

GAW | POE LLP

Randolph Gaw (S.B. # 223718)
rgaw@gawpoe.com
One Embarcadero Ctr., Suite 1200
San Francisco, CA 94111
Tel: (415) 766-7451

SLARSKEY LLC

David Slarskey (*pro hac vice*)
dslarskey@slarskey.com
Adam Hollander (*pro hac vice pending*)
ahollander@slarskey.com
Deepa Devanathan (*pro hac vice pending*)
ddevanathan@slarskey.com
Phillip Gasperetti (*pro hac vice pending*)
pgasperetti@slarskey.com
767 Third Ave, 14th Floor
New York, NY 10017
Tel: (212) 658-0661

ARCHSTONE LAW GROUP, P.C.

Brenda Ulrich (*pro hac vice pending*)
bulrich@archstonelaw.com
Zick Rubin (*pro hac vice pending*)
zrubin@archstonelaw.com
Riverside Center
275 Grove Street, Suite 2-400
Newton, MA 02466
Tel: (781) 314-0100
Attorneys for Plaintiffs and the Proposed Class

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

MICHAEL SULLIVAN, KENNETH S.
SALADIN, ZVI BODIE, ALAN J. MARCUS,
KEVIN T. PATTON, LION DEN INC., and
DANA LOEWY individually and on behalf of
all others similarly situated,

Plaintiffs,

v.

META PLATFORMS, INC. and MARK
ZUCKERBERG,

Defendants.

Case No. 26-6793

CLASS ACTION COMPLAINT

Class Action

Demand for Jury Trial

1 Plaintiffs Michael Sullivan, Kenneth S. Saladin, Zvi Bodie, Alan J. Marcus, Kevin Patton, Lion
2 Den Inc., and Dana Loewy (“Plaintiffs”), individually and on behalf of all others similarly situated,
3 by and through undersigned counsel, for their Complaint against Defendants Meta Platforms, Inc.
4 (“Meta”) and Mark Zuckerberg (“Zuckerberg,” and together with Meta, “Defendants”), allege as
5 follows:

6 **INTRODUCTION**

7 1. This is a class action brought by textbook authors (the “Class”) whose copyrighted
8 textbooks Meta copied, reproduced, distributed, and used without authorization to build and train its
9 generative large language model (“LLM” or “model”) known as Llama.

10 2. Without license, consent, credit, or compensation to the textbook authors, Meta
11 obtained Plaintiffs’ and the Class’s copyrighted textbooks from pirated “shadow library” sources,
12 reproduced them, distributed them to others while torrenting them, removed copyright-management
13 information from them, and used them to train its Llama models.

14 3. Plaintiffs are textbook authors. Textbook authors occupy a distinct position in the
15 market for written works, as textbooks are created, marketed, and sold for the specific purpose of
16 teaching academic subjects to students in classroom and course settings at educational institutions.
17 Textbooks are reviewed and selected by instructors and through formal adoption processes, are
18 typically updated, revised, and published in successive editions on a regular basis, and are
19 commercialized in educational markets that fundamentally differ from the markets for trade books and
20 scholarly works.

21 4. Those distinctions are central to this litigation. The market for textbooks and the risk
22 that AI models trained on textbooks will substitute for and dilute that market turns in part upon
23 evidence of textbook adoption, revision economics, and customer purchasing decisions that are unique
24 to textbook authors as compared to other categories of authors.

5. Each Plaintiff assigned exclusive publication and distribution rights under the Copyright Act to his or her respective publisher, subject to the publisher's contractual obligation to pay royalties to the Plaintiff based on the publisher's exploitation of those exclusive rights. Accordingly, each Plaintiff is the beneficial owner of copyright in his or her work and is entitled to institute an action for infringement of the copyright, pursuant to 17 U.S.C. § 501(b).

6. Plaintiffs assert claims for direct copyright infringement, contributory copyright infringement, and removal and distribution of copyright-management information under the Digital Millennium Copyright Act, on behalf of themselves and a class of textbook authors.

JURISDICTION AND VENUE

7. This Court has subject-matter jurisdiction over this action under 28 U.S.C. §§ 1331 and 1338 because this action arises under the Copyright Act, 17 U.S.C. §§ 101 et seq., and the Digital Millennium Copyright Act, 17 U.S.C. §§ 501, 1202, and 1203.

8. This Court has personal jurisdiction over Defendants.

9. Defendant Meta maintains its principal place of business in this District and has purposefully availed itself of this forum, and the conduct giving rise to Plaintiffs' claims occurred substantially in this District.

