

DANIEL J. MULLER, SBN 193396  
dmuller@venturahersey.com  
VENTURA HERSEY & MULLER, LLP  
1506 Hamilton Avenue  
San Jose, California 95125  
Telephone: (408) 512-3022  
Facsimile: (408) 512-3023

*Attorneys for Plaintiff and the Class*

**UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA – SAN FRANCISCO DIVISION**

MICHAEL CHABON, TA-NEHISI  
COATES, JUNOT DÍAZ, ANDREW  
SEAN GREER, DAVID HENRY  
HWANG, MATTHEW KLAM, LAURA  
LIPPMAN, RACHEL LOUISE SNYDER,  
AYELET WALDMAN, AND  
JACQUELINE WOODSON,

individually and on behalf of all others  
similarly situated,

Plaintiffs,

v.

META PLATFORMS, INC., a Delaware  
Corporation,

Defendant.

Case No. 3:23-cv-04663-VC

**FIRST AMENDED CLASS ACTION  
COMPLAINT**

CLASS ACTION

**JURY TRIAL DEMANDED**

1 Plaintiffs Michael Chabon, Ta-Nehisi Coates, Junot Díaz, Andrew Sean Greer, David  
2 Henry Hwang, Matthew Klam, Laura Lippman, Rachel Louise Snyder, Ayelet Waldman, and  
3 Jacqueline Woodson (“Plaintiffs”), individually and on behalf of all others similarly situated,  
4 bring this class action against Defendant Meta Platforms, Inc. Plaintiffs allege as follows based  
5 upon personal knowledge as to themselves and their own acts, and upon information and belief  
6 as to all other matters:

7 **NATURE OF ACTION**

8 1. This is a class action lawsuit brought by Plaintiffs on behalf of themselves and a  
9 Class of authors holding copyrights in their published works arising from Meta’s clear  
10 infringement of their intellectual property.

11 2. Meta’s LLaMA (Large Language Model Meta AI) is a set of large language  
12 models created and maintained by Meta Platforms, Inc. A *large language model* is an AI  
13 software program designed to produce convincingly natural texts outputs in response user  
14 prompts.

15 3. Rather than being programmed in the traditional manner, a large language model  
16 is “trained” by copying massive amounts of text and extracting expressive information from it.  
17 The body of text is referred to as the *training dataset*.

18 4. Accordingly, a large language model’s output is therefore entirely and uniquely  
19 reliant on the material in its training dataset. Every time it assembles a text output, the model  
20 relies on the information it extracted from its training dataset. Therefore, the decisions about the  
21 textual information it includes in the training dataset are deliberate and important choices.

22 5. Plaintiffs and Class members are authors of books, screenplays, novels, and other  
23 written works. Plaintiffs and Class members possess copyrights for the books and written works  
24 they created and published. Plaintiffs and Class members did not consent to the use of their  
25 copyrighted books as training materials for LLaMA.

26 6. Nevertheless, their copyright protected works were copied and ingested as part  
27 of training LLaMA. Plaintiffs’ copyrighted books appear in the dataset that Meta has admitted  
28 to using to train LLaMA.



1           11. Plaintiff Ta-Nehisi Coates (“Plaintiff Coates”) is a resident of New York.  
2 Plaintiff Coates is an author who owns registered copyrights in many works, including but not  
3 limited to, *Between the World and Me*, *The Beautiful Struggle*, *The Water Dancer*, and *We Were*  
4 *Eight Years in Power*. Plaintiff Coates is the recipient of the National Book Award and a  
5 MacArthur Fellowship. Plaintiff Coates’s works include copyright-management information  
6 that provides information about the copyrighted work, including the title of the work, its ISBN  
7 or copyright registration number, the name of the author, and the year of publication.

