information (Oct. 19, 2023), *available at* https://www.theinformation.com/articles/thrive-capital-to-lead-purchase-of-openai-employeeshares-at-80-billion-plus-valuation (last accessed Jan. 2, 2024).

V.   The Large Language Models

76.   OpenAI's LLMs are collectively referred to as "GPT-N," which stands for "Generative Pre-trained Transformer" (a specific type of LLM architecture), followed by a version number.

77.   GPT-3 was released in 2020 and exclusively licensed to Microsoft the same year.

78.   OpenAI further refined GPT-3 into GPT-3.5, which was released in 2022.

79.   In November 2022, OpenAI released ChatGPT, a consumer-facing chatbot application built on GPT-3.5.

80.   ChatGPT's popularity exploded virtually overnight. By January 2023, less than three months after its release, the application had an estimated 100 million monthly active users, making it one of the fastest-growing consumer applications in history.

81.   GPT-4, the successor to GPT-3.5, was released in March 2023.

82.   GPT-4 underlies OpenAI's new subscription-based chatbot, called ChatGPT Plus, which is available to consumers for $20 per month.

83.   Defendants intend to earn billions of dollars from this technology.

84.   When announcing the release of ChatGPT Enterprise, a subscription-based high-capability GPT-4 application targeted for corporate clients, in August 2023, OpenAI claimed that teams in "over 80% of Fortune 500 companies" were using its products.[25]

85.   GPT-4 also underlies Microsoft's Bing Chat product, offered through its Bing Internet search engine, and is integrated into its sales and marketing software, coding tools, productivity software, and cloud storage services.

---

[25] OpenAI, *Introducing ChatGPT Enterprise* (Aug. 28, 2023),
available at https://openai.com/blog/introducing-chatgpt-enterprise (last accessed Jan. 2, 2024).

86. OpenAI CEO Sam Altman recently reported to OpenAI employees that OpenAI is on track to generate $1.3 billion in revenue in 2023.[26]

87. Analysts estimate that Microsoft could earn more than $10 billion in annual revenue by 2026 *only* from AI add-ons to its Microsoft 365 productivity software, "at the core" of which lies OpenAI technology.[27]

VI. Knowingly "Training" on Copyrighted Books

88. OpenAI does not disclose or publicize with specificity what datasets GPT-3, GPT-3.5, or GPT-4 were "trained" on. OpenAI treats that information as proprietary.

89. To "train" its LLMs—including GPT-3, GPT-3.5, and GPT-4—OpenAI has reproduced copyrighted books without their authors' consent. OpenAI has publicly admitted this.

90. It has admitted that it has "trained" its LLMs on "large, publicly available datasets that include copyrighted works."[28]

91. It has admitted that "training" LLMs "require[s] large amounts of data," and that "analyzing large corpora" of data "necessarily involves first making copies of the data to be analyzed."[29]

92. It has admitted that, if it refrained from using copyrighted works in its LLMs'

---

[26] Amir Efrati, *OpenAI's Revenue Crossed $1.3 Billion Annualized Rate, CEO Tells Staff*, The Information (Oct. 12, 2023), *available at* https://www.theinformation.com/articles/openaisrevenue-crossed-1-3-billion-annualized-rate-ceo-tells-staff (last accessed Jan. 2, 2024).

[27] Jordan Novet, *Microsoft starts selling AI tool for Office, which could generate $10 billion a year by 2026*, CNBC (Nov. 1, 2023), *available at* https://www.cnbc.com/2023/11/01/microsoft-365-copilot-becomes-generally-available.html (last accessed Jan. 2, 2024).

[28] OpenAI, *Comment Regarding Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation*, U.S. Patent and Trademark Office Dkt. No. PTO-C-2019-0038, at 1 (2019), *available at* https://www.uspto.gov/sites/default/files/documents/OpenAI_RFC-84-FR-58141.pdf (last accessed Jan. 2, 2024).

[29] *Id.*

"training," it would "lead to significant reductions in model quality."[30]

93.   Accordingly, OpenAI has openly admitted to reproducing copyrighted works in the course of "training" its LLMs because such reproduction is central to the quality of its products.