10. Defendant Mark Zuckerberg resides in and directs Meta's operations from this District and has purposefully availed himself of this forum, and the conduct giving rise to Plaintiffs' claims occurred substantially in this District. Zuckerberg oversaw, approved, and/or directed the illegal conduct alleged in this Complaint, and foreseeably caused injury to persons residing in, located in, or doing business in this District.

11. Venue is proper in this District under 28 U.S.C. § 1391 because Defendant Meta is headquartered in this District, Defendant Zuckerberg resides in this District, and a substantial part of the events or omissions giving rise to the claims occurred in this District, including alleged copyright infringement.

INTRADISTRICT ASSIGNMENT

12. Assignment of this action is governed by the District’s Assignment Plan. Because this is an Intellectual Property Action, it is assigned on a district-wide basis under Civil L.R. 3-2(c) and General Order No. 44, rather than to a particular division under Civil L.R. 3-2(d).

PARTIES

13. Plaintiff Michael Sullivan is an individual residing in Florida and is Professor Emeritus of Mathematics at Chicago State University, where he taught for more than 35 years.

14. Dr. Sullivan is the author of more than 120 mathematics textbooks and editions published by Pearson Education, Inc., W.H. Freeman & Co. and BFW Publishing Group (subsidiaries of Macmillan Publishers) and John Wiley & Sons, Inc., including the widely adopted titles *College Algebra*, *Algebra and Trigonometry*, *Precalculus*, *Trigonometry: A Unit Circle Approach* (12th Edition, Pearson Education), 3-book *Precalculus series*, *Concepts Through Function* (5th Edition, Pearson Education), 3-book *Precalculus Enhanced with Graphing Utilities series* (9th Edition, Pearson Education), and *Calculus for the AP Course*, 4th Edition, BFW Publishers.

15. Dr. Sullivan has beneficial interests, copyrights, exclusive rights, royalty interests, contractual rights, or other legally protected interests in these and other works that give him standing within the class.

16. Dr. Sullivan is a past president of the Textbook & Academic Authors Association (“TAA”), a member organization whose mission is to advance the interests of textbook and academic authors. TAA’s “Michael Sullivan Lecture on Textbook & Academic Authoring” is named in his honor. He currently serves as TAA’s Treasurer.

17. Plaintiff Kenneth S. Saladin is an individual residing in Georgia and is Distinguished Professor of Biology, Emeritus, at Georgia College & State University.

18. Dr. Saladin is the author of 21 titles and editions of English-language textbooks of human anatomy and physiology published by McGraw Hill, LLC. Dr. Saladin’s textbooks include

1 *Anatomy & Physiology: The Unity of Form and Function, Human Anatomy, and Essentials of Anatomy*
2 *& Physiology.*

3 19. Dr. Saladin has beneficial interests, copyrights, exclusive rights, royalty interests,
4 contractual rights, or other legally protected interests in these and other works that, give him standing
5 within the class.

6 20. Plaintiff Zvi Bodie is an individual residing in Massachusetts and is Professor Emeritus
7 of Finance at the Questrom School of Business at Boston University, where he taught for more than
8 four decades.

9 21. Dr. Bodie is a co-author of the leading finance textbooks published in multiple editions,
10 including *Investments and Essentials of Investments*, published by McGraw Hill, LLC, and *Financial*
11 *Economics*, published by Cambridge University Press.

12 22. Dr. Bodie has beneficial interests, copyrights, exclusive rights, royalty interests,
13 contractual rights, or other legally protected interests in these and other works that give him standing
14 within the class.

15 23. Plaintiff Alan J. Marcus is an individual residing in Massachusetts and is the Mario J.
16 Gabelli Endowed Professor of Finance at the Carroll School of Management at Boston College.

17 24. Dr. Marcus is a co-author of leading finance textbooks in multiple editions, including
18 *Investments and Essentials of Investments* and *Fundamentals of Corporate Finance*, all published by
19 McGraw Hill, LLC.

20 25. Dr. Marcus has beneficial interests, copyrights, exclusive rights, royalty interests,
21 contractual rights, or other legally protected interests in these and other works that, give him standing
22 within the class.

23 26. Kevin T. Patton is an individual residing in Missouri who worked as an anatomy &
24 physiology (A&P) professor for several decades, teaching at the high school, community college,
25 undergraduate, and graduate levels. Dr. Patton is President Emeritus of the Human Anatomy &

1 Physiology Society and a founder of the HAPS Institute, a continuing education program for A&P
2 professors. He is also a past President of the Textbook & Academic Authors Association.