8           12. Plaintiff Junot Díaz (“Plaintiff Diaz”) is a resident of Massachusetts. Plaintiff  
9 Díaz is an author who owns registered copyrights in many works, including but not limited to,  
10 *Islandborn*, *The Brief Wondrous Life of Oscar Wao*, *This is How You Lose Her*, and *Drown*.  
11 Plaintiff Díaz is the recipient of the Pulitzer Prize for Fiction, the National Book Critics Circle  
12 Award, a MacArthur “Genius” Fellowship, PEN/Malamud Award, Dayton Literary Peace Prize,  
13 Guggenheim Fellowship, and PEN/O. Henry Award. Plaintiff Díaz is the Rudge and Nancy  
14 Allen Professor of Writing at the Massachusetts Institute of Technology. Plaintiff Díaz’s works  
15 include copyright-management information that provides information about the copyrighted  
16 work, including the title of the work, its ISBN or copyright registration number, the name of the  
17 author, and the year of publication.

18           13. Andrew Sean Greer (“Plaintiff Greer”) is a resident of California. Plaintiff Greer  
19 is an author who owns registered copyrights in many works, including but not limited to, *The*  
20 *Confessions of Max Tivoli*, *How it Was for Me*, *The Story of a Marriage*, *The Impossible Lives*  
21 *of Greta Wells*, *The Path of Minor Planets*, *Less is Lost*, and *Less, a novel*. Plaintiff Greer is the  
22 recipient of the Pulitzer Prize for Fiction, a Guggenheim Fellowship, and an NEA grant. Plaintiff  
23 Greer’s works include copyright-management information that provides information about the  
24 copyrighted work, including the title of the work, its ISBN or copyright registration number, the  
25 name of the author, and the year of publication.

26           14. Plaintiff David Henry Hwang (“Plaintiff Hwang”) is a resident of New York.  
27 Plaintiff Hwang is a playwright and screenwriter who owns registered copyrights in several  
28 plays, including but not limited to, *M. Butterfly*, *Chinglish*, *Yellow Face*, *Golden Child*, *the*

1 *Dance and the Railroad*, and *FOB*, as well as the Broadway musicals *Aida*, *Flower Drum Song*  
2 (2002 revival). Plaintiff Hwang is a Tony Award winner and three-time nominee, a Grammy  
3 Award winner and two time nominee, a three-time OBIE Award winner, and a three-time finalist  
4 for the Pulitzer Prize in Drama. Plaintiff Hwang's works include copyright-management  
5 information that provides information about the copyrighted work, including the title of the  
6 work, its ISBN or copyright registration number, the name of the author, and the year of  
7 publication.

8 15. Plaintiff Matthew Klam ("Plaintiff Klam") is a resident of Washington D.C.  
9 Plaintiff Klam is an author who owns registered copyrights in several works, including but not  
10 limited to, *Who is Rich?*, and *Sam the Cat and Other Stories*. Plaintiff Klam is a recipient of a  
11 Guggenheim Fellowship, a Robert Bingham/PEN Award, a Whiting Writer's Award, and a  
12 National Endowment of the Arts. Plaintiff Klam's works have been selected as Notable Books  
13 of the year by *The New York Times*, *The Los Angeles Times*, *the Kansas City Star*, and *the*  
14 *Washington Post*. His work has appeared in *The New York Times*, *The New Yorker*, *Harper's*  
15 *Magazine*, and elsewhere. Plaintiff Klam's works include copyright-management information  
16 that provides information about the copyrighted work, including the title of the work, its ISBN  
17 or copyright registration number, the name of the author, and the year of publication.

18 16. Plaintiff Laura Lippman ("Plaintiff Lippman") is a resident of Maryland. Plaintiff  
19 Lippman is an author who owns registered copyrights in several works, including but not limited  
20 to, *I'd Know you Anywhere*, *What the Dead Know*, *Hardly Knew Her*, *Another Thing to Fall*,  
21 *Butchers Hill*, and *Baltimore Blues*. Plaintiff Lippman is a *New York Times* bestseller and the  
22 recipient of the Edgar Award, the Anthony Award, the Agatha Award, the Quill Award, the  
23 Shamus Award, and the Nero Wolfe Award. Plaintiff Lippman's works include copyright-  
24 management information that provides information about the copyrighted work, including the  
25 title of the work, its ISBN or copyright registration number, the name of the author, and the year  
26 of publication.