94.   Instead of "verbatim excerpts," ChatGPT now offers to produces summaries of copyrighted books, which usually contains details not available in reviews and other publicly available material—again suggesting that the underlying LLM must have ingested the entire book during its "training."

95.   OpenAI is characteristically opaque about where and how it procured the entirety of these books.

96.   OpenAI admits that among the "training" datasets it used to "train" were "Common Crawl," and two "high-quality," "internet-based books corpora" which it calls "Books1" and "Books2."[31]

97.   Common Crawl is a vast and growing corpus of "raw web page data, metadata extracts, and text extracts" scraped from billions of web pages. It is widely used in "training" LLMs, and has been used to "train," in addition to GPT-N, Meta's LlaMa, and Google's BERT. It is known to contain text from books copied from pirate sites.[32]

98.   Some independent AI researchers suspect that Books2 contains or consists of ebook files downloaded from large pirate book repositories such as Library Genesis or

---

[30] *Id.* at 7 n.33.

[31] Tom B. Brown *et al.*, *Language Models Are Few-Shot Learners* 8 (2020), *available at* https://arxiv.org/pdf/2005.14165.pdf (last accessed Jan. 2, 2024).

[32] Alex Hern, *Fresh Concerns Raised Over Sources of Training Material for AI Systems*, The Guardian (Apr. 20, 2023), *available at* https://www.theguardian.com/technology/2023/apr/20/fresh-concerns-training-material-aisystems-facist-pirated-malicious (last accessed Jan. 2, 2024).

"LibGen," "which offers a vast repository of pirated text."[33]

99. LibGen is already known to this Court as a notorious copyright infringer.[34]

100. Other possible candidates for Books2's sources include Z-Library, another large pirate book repository that hosts more than 11 million books, and pirate torrent trackers like Bibliotik, which allow users to download ebooks in bulk.

101. Websites linked to Z-Library appear in the Common Crawl corpus and have been included in the "training" dataset of other LLMs.[35]

102. Z-Library's Internet domains were seized by the FBI in February 2022, only months after OpenAI stopped "training" GPT-3.5 in September 2021.

103. The disclosed size of the Books2 dataset (55 billion "tokens," the basic units of textual meaning such as words, syllables, numbers, and punctuation marks) suggests it comprises over 100,000 books.

104. "Books3," a dataset compiled by an independent AI researcher, is comprised of nearly 200,000 books downloaded from Bibliotik, and has been used by other AI developers to "train" LLMs.

105. The similarities in the sizes of Books2 and Books3, and the fact that there are only a few pirate repositories on the Internet that allow bulk ebook downloads, strongly indicates that the books contained in Books2 were also obtained from one of the notorious repositories discussed above.

---

[33] Kate Knibbs, *The Battle Over Books3 Could Change AI Forever*, Wired (Sept. 4, 2023), *available at* https://www.wired.com/story/battle-over-books3 (last accessed Jan. 2, 2024).
[34] *See Elsevier Inc. v. Sci-Hub*, No. 1:15-cv-4282-RWS (S.D.N.Y.).
[35] Kevin Schaul *et al.*, *Inside the Secret List of Websites that Make AI Like ChatGPT Sounds Smart*, The Washington Post (Apr. 19, 2023), *available at* https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning (last accessed Jan. 2, 2024).

106.   OpenAI has not discussed the datasets used to "train" GPT-3.5, GPT-4, or their source or sources.

107.   GPT-3.5 and GPT-4 are significantly more powerful than their predecessors. GPT 3.5 contains roughly 200 billion parameters, and GPT-4 contains roughly 1.75 trillion parameters, compared to GPT-3's roughly 175 billion parameters.

108.   The growth in power and sophistication from GPT-3 to GPT-4 suggests a correlative growth in the size of the "training" datasets, raising the inference that one or more very large sources of pirated ebooks discussed above must have been used to "train" GPT-4.

109.   There is no other way OpenAI could have obtained the volume of books required to "train" a powerful LLM like GPT-4.