3 27. Dr. Patton is the author of multiple successive editions of widely adopted anatomy and
4 physiology textbooks published by Elsevier, Inc., including *Anatomy & Physiology, Structure &*
5 *Function of the Body*, and *The Human Body in Health & Disease*. Dr. Patton authored his textbooks
6 as an employee of Lion Den Inc., an S-Corporation of which he is the president and owner.

7 28. Lion Den Inc. is a Missouri S-Corporation solely owned by the Patton Trust dated May
8 13, 2005, of which Dr. Patton and his wife are both grantors and trustees.

9 29. Lion Den Inc., together with Dr. Patton, have beneficial interests, copyrights, exclusive
10 rights, royalty interests, contractual rights, or other legally protected interests in these and other works
11 that give him and his company standing within the class.

12 30. Dana Loewy is an individual residing in California and is Professor Emerita of Business
13 Communication at Cal State Fullerton.

14 31. Dr. Loewy is the co-author of multiple successive editions of leading business
15 communications textbooks published by Cengage Learning, Inc., including *Business Communication:*
16 *Process & Product* and *Essentials of Business Communication*.

17 32. Dr. Loewy has beneficial interests, copyrights, exclusive rights, royalty interests,
18 contractual rights, or other legally protected interests in these and other works that give her standing
19 within the class.

20 33. Each Plaintiff assigned exclusive publication and distribution rights to his or her
21 respective publisher via a publishing contract with the publisher. The grant of rights is subject to the
22 publisher's contractual obligation to pay royalties to the Plaintiff based on the publisher's exploitation
23 of those exclusive rights. Accordingly, each Plaintiff is the beneficial owner of copyright in his or her
24 work and is entitled to institute an action for infringement of the copyright, pursuant to 17 U.S.C.

§ 501(b). If the publishing contract is terminated, the grant of copyright typically reverts to the Author, who then holds both legal and beneficial copyright.

34. A schedule of named Plaintiffs' representative works is attached hereto as Exhibit A. One or more edition and/or version of each of the representative works has been registered as a literary work in the U.S. Copyright Office. Each of the registrations is valid and subsisting.

35. Defendant Meta is a Delaware corporation with its principal place of business in Menlo Park, California. Meta is alleged to have developed and trained the Llama generative AI system using copyrighted works without authorization.

36. Defendant Mark Zuckerberg is Meta's founder, Chairman, and Chief Executive Officer, and resides in this District. On information and belief, Zuckerberg oversaw Meta's use of pirated shadow-library datasets to train the Llama models notwithstanding internal warnings that those datasets were known to be pirated.

FACTUAL ALLEGATIONS

I. General Allegations

37. Plaintiffs, individually and on behalf of the Class, allege that Defendants reproduced and distributed their copyrighted works without authorization in connection with sourcing content for, developing, and training Meta's Large Language Model Meta AI (hereinafter and inclusive of its later iterations, "Llama") generative AI platform.

38. The first version of Llama was initially released by Meta in February 2023.

39. At the time of this Complaint, Meta maintains five active versions of Llama (Llama 3, Llama 3.1, Llama 3.2, Llama 3.3, and Llama 4).

40. All of these active versions of Llama are available to consumers and businesses for commercial and/or personal usage.

1 41. Like other LLMs, such as ChatGPT and Claude, Llama is a type of artificial
2 intelligence trained on massive text datasets to understand, summarize, translate, and generate human-
3 like language.

4 42. Llama, like other LLMs, functions as the engine behind modern chatbots by predicting
5 probable words in a sequence based on patterns learned from billions of text sources.

6 43. The training of an LLM is critical since the more data an LLM has to train on, the better
7 it is at predicting word patterns.

8 44. Thus, the more data a model is trained on, the more commercial viability it has.

9 45. During initial release of this model, Meta claimed that it used only publicly available
10 material to train Llama.

11 46. To assemble that training data, Meta obtained copyrighted books and other written
12 works through torrenting from pirate “shadow library” sources, including Books3 (a component of
13 “The Pile”), Library Genesis (LibGen), Z-Library, and Sci-Hub, and through unauthorized web-
14 scraped datasets, and then copied those works for use in training the Llama models.

15 47. Meta also obtained copyrighted works from additional shadow-library and bulk
16 sources, including Anna’s Archive, Bibliotik, and Common Crawl.

17 48. Meta deployed its own web crawler, known as SpiderMate, to scrape and copy
18 additional copyrighted works, including books, for use as training data.

19 49. Meta first downloaded books from Library Genesis in or about October 2022 and
20 initially withheld LibGen from use as training data because it knew LibGen to be a well-known
21 repository of pirated books.