1           17. Plaintiff Rachel Louise Snyder (“Plaintiff Snyder”) is a resident of Washington,  
2 D.C. Plaintiff Snyder is an author who owns registered copyrights in several works, including  
3 but not limited to, *Women We Buried*, *Women We Burned*, *No Visible Bruises – What We Don’t*  
4 *Know About Domestic Violence Can Kill Us*, *What We’ve Lost is Nothing*, and *Fugitive Denim:*  
5 *A Moving Story of People and Pants in the Borderless World of Global Trade*. Plaintiff Snyder  
6 is the recipient of the J. Anthony Lukas Work-in-Progress Award, the Hillman Prize, and the  
7 Helen Bernstein Book Award, and finalist for the National Book Critics Circle Award, *Los*  
8 *Angeles Times* Book Prize, and Kirkus Award. Her work has appeared in *The New*  
9 *Yorker*, *The New York Times*, *Slate*, and elsewhere. Plaintiff Snyder’s works include copyright-  
10 management information that provides information about the copyrighted work, including the  
11 title of the work, its ISBN or copyright registration number, the name of the author, and the year  
12 of publication.

13           18. Plaintiff Ayelet Waldman (“Plaintiff Waldman”) is a resident of California.  
14 Plaintiff Waldman is an author and screen and television writer who owns registered copyrights  
15 in several works, including but not limited to, *Love and other Impossible Pursuits*, *Red Hook*  
16 *Road*, *Love and Treasure*, *Bad Mother*, *Daughter’s Keeper*, *A Really Good Day*, and *Mommy*  
17 *Track Mysteries*. Plaintiff Waldman has been nominated for an Emmy and Golden Globe and is  
18 the recipient of numerous awards including a Peabody, AFI award, and a Pen Award, among  
19 others. Plaintiff Waldman’s works include copyright-management information that provides  
20 information about the copyrighted work, including the title of the work, its ISBN or copyright  
21 registration number, the name of the author, and the year of publication.

22           19. Plaintiff Jacqueline Woodson (“Plaintiff Woodson”) is a resident of New York.  
23 Plaintiff Woodson is an author who owns registered copyrights in many works, including but  
24 not limited to, *Brown Girl Dreaming*, *After Tupac & D Foster*, *Harbor Me*, *Before the Ever*  
25 *After*, *Day You Begin*, and *Another Brooklyn*. Plaintiff Woodson is the recipient of the  
26 MacArthur “Genius” Award, a Guggenheim Fellowship, the National Book Award, a Caldecott  
27 Honor Medal, the Coretta Scott King Award, the Newbery Honor Medal, and the Hans Christian  
28 Andersen Award, among others. Plaintiff Woodson’s works include copyright-management

1 information that provides information about the copyrighted work, including the title of the  
2 work, its ISBN or copyright registration number, the name of the author, and the year of  
3 publication.

4 20. At all times relevant hereto, Plaintiffs have been and remain the holders of the  
5 exclusive rights under the Copyright Act of 1976 (17 U.S.C. §§ 101, *et seq.* and all amendments  
6 thereto) to reproduce, distribute, display, or license the reproduction, distribution, and/or display  
7 the works identified in paragraphs 10-19, *supra*.

8 **B. Defendant**

9 21. Defendant Meta is a Delaware corporation with its principal place of business at  
10 1601 Willow Road, Menlo Park, California 94025.

11 **AGENTS AND CO-CONSPIRATORS**

12  
13 22. The unlawful acts alleged against Meta in this class action complaint were  
14 authorized, ordered, or performed by the Defendant's respective officers, agents, employees,  
15 representatives, or shareholders while actively engaged in the management, direction, or control  
16 of the Defendant's businesses or affairs. The Defendant's agents operated under the explicit and  
17 apparent authority of their principals. Each Defendant, and its subsidiaries, affiliates, and agents  
18 operated as a single unified entity.

19 23. Various persons and/or firms not named as Defendants may have participated as  
20 co-conspirators in the violations alleged herein and may have performed acts and made  
21 statements in furtherance thereof. Each acted as the principal, agent or joint venture of, or for  
22 other Defendants with respect to the acts, violations, and common course of conduct alleged  
23 herein.

1  
2 **FACTUAL ALLEGATIONS**

3 **A. Meta Platform’s Artificial Intelligence Products**

4 24. Meta creates, markets, and sells software and hardware technology products,  
5 including Facebook, Instagram, and Horizon Worlds. Meta also has a large artificial-intelligence  
6 group called Meta AI that creates and distributes artificial-intelligence software products.