110.   In short, OpenAI admits it needs[36] and uses[37] "large, publicly available datasets that include copyrighted works"[38]—and specifically, "high-quality"[39] copyrighted books—to "train" its LLMs; pirated sources of such "training" data are readily available; and one or more of these sources contain Plaintiffs' works.

111.   Defendants knew that their "training" data included texts protected by copyright but willfully proceeded without obtaining authorization.

112.   OpenAI's "training" its LLMs could not have happened without Microsoft's financial and technical support. In 2020, Microsoft announced that it had developed "one of the top five publicly disclosed supercomputers in the world" that had been "[b]uilt in collaboration with and exclusively for OpenAI," and "designed specifically to train that company's AI

---

[36] OpenAI, *Comment Regarding Request for Comments*, *supra*, at 28 n.33.
[37] *Id.* at 2.
[38] *Id.* at 1.
[39] Brown *et al.*, *Few-Shot Learners*, *supra*, at 31.

models."[40] And in 2023, Microsoft CEO Satya Nadella reminded the world that the "heavy lifting" for OpenAI's LLM "training" was done by Microsoft "compute infrastructure."[41]

VII.   GPT-N's and ChatGPT's Harm to Authors

113.   ChatGPT and the LLMs underlying it seriously threaten the livelihood of Plaintiffs, on whose works they were "trained" without consent.

114.   Goldman Sachs estimates that generative AI could replace 300 million full-time jobs in the near future, or one-fourth of the labor currently performed in the United States and Europe.

115.   Already, writers report losing income from copywriting, journalism, and online content writing—important sources of income for many authors.

116.   Recently, the owner of popular online publications such as *Gizmodo*, *Deadspin*, *The Root*, *Jezebel* and *The Onion* came under fire for publishing an error-riddled, AI-generated piece, leading the Writers Guild of America to demand "an immediate end of AI-generated articles" on the company's properties.[42]

117.   Until recently, ChatGPT provided verbatim quotes of copyrighted text. Currently, it instead readily offers to produce summaries of such text. These summaries are themselves

---

[40] Jennifer Langston, *Microsoft announces new supercomputer, lays out vision for future AI work*, Microsoft (May 19, 2020), *available at* https://news.microsoft.com/source/features/ai/openai-azure-supercomputer/ (last accessed Jan. 2, 2024).

[41] CNBC, *First on CNBC: CNBC Transcript: Microsoft CEO Satya Nadella Speaks with CNBC's Jon Fortt on "Power Lunch" Today* (Fed. 7, 2023), *available at* https://www.cnbc.com/2023/02/07/first-on-cnbc-cnbc-transcript-microsoft-ceo-satya-nadellaspeaks-with-cnbcs-jon-fortt-on-power-lunch-today.html (last accessed Jan. 2, 2024).

[42] Todd Spangler, *WGA Slams G/O Media's AI-Generated Articles as 'Existential Threat to Journalism,' Demands Company End Practice*, Variety (July 12, 2023), https://variety.com/2023/digital/news/wga-slams-go-media-ai-generated-articles-existentialthreat-1235668496 (last accessed Jan. 2, 2024).

derivative works, the creation of which is inherently based on the original unlawfully copied

work and could be—but for ChatGPT—licensed by the authors who created these underlying

works to willing, *paying* licensees.

118.   Plaintiffs are thus reasonably concerned about the risks OpenAI's conduct poses

to their livelihoods.

119.   In short, the success and profitability of OpenAI are predicated on mass copyright

infringement, *i.e.* use without permission from or compensation to copyright owners, including

Plaintiffs' here.

CLASS ALLEGATIONS

I.   Class Definitions

120.   Plaintiffs bring this action as Proposed Class Representatives for and on behalf of

the Proposed Class and Proposed Class Members, as defined below, under Federal Rules of Civil

Procedure 23(b)(2), 23(b)(3), and 23(c)(4).

121.   The Proposed Class is defined as follows:

> All natural persons in the United States who are authors and legal or
> beneficial owners, in full or in part, of copyrights in, one or more written
> works which has been, or is being, used by Defendants' to train their large
> language models.