22 50. Meta’s growing need for high-quality data led it to approve the use of LibGen to train
23 the Llama models. In or about April 2023, Meta downloaded millions of additional pirated books from
24 LibGen through the BitTorrent protocol using software known as LibTorrent.

51. The scale of Meta’s copying and distribution was massive. The Books3 dataset alone comprised approximately 196,640 books. In the spring of 2024, Meta torrented at least 81.7 terabytes of data across multiple shadow libraries through Anna’s Archive (including at least 35.7 terabytes from Z-Library and LibGen) and it had previously torrented at least 80.6 terabytes from LibGen. And during Meta’s 2022 torrenting from Library Genesis, approximately 4,500 other users in the torrent “swarm” obtained files from Meta. The billing records for a single portion of the cloud services Meta used to torrent these works in 2024 reflected hundreds of thousands of dollars in charges, supplying material support to the very shadow libraries from which Meta took the works.

52. Meta uploaded at least 40 terabytes of books via torrenting.

53. The Llama models include all versions, iterations, and relatives of Llama in any stage of their development and lifecycle, including Llama 1, Llama 2, Llama 3, and Llama 4, and the development of further versions.

54. Meta trained these models on copyrighted works obtained from shadow libraries on a continuing basis throughout the Class Period, including by training Llama 1 between approximately December 2022 and February 2023, Llama 2 between approximately January and July 2023, and Llama 3 for release on or about April 18, 2024, and by continuing to use such works in the development of Llama 4 and subsequent models. This continuing course of conduct spans the period during which Plaintiffs’ works were allegedly copied and distributed.

55. Defendants’ copying occurred in multiple stages, including downloading works, copying works into computer memory, converting works into formats usable by AI systems, and copying works into the training materials used to build the Llama models.

56. By downloading works via torrenting, Meta also made available, uploaded, and distributed copies of those works to other users worldwide, thereby further facilitating copyright infringement.

1 57. Separately, Meta developed and used software to remove copyright-management
2 information (“CMI”), including copyright notices, titles, International Standard Book Numbers
3 (“ISBNs”), and similar identifying information, from books it obtained from shadow libraries. Meta
4 did this to facilitate training and to conceal its use of copyrighted works.

5 58. Specifically, Meta developed and deployed a software script that removed CMI from
6 the beginning and end of works obtained from shadow libraries, including ISBNs, copyright notices,
7 the © symbol, statements of “all rights reserved,” and Digital Object Identifiers (“DOIs”), as well as
8 language indicating that unauthorized copying constituted infringement.

9 59. Beyond stripping CMI, Meta took further steps to conceal its use of copyrighted works,
10 including by addressing instances in which the Llama models, in response to user prompts, generated
11 answers acknowledging that their training data might include pirated, illegal, or copyrighted material;
12 Meta’s personnel introduced curated fine-tuning samples designed to suppress such answers and to
13 reduce the risk that users or the public would learn that Meta had trained Llama on infringing material.

14 60. Defendants’ conduct affected works across a broad range of categories, including
15 fiction, nonfiction, children’s books, poetry, memoirs, dramatic works, scholarly works, as well as
16 textbooks.

17 61. The putative class consists of all textbook authors holding legal or beneficial copyrights
18 in works reproduced, distributed, or otherwise used by Defendants without authorization.

19 62. On information and belief, each copied work is an original work, comprises
20 copyrightable subject matter, is registered with the U.S. Copyright Office, and remains subject to valid
21 copyrights in which class members have a legal or beneficial interest.

22 63. On information and belief, Defendants’ unauthorized use displaced legitimate sales and
23 licenses, deprived authors and publishers of revenues, and usurped the market for licensing
24 copyrighted works for AI training.

64. On information and belief, a market exists for licensing copyrighted works for AI training. AI companies could have tried to negotiate agreements with publishers, media companies, academic institutions, or individual authors for access to books, articles, and journals for model development and training. Meta did not take these actions with respect to the textbooks it copied and distributed.

65. Throughout, Defendants acted willfully and with knowledge that their conduct was unlawful.

66. The unlawful acts alleged against Meta were authorized, ordered, approved, or performed by Meta's officers, directors, employees, agents, and representatives, including Zuckerberg who is ultimately in charge of the management, direction, and control of Meta's business and affairs, and who made the decision or knew the decision was made to undertake the piracy by Meta to train its Llama models.

67. Indeed, Meta employees internally described the sources Meta used as illegal pirated websites and discussed the risk of being treated as accomplices to piracy.

68. Meta cross-referenced pirated titles in its Library Genesis collection against books available for license in order to decide whether to pursue licenses.

69. Meta's own researchers were advised, in discussions concerning the dataset known as The Pile, that the Books3 component contained works under active copyright and would be the principal concern of Meta's legal department.