7 25. AI software is designed to algorithmically simulate human reasoning or  
8 inference, often based upon statistical models or methods.

9 26. In February 2023, Meta released an AI product called LLaMA. LLaMA is a set  
10 of large language models. A large language model (or “LLM” for short) is AI software designed  
11 to parse and emit natural language. Though a large language model is a software program, it is  
12 not created the way most software programs are—that is, by human software engineers writing  
13 code. Rather, a large language model is “trained” by copying massive amounts of text from  
14 various sources and feeding these copies into the model. This corpus of input material is called  
15 the training dataset. During training, the large language model copies each piece of text in the  
16 training dataset and extracts expressive information from it. The large language model  
17 progressively adjusts its output to more closely resemble the sequences of words copied from  
18 the training dataset. Once the large language model copies and ingests the all of this text, it is  
19 able to generate and produce convincing simulations of natural written language as it appears in  
20 the training dataset.

21 27. Much of the material in Meta’s training dataset, however, comes from  
22 copyrighted works—including works written by Plaintiffs—that were copied by Meta without  
23 consent, without credit, and without compensation.

24 28. Plaintiffs’ published written works contain certain copyright management  
25 information. This information includes the written work’s title, the ISBN number or copyright  
26 number, the author’s name the copyright holder’s name, and terms and conditions of use.

27 29. Meta introduced LLaMA in a paper called “LLaMA: Open and Efficient  
28 Foundation Language Models”. In the paper, Meta describes the LLaMA training dataset as “a

1 large quantity of textual data” that was chosen because it was “publicly available, and  
2 compatible with open sourcing.”

3 30. *Open sourcing* refers to putting data under a permissive style of copyright license  
4 called an *open-source license*. Copyrighted materials, however, are not ordinarily “compatible  
5 with open sourcing” unless and until the copyright owner first places the material under an open-  
6 source license, thereby enabling others to do so later.

7 31. In a table describing the composition of the LLaMA training dataset, Meta notes  
8 that 85 gigabytes of the training data comes from a category called “Books.” Meta further  
9 elaborates that “Books” comprises the text of books from two internet sources: (1) Project  
10 Gutenberg, an online archive of approximately 70,000 books that are out of copyright, and (2)  
11 “the Books3 section of ThePile . . . a publicly available dataset for training large language  
12 models.” Meta’s paper on LLaMA does not further describe the contents of Books3 or ThePile.

13 32. In a table describing the composition of the LLaMA training dataset, Meta notes  
14 that 85 gigabytes of the training data comes from a category called “Books.” Meta further  
15 elaborates that “Books” comprises the text of books from two internet sources: (1) Project  
16 Gutenberg, an online archive of approximately 70,000 books that are out of copyright, and (2)  
17 “the Books3 section of ThePile . . . a publicly available dataset for training large language  
18 models.” Meta’s paper on LLaMA does not further describe the contents of Books3 or ThePile.

19 33. But that information is available elsewhere. ThePile is a dataset assembled by a  
20 research organization called EleutherAI. In December 2020, EleutherAI introduced this dataset  
21 in a paper called “The Pile: An 800GB Dataset of Diverse Text for Language Modeling”.

22 34. The EleutherAI paper reveals that the Books3 dataset comprises 108 gigabytes  
23 of data, or approximately 12% of the dataset, making it the third largest component of The Pile  
24 by size.

25 35. The EleutherAI paper describes the contents of Books3:

26 Books3 is a dataset of books derived from a copy of the contents of the  
27 Bibliotik private tracker . . . Bibliotik consists of a mix of fiction and  
28 nonfiction books and is almost an order of magnitude larger than our next  
largest book dataset (BookCorpus2). We included Bibliotik because

1 books are invaluable for long-range context modeling research and  
2 coherent storytelling.

3 36. Bibliotik is one of a number of notorious “shadow library” websites that also  
4 includes Library Genesis (aka LibGen), Z-Library (aka B-ok), and Sci-Hub. The books and other  
5 materials aggregated by these websites have also been available in bulk via torrent systems.  
6 These shadow libraries have long been of interest to the AI-training community because of the  
7 large quantity of copyrighted material they contain. For that reason, these shadow libraries are  
8 also flagrantly illegal.