122.   Excluded from the class definitions above are Defendants; Defendants'

coconspirators, aiders and abettors, and members of their immediate families; Defendants'

corporate parents, subsidiaries, and affiliates; Defendants' directors, officers, employees, and

other agents, as well as members of their immediate families; and any judge who may preside

over this action, the judge's staff, and members of their immediate families.

II.   Rules 23(a) and 23(g)

123.   The Proposed Class is sufficiently numerous because it is estimated to have tens

of thousands of members.

124. The identities of the Proposed Class Members are objectively ascertainable because Defendants know, and can produce in discovery, which texts they used to "train" their large language models; and because information regarding copyright ownership, copyright registration, and book sales is determinable from public or other objective sources and measures.

125. The Proposed Class Representatives' claims are typical of the claims of the Proposed Class because their copyrights were infringed in materially the same way and their interests in preventing future infringement and redressing past infringement are materially the same.

126. The Proposed Class Representative will adequately represent the Proposed Class, and Plaintiffs' counsel is experienced, knowledgeable, well resourced, and will zealously and faithfully represent Plaintiffs and the Proposed Class.

127. There are questions of law or fact common to the Proposed Class, including (a) whether Defendants copied Plaintiffs' and Proposed Class Members' copyrighted works in "training" their LLMs; (b) whether Defendants' copying of Plaintiffs' and Proposed Class Members' copyrighted works constitutes direct, vicarious, or contributory infringement under the Copyright Act; and (c) whether Defendants' copying of Plaintiffs' and Proposed Class Members' copyrighted works was willful.

III. Rule 23(b)

128. Defendants have acted on grounds common to Plaintiffs and the Proposed Class by treating all Plaintiffs' and Proposed Class Members' works equally, in all material respects, in their LLM "training."

129. Common questions of liability for infringement predominate over any

individualized damages determinations as may be necessary. To decide liability, the Court will necessarily apply the same law to the same conduct, which Defendants engaged in indiscriminately with respect to Plaintiffs and all Proposed Class Members.

130.   Further, to the extent Plaintiffs elect to pursue statutory rather than actual damages before final judgment, the damages inquiry will likewise be common, if not identical, for Plaintiffs and Proposed Class Members.

131.   A class action is superior to any individual litigation of Plaintiffs' and Proposed Class Members' claims. Proposed Class Members have little interest, distinct from Plaintiffs' and other Proposed Class Members', in prosecuting individual actions. It would waste judicial resources to decide the same legal questions repeatedly, thousands of times over, on materially indistinguishable facts. The Proposed Class presents no special manageability problems.

IV.   Rule 23(c)(4)

132.   In the alternative to certification under Rule 23(b)(3), common questions predominate within the determination of liability for infringement, and therefore the issue of liability may be separately certified for class treatment even if the entire action is not.

CLAIMS TO RELIEF

COUNT I: DIRECT COPYRIGHT INFRINGEMENT (17 U.S.C. § 501)
Against All Defendants

133.   Plaintiffs incorporate and reallege paragraphs 1 through 129 above.

134.   Plaintiffs and Proposed Class Members are the rightful and lawful legal or beneficial owners of the copyrights in and to their written works.

135.   Plaintiffs' and Proposed Class Members' written works are original to their authors and are fixed in tangible mediums of expression as works under 17 U.S.C. § 102(a)(1).

136.   Plaintiffs and Proposed Class Members have duly and timely registered their

copyrights in their works with the U.S. Copyright Office.

137. Plaintiffs and Proposed Class Members are legal or beneficial owners of the exclusive right to reproduce their copyrighted works in copies under 17 U.S.C. § 106(1), as well as the right to refrain from such reproduction.

138. Defendants had access to Plaintiffs' and Proposed Class Members' copyrighted works, including by way of the various unauthorized datasets discussed above.

139. Defendants violated Plaintiffs' and Proposed Class Members' exclusive rights by reproducing their copyrighted works in copies for the purpose of "training" their LLMs and ChatGPT.