70. Meta continued to download from shadow libraries that it knew had been repeatedly enjoined or shut down, including as recently as April 2024.

71. Meta publicly acknowledged in its securities filings that it would be susceptible to litigation challenging its alleged use of copyright-protected content to train its AI models.

72. Meta's own representatives have generally stated that books are good data for training LLMs' memories because they are long, consistent, and maintain a particular style and coherent structure.

73. Meta's own representatives have recognized that books also provide high-quality training materials in the sense that they generally are well written and use proper grammar (especially compared to text from the internet, which varies widely on these metrics).

74. The Llama models can generate outputs that substitute for copyrighted works, including outlines, summaries, low-quality alternative versions, imitations, and derivative works. This impacts the market for textbooks in particular, insofar as textbooks are designed to teach and impart knowledge and structures for learning.

75. Using copyrighted textbooks to train an LLM harms the market for those works by helping to enable the rapid generation of countless works that compete with the originals, even if those works themselves are not necessarily infringing. LLMs can generate massive amounts of text in significantly less time that it would take to write that text, and using a fraction of the creativity.

76. Meta has integrated the Llama models into its principal commercial products and services, including Facebook, Instagram, WhatsApp, and its Ray-Ban Meta smart glasses, and has derived substantial commercial benefit from models trained on the copied works.

77. At the same time, Meta asserted and enforced its own copyright in the Llama models, including by issuing a takedown notice directed at a tool used to download a leaked version of Llama 1 while simultaneously copying and distributing Plaintiffs' copyrighted works without authorization.

II. TEXTBOOKS AND TEXTBOOK AUTHORS

78. Textbooks are written works created for the express purpose of teaching academic subjects to primary school, high school, college, graduate, professional, and continuing education students and are marketed, offered, or selected for assignment to students in classroom or course settings.

79. The primary goal of textbooks is to teach the content of a field or subfield in a knowledgeable, coherent, and approachable way that serves the purposes of instructors and students in that field or subfield. Over decades of writing these textbooks, each author develops a unique voice that becomes their distinctive style. A particular textbook is selected and adopted not only because of the title or subject matter but because the author has a style that works for the intended learners.

80. For example, Plaintiff Michael Sullivan's *Precalculus*, now in its 12th Edition, teaches the content of the undergraduate precalculus course, including graphs, functions, and analytic geometry, in a way that is accessible to students and that prepares those students to go on to study calculus.

81. Plaintiff Kenneth S. Saladin's *Anatomy and Physiology: The Unity of Form and Function*, now in its 10th Edition, has achieved its large market share mainly on the strength of a unique style and pedagogic organization not seen in competing textbooks on the same subject.

82. Plaintiffs Zvi Bodie and Alan J. Marcus's *Essentials of Investments*, now in its 13th Edition, presents investment analysis and investment theory to undergraduate business students in a way that is consistent with the approach of the CFA (Chartered Financial Analyst) Institute.

83. Plaintiff Kevin Patton's *Anatomy & Physiology* is designed as an accessible course book, especially for teaching nursing students or others entering the biological sciences, and for instructors desiring a more conversational, narrative approach to the subject matter.

84. Plaintiff Dana Loewy's *Essentials of Business Communication*, now in its 13th edition, teaches best practices in business communication supported by the latest research, with a practical focus on current technology and transferable workplace skills.

85. Most textbook authors are experienced classroom instructors, skilled writers, and subject matter experts in their fields. They each bring their unique combination of knowledge, experience, and skills to writing their textbooks.

1 86. Textbooks in any particular academic area cover the field in a variety of ways,
2 reflecting the unique backgrounds, interests, and styles of their authors, as well as the needs of their
3 intended audience of instructors and students.

4 87. The subject matter that a textbook conveys, whether trigonometric functions, the
5 respiratory system, or the operation of security markets, is not itself protected by copyright. What is
6 protected, and what gives a textbook its value, is the author's selection, organization, and sequencing
7 of material, the examples and explanations that the author develops to effectively convey the
8 information to the student, and the author's voice and pedagogical style.

9 88. Unlike trade books, which are seldom revised, textbooks typically have regular revision
10 cycles to keep their treatment of the subject matter up to date with developments in the field, their
11 pedagogy current, and to stay aligned with instructor preferences, student needs, and adoption
12 standards. Each revision requires a substantial investment of time, research, and resources by the
13 author. One result of illegal training on pirated textbooks, and market substitution of students who turn
14 to artificial intelligence alternatives, is that students can receive inferior, incomplete, or out-of-date
15 guidance—but which appears to be authoritative. Searches on Llama for the named Plaintiffs and their
16 publications resulted in exactly such responses, providing feedback to users that purports to be “in the
17 style of” or “taking the approach of” specifically named textbook authors.