9 37. The person who assembled the Books3 dataset has confirmed in public  
10 statements that it represents “all of Bibliotik” and contains 196,640 books. EleutherAI currently  
11 distributes copies of Books3 through its website (<https://pile.eleuther.ai/>).

12 38. The Books3 dataset is also available from a popular AI project hosting service  
13 called Hugging Face ([https://huggingface.co/datasets/the\\_pile\\_books3](https://huggingface.co/datasets/the_pile_books3)).

14 39. Many of Plaintiffs’ written works appear in the Books3 dataset, these written  
15 works are referred to as the Infringed Works.

16 40. For example, Books3 contains a significant amount of Plaintiff Chabon’s works,  
17 including, but not limited to, *The Final Solution*, *Bookends: Collected Intros and Outros*,  
18 *Kingdom of Olives and Ash*, *Manhood for Amateurs: The Pleasures and Regrets of a Husband*,  
19 *Father, and Son*, *Maps and Legends*, *McSweeney’s Mammoth Treasury of Thrilling Tales*,  
20 *Werewolves in Their Youth*, *Moonglow*, *Pops Fatherhood in Pieces*, *The Amazing Adventures*  
21 *of Kavalier & Clay*, and *The Yiddish Policemen’s Union*.

22 41. Books3 also contains Plaintiff Coates’s works, including, but not limited to *The*  
23 *Water Dancer*, *We Were Eight Years in Power*, and *The Beautiful Struggle*.

24 42. Likewise, Books3 contained Plaintiff Díaz’s works, including, but not limited to,  
25 *Drown*, *Asi es como la pierdes*, *La breve y maravillosa vida de Oscar Wao*, *Los boys*, and  
26 *Negocios*.

27 43. Plaintiff Greer’s works likewise are found in Books3, including but not limited  
28 to *How it Was for Me*, *Less*, *The Confessions of Max Tivoli*, and the *Path of Minor Planets*.

1           44. Books3 similarly contains Plaintiff Hwang’s written works, including, but not  
2 limited to, *Golden Child*, *M. Butterfly*, and *Trying to Find Chinatown*.

3           45. Plaintiff Klam’s works are similarly found in the Books3 dataset, including, but  
4 not limited to, *Who is Rich?* and *Sam the Cat*.

5           46. Book3 also contains Plaintiff Lippman’s written works, including, but not  
6 limited to, *After I’m Gone*, *In a Strange City*, *Lady in the Lake*, *Sunburn*, *What the Dead know*,  
7 and *Wilde Lake*.

8           47. Plaintiff Snyder’s works also are contained in the Books3 dataset, including, but  
9 not limited to, *No Visible Bruises: What We Don’t Know about Domestic Violence Can Kill Us*.

10           48. In the same vein, Plaintiff Waldman’s works appear in the Books3 dataset,  
11 including, but not limited to, *A Really Good Day*, *Bad Mother*, *Love and Other Impossible*  
12 *Pursuits*, and *Love and Treasure*.

13           49. Plaintiff Woodson’s written works also appear in the Books3 dataset, including,  
14 but not limited to, *Another Brooklyn*, *Behind You*, *Beneath a Meth Moon*, *Brown Girl Dreaming*,  
15 *Feathers*, *Harbor Me*, *Hush*, *If You Come Softly*, *Red at the Bone*, and *Miracle Boys*.

16           50. Since the launch of the LLaMA language models in February 2023, Meta has  
17 made these models selectively available to organizations that request access, saying:

18           To maintain integrity and prevent misuse, we are releasing our model  
19 under a noncommercial license focused on research use cases. Access to  
20 the model will be granted on a case-by-case basis to academic  
21 researchers; those affiliated with organizations in government, civil  
22 society, and academia; and industry research laboratories around the  
23 world. People interested in applying for access can find the link to the  
24 application in our research paper.

25           51. Meta has not disclosed what criteria it uses to decide who is eligible to receive  
26 the LLaMA language models, nor who has actually received them, or whether Meta has in fact  
27 adhered to its stated criteria. On information and belief, Meta has in fact distributed the LLaMA  
28 models to certain people and entities, continues to do so, and has benefited financially from  
these acts.