140. Defendant's violation of Plaintiffs' and Proposed Class Members' exclusive rights was willful because Defendants knew the datasets on which it "trained" its large language models contained copyrighted works.

COUNT II: VICARIOUS COPYRIGHT INFRINGEMENT
Against Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC, OpenAI Holdings, LLC and OpenAI Global, LLC

141. Plaintiffs incorporate and reallege paragraphs 1 through 129 above.

142. Defendants Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC, OpenAI Holdings, LLC and OpenAI Global, LLC controlled, directed and profited from the direct infringement alleged in Count I.

143. Microsoft controlled, directed, and profited from the infringement perpetrated by the OpenAI Defendants. Microsoft controls and directs the supercomputing platform used to store, process and reproduce the training datasets. Microsoft profited from the infringement alleged in this Complaint because it incorporated the infringing GPT models trained on Plaintiffs' copyrighted works into its own commercial product offerings, including Bing Chat.

144.   Defendants OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC, OpenAI

Holdings, LLC and OpenAI Global, LLC are vicariously liable for the direct infringement

alleged in Count I.

<div align="center">

COUNT III: CONTRIBUTORY COPYRIGHT INFRINGEMENT
Against Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC, OpenAI
Holdings, LLC and OpenAI Global, LLC

</div>

145.   Plaintiffs incorporate and reallege paragraphs 1 through 129 above.

146.   Defendants Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC,

OpenAI Holdings, LLC and OpenAI Global, LLC materially contributed to and directly assisted

in the direct infringement alleged in Count I by funding the direct infringement by way of

capital, technology, personnel, and other resources; controlling or managing the property or other

assets with which the direct infringement was accomplished; or providing business, legal,

strategic, or operational guidance to accomplish the direct infringement.

147.   Defendants Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC,

OpenAI Holdings, LLC and OpenAI Global, LLC knew or had reason to know of the direct

infringement alleged in Count I, because these Defendants share management personnel and

operational plans with Defendants and are fully aware of the capabilities of their own product

and the materials upon which it was "trained," including known caches of pilfered copyrighted

works.

148.   Defendants Microsoft, OpenAI, Inc., OpenAI GP, L.L.C., OAI Corporation, LLC,

OpenAI Holdings, LLC and OpenAI Global, LLC are contributorily liable for the direct

infringement alleged in Count I.

<div align="center">

PRAYER FOR RELIEF

</div>

149.   Plaintiffs, on behalf of themselves and all others similarly situated, pray for the

following relief:

(a)   Certification of this action as a class action under Federal Rule of Civil

Procedure 23;

(b)   Designation of Mr. Basbanes and Mr. Gage as class representatives;

(c)   Designation of Mr. Basbanes' and Mr. Gage's counsel as class counsel;

(d)   An injunction prohibiting Defendants from infringing Plaintiffs' and class

members' copyrights, including without limitation enjoining Defendants from

using Plaintiffs' and class members' copyrighted works in "training" Defendants'

large language models without express authorization;

(e)   An award of actual damages to Plaintiffs and class members;

(f)   An award of Defendants' additional profits attributable to infringement to

Plaintiffs and class members;

(g)   An award of statutory damages up to $150,000 per infringed work to

Plaintiffs and class members, in the alternative to actual damages and profits, at

Plaintiffs election before final judgment;

(h)   Reasonable attorneys' fees and costs, as allowed by law;

(i)   Pre-judgment and post-judgment interest, as allowed by law; and

(j)   Such further relief as the Court may deem just and proper.

JURY DEMAND

150.   Plaintiffs demand a trial by jury as to all issues so triable.

DATED: January 5, 2024
      New York, New York

GRANT HERRMANN SCHWARTZ & KLINGER LLP

By:*/s/ Michael P. Richter*
    Michael P. Richter (MR2230)
    Bryan M. Goldstein (4980702)
    107 Greenwich Street, 25th Floor
    New York, New York 10006
    Tel.: (212) 682-1800
    mrichter@ghsklaw.com
    bgoldstein@ghsklaw.com

Attorneys for Mr. Basbanes, Mr. Gage and the
 Proposed Class