18 89. Many textbooks, including those authored by Plaintiffs, have lives that extend over
19 many editions and many decades. These textbooks become a central part of the authors' life work,
20 commitments, reputations, and livelihoods.

21 90. Although different textbooks in any field have different emphases, viewpoints, and
22 styles, successful textbooks share a reputation in their fields, and in particular among the adopters and
23 instructors who select them for use in the classrooms, for accessibility, reliability, and effectiveness.

24 91. Textbooks play an important role in shaping the academic fields and subfields that they
25 cover, and in shaping the education of the students who use them.

92. Because of the importance of textbooks in academia, many textbook authors establish their professional reputation not only on the basis of their teaching and research, but also on the quality, impact, and adoptions of their textbooks.

III. THE TEXTBOOK MARKET

93. The markets for textbooks, on the one hand, and trade books (both fiction and non-fiction), on the other, are fundamentally different. The trade book market consists primarily of people who choose to purchase books for personal reasons, e.g., for their own edification, curiosity, or entertainment. A casual reader looking for a novel by J.K. Rowling or a popular history book by David McCullough will not accept a substitute by another author or generated by an LLM.

94. In the textbook market, by contrast, the final purchaser is *not* the person who selects the book. The market for textbooks is driven by adoptions, i.e., instructors, schools, or educational systems who consider the textbook options available in a specific subject area and then “adopt” the one they think best suits the needs of their instructors and students. The textbooks are then assigned to students in specific courses.

95. Purchasing decisions are fundamentally different in the textbook market than in the trade book market because the person who *chooses* the textbook is not the same person who is *required* to purchase it. The purchaser at the elementary or secondary school level is typically a school system or individual school required to make the purchase and is continually looking for ways to cut costs from its budget. In the higher education space, the purchaser is the student assigned the textbook by their instructor. In either situation, the purchaser has a powerful incentive to avoid the cost of purchasing a textbook if a “free” version of that same book, or a “good enough” substitute, can be found via an LLM.

IV. DEFENDANTS’ USE OF TEXTBOOKS TO TRAIN THEIR LLAMA MODELS THREATENS TO DISPLACE THE TEXTBOOK MARKET.

96. Because textbooks cover academic fields and subfields in engaging, effective and up-

1 to-date ways, they are uniquely valuable in the training of LLMs.

2 97. On information and belief, textbooks constitute a large proportion of the books that
3 Defendants used to train their Llama models.

4 98. Defendants' Llama models trained on sets of textbooks in various fields are capable of
5 feeding back to users comprehensive information about those fields, often with detailed references to
6 specific authors and their organization of and approach to the subject matter.

7 99. Textbooks are required purchases by students (in higher education) and schools (in K-
8 12 settings). Students, instructors, and schools feel strong pressures to reduce the cost of educational
9 materials. As a result, textbooks are uniquely susceptible to displacement by LLMs trained on
10 textbooks in that field.

11 100. As LLMs become more powerful and widespread, and if they are allowed to
12 continually update their content by ingesting new editions of textbooks as they come out, instructors,
13 schools, and adoption and curriculum committees may consider dropping textbook adoptions
14 altogether in favor of substitutes that rely on the use of LLM models.

15 101. LLMs threaten the textbook market by destroying the incentives for textbook authors
16 to continue writing and revising high quality textbooks, to the detriment of the textbook authors on
17 whose work the LLMs rely and the student learners whose education may be compromised by their
18 reliance on outputs.

19 102. Llama and other generative AI agents threaten to displace human-authored textbooks
20 with cheaper AI-generated textbooks or textbook equivalents, at the expense of losing the valuable
21 authorial voice, research, expertise, and perspective provided by textbook authors.

22 103. On information and belief, licensing markets exist, or could be developed, that would
23 permit Defendants to make use of Plaintiffs' textbooks in exchange for fair compensation. Defendants
24 bypassed such licensing efforts by obtaining Plaintiffs' textbooks without license, thus threatening to
25 displace the very sources on which they relied.

104. On information and belief, when Defendants used massive numbers of books without permission or compensation to their authors, they were aware that such use would likely displace and dilute the market for the works used in the training.

V. Plaintiffs' Claims Are Timely

105. Plaintiffs and the members of the Class did not know, and in the exercise of reasonable diligence could not have known, that Meta had copied their textbooks until the filing of *Kadrey v. Meta Platforms, Inc.* on July 7, 2023, first publicly disclosed and made identifiable Meta's copying of copyrighted books to train Llama.