1 questions affecting individual Class members, including, without limitation:

- 2 a. Whether Defendant violated the copyrights of Plaintiffs and the Class when they  
3 downloaded copies of Plaintiffs' and the Class's Infringed Works and used them  
4 to train the LLaMA language models;
- 5 b. Whether the LLaMA language models are themselves infringing derivative  
6 works based on Plaintiffs' and the Class's Infringed Works;
- 7 c. Whether the text outputs of the LLaMA language models are infringing  
8 derivative works based on Plaintiffs' Infringed Works;
- 9 d. Whether Defendant violate the DMCA by removing copyright-management  
10 information from Plaintiffs' and the Class's Infringed Works;
- 11 e. Whether Defendant was unjustly enriched by the unlawful conduct alleged  
12 herein;
- 13 f. Whether Defendant's conduct allege herein constitutes Unfair Competition under  
14 California Business and Professions Code Secion 17200 *et seq.*
- 15 g. Whether Defendant's conduct alleged herein constitute common unfair  
16 competition;
- 17 h. Whether any affirmative defense excuses Defendant's conduct;
- 18 i. Whether any statutes of limitation limits Plaintiffs' and the Class's potential for  
19 recovery;
- 20 j. Whether Plaintiffs and the other Class members are entitled to equitable relief,  
21 including, but not limited to, restitution or injunctive relief; and
- 22 k. Whether Plaintiffs and the other Class members are entitled to damages and other  
23 monetary relief and, if so, in what amount.

24 48. Typicality. Federal Rule of Civil Procedure 23(a)(3): Plaintiffs' claims are  
25 typical of the other Class members' claims because, among other things, all Class members were  
26 comparably injured through Defendant's wrongful conduct as described above.

27 49. Adequacy. Federal Rule of Civil Procedure 23(a)(4): Plaintiffs are adequate  
28 Class representative because their interests do not conflict with the interests of the other

1 members of the Class they seeks to represent; Plaintiffs have retained counsel competent and  
2 experienced in complex class action litigation; and Plaintiffs intend to prosecute this action  
3 vigorously. The interests of the Class will be fairly and adequately protected by Plaintiffs and  
4 their counsel.

5 50. Declaratory and Injunctive Relief. Federal Rule of Civil Procedure 23(b)(2):  
6 Defendants have acted or refused to act on grounds generally applicable to Plaintiffs and the  
7 other members of the Class, thereby making appropriate final injunctive relief and declaratory  
8 relief with respect to the Class as a whole.

9 51. Superiority. Federal Rule of Civil Procedure 23(b)(3): A class action is superior  
10 to any other available means for the fair and efficient adjudication of this controversy, and no  
11 unusual difficulties are likely to be encountered in the management of this class action. The  
12 damages or other financial detriment suffered by Plaintiffs and the other Class members are  
13 relatively small compared to the burden and expense that would be required to individually  
14 litigate their claims against Defendants, so it would be impracticable for the members of the  
15 Class to individually seek redress for Defendant's wrongful conduct. Even if Class members  
16 could afford individual litigation, the court system could not. Individualized litigation creates a  
17 potential for inconsistent or contradictory judgments, and increases the delay and expense to all  
18 parties and the court system. By contrast, the class action device presents far fewer management  
19 difficulties, and provides the benefits of single adjudication, economy of scale, and  
20 comprehensive supervision by a single court.

21  
22  
23 **CAUSES OF ACTION**

24 **FIRST CAUSE OF ACTION**

25 **DIRECT COPYRIGHT INFRINGEMENT,**

26 **17 U.S.C. § 106, *et seq.***

27 52. Plaintiffs hereby incorporate by reference the allegations contained in the  
28 preceding paragraphs of this Complaint.



1 models is an infringing derivative work, made without Plaintiffs’ permission and in violation of  
2 their exclusive rights under the Copyright Act.

3 63. Meta has the right and ability to control the output of the LLaMA language  
4 models. Meta has benefited financially from the infringing output of the LLaMA language  
5 models. Therefore, every output from the LLaMA language models constitutes an act of  
6 vicarious copyright infringement.

7 64. Plaintiffs and the Class have been injured by Meta’s acts of vicarious copyright  
8 infringement. Plaintiffs and