106. Meta concealed the conduct at the heart of these claims (the acquisition of works by torrenting the LibGen and Anna's Archive shadow libraries and the distribution of those works by seeding, and its removal of copyright-management information) until its own internal records were produced and unsealed in this District in January 2025.

107. This action is therefore brought within three years after Plaintiffs' claims accrued under 17 U.S.C. § 507(b), and Meta's infringing conduct has in any event continued throughout the Class Period.

CLASS ACTION ALLEGATIONS

108. Plaintiffs bring this action under Federal Rules of Civil Procedure 23(a), 23(b)(2), and 23(b)(3) on behalf of the following Class: All textbook authors who hold legal or beneficial ownership of a United States copyright in a textbook that Meta downloaded, copied, reproduced, uploaded, and/or otherwise distributed via BitTorrent, or otherwise used to train its Llama models between July 7, 2020, and the present (the "Class Period"), and that was registered with the United States Copyright Office (i) within five years of the work's first publication and before Meta's copying, or (ii) within three months of the work's first publication.

109. For the purposes of the Class, a textbook is defined as a book in any medium, including printed and digital, created for the express purpose of teaching academic subject matter to primary

1 school, secondary school, college, university, graduate, professional, or continuing education students
2 and marketed, offered, or selected for educational instruction in a classroom or course setting.

3 110. Excluded from the Class are Defendants; their officers, directors, employees, parents,
4 subsidiaries, and affiliates; all governmental entities; and the judge and chambers staff assigned to this
5 case and their immediate families.

6 111. The Class is so numerous that joinder of all members is impracticable. On information
7 and belief, textbook authors number in at least the thousands, and textbooks constitute approximately
8 twenty-five percent of the works within the broader population of works Meta copied.

9 112. Questions of law and fact common to the Class predominate over any individual
10 questions, including whether Meta reproduced the Class's textbooks, whether Meta distributed them
11 by torrenting, whether Meta removed copyright-management information, whether Meta's conduct
12 was willful, whether such use infringed the Class's exclusive rights, whether Meta's use harmed the
13 market for textbooks through substitution or dilution, and the appropriate relief.

14 113. Plaintiffs' claims are typical of the Class's claims because Plaintiffs and the Class are
15 textbook authors whose registered textbook works were copied, distributed, and used by Meta in the
16 same manner and who suffered the same type of harm.

17 114. Plaintiffs will fairly and adequately protect the interests of the Class. They are textbook
18 authors with no conflicts with other Class members, and they have retained counsel experienced in
19 copyright and class-action litigation, including attorneys at Slarskey LLC and Archstone Law P.C.
20 who have represented textbook authors for many years.

21 115. A class action is superior to other available methods for fairly and efficiently
22 adjudicating this controversy, and Defendants have acted on grounds generally applicable to the Class,
23 making final relief appropriate for the Class as a whole.

116. Plaintiffs' counsel, Slarskey LLC, Archstone Law P.C., and Gaw | Poe LLP, are qualified and experienced in representing textbook authors and litigating class actions and should be appointed Class Counsel under Federal Rule of Civil Procedure 23(g).

CAUSES OF ACTION

COUNT I — Direct Copyright Infringement (17 U.S.C. §§ 106(1), 106(2), 106(3), 501) against all Defendants

117. Plaintiffs incorporate by reference each of the foregoing paragraphs.

118. Plaintiffs and the Class own or hold beneficial interests in the exclusive rights in their registered textbook works under 17 U.S.C. § 106, including the rights of reproduction and distribution, and preparation of derivative works.

119. Meta reproduced Plaintiffs' and the Class's textbook works without authorization by downloading them from shadow libraries, copying them into datasets and computer memory, converting them into training-ready formats, and copying them into the materials used to train the Llama models.

120. Meta further distributed Plaintiffs' and the Class's authored textbooks without authorization by uploading and making them available to others through the BitTorrent protocol while torrenting them.

121. Zuckerberg oversaw and directed Meta to undertake the behavior described in Count I.

122. Meta's infringement was willful. Plaintiffs and the Class are entitled to statutory damages, or at their election actual damages and Meta's profits, together with injunctive relief, costs, and attorneys' fees under 17 U.S.C. §§ 504, 505.

COUNT II — Contributory Copyright Infringement against all Defendants

123. Plaintiffs incorporate by reference each of the foregoing paragraphs.

124. By uploading and making available Plaintiffs' and the Class's works to other users in BitTorrent "swarms" as it torrented them, Meta induced, caused, and materially contributed to the reproduction and distribution of those works by thousands of third parties.

125. Meta knew or had reason to know that its conduct furthered the infringement of Plaintiffs' and the Class's works, and Meta could have avoided distributing the works but chose not to. Meta is liable for contributory copyright infringement, and its conduct was willful.

126. Zuckerberg oversaw and directed Meta to undertake the behavior described in Count II.

127. Plaintiffs and the Class are entitled to statutory damages, or at their election actual damages and Meta's profits, together with injunctive relief, costs, and attorneys' fees.

**COUNT III — Removal of Copyright-Management Information (17 U.S.C. § 1202(b))
against all Defendants**

128. Plaintiffs incorporate by reference each of the foregoing paragraphs.

129. Plaintiffs' and the Class's textbook works contained copyright-management information, including copyright notices, the © symbol, titles, ISBNs, and the names of the authors and rights holders.

130. Meta developed and deployed software to remove that copyright-management information from the works it obtained from shadow libraries, and did so intentionally, knowing or having reasonable grounds to know that doing so would conceal and facilitate infringement, in violation of 17 U.S.C. § 1202(b).

131. Zuckerberg oversaw and directed Meta to undertake the behavior described in Count III.

132. Plaintiffs and the Class are entitled to statutory or actual damages under 17 U.S.C. § 1203, together with injunctive relief, costs, and attorneys' fees.

**COUNT IV — Distribution of Copyright-Management Information (17 U.S.C. § 1202(b))
against all Defendants**

133. Plaintiffs incorporate by reference each of the foregoing paragraphs.

134. Plaintiffs' and Class's textbook works contained copyright-management information, including copyright notices, the © symbol, titles, ISBNs, and the names of the authors and rights holders.

135. Meta distributed copies of the works (via BitTorrent seeding to swarm peers) knowing CMI had been removed from the Plaintiffs' and Class's works without authority of the copyright owner or the law.

136. The distributed copies were identical to the originals and were more accessible to other pirates because of Meta's seeding of Plaintiffs' and Class's works.

137. Meta knew or had reason to know the distribution would induce, enable, facilitate, or conceal infringement.

138. Zuckerberg oversaw and directed Meta to undertake the behavior described in Count IV.

139. Plaintiffs and the Class are entitled to statutory or actual damages under 17 U.S.C. § 1203, together with injunctive relief, costs, and attorneys' fees.

PRAYER FOR RELIEF

WHEREFORE, Plaintiffs, individually and on behalf of the Class, respectfully request that the Court enter judgment and grant relief as follows:

- a. Certifying this action as a class action, appointing Plaintiffs as Class Representatives, and appointing Slarskey LLC, Archstone Law P.C., and Gaw | Poe LLP as Class Counsel;
- b. Entering judgment that Defendants infringed Plaintiffs' and the Class's copyrights and violated 17 U.S.C. § 1202(b);
- c. Awarding statutory damages, or at Plaintiffs' and the Class's election, actual damages and Defendants' profits, under 17 U.S.C. §§ 504 and 1203;

- 1 d. Awarding declaratory and permanent injunctive relief, including relief restraining
2 further infringement and requiring restoration of copyright-management information
3 under 17 U.S.C. § 1203(b)(1);
- 4 e. Awarding pre- and post-judgment interest, costs, and reasonable attorneys' fees as
5 permitted by law; and
- 6 f. Granting such other and further relief as the Court deems just and proper.

7 Plaintiffs demand a trial by jury on all claims and issues so triable.

8 Date: July 2, 2026

9 **SLARSKEY LLC**

10 By: /s/ David Slarskey

David N. Slarskey *pro hac vice application pending*

11 dslarskey@slarskey.com

Adam Hollander (*pro hac vice application pending*)

12 Deepa Devanathan (*pro hac vice application pending*)

Phillip Gasperetti (*pro hac vice application pending*)

13 767 Third Ave, 14th Floor

New York, NY 10017

14 Tel: (212) 658-0661

15 **ARCHSTONE LAW GROUP, P.C.**

Brenda Ulrich (*pro hac vice pending*)

16 Zick Rubin (*pro hac vice pending*)

bulrich@archstonelaw.com

17 zrubin@archstonelaw.com

Riverside Center

18 275 Grove Street, Suite 2-400

Newton, MA 02466

19 Tel: (781) 314-0100

20 **GAW POE LLP**

21 
Randolph Gaw (S.B. #223718)

22 rgaw@gawpoe.com

One Embarcadero Ctr., Suite 1200

23 SAN FRANCISCO, CA 94111

Telephone: (415) 766-7451

24 Facsimile: (415) 737-0642

Counsel for Plaintiffs and the Proposed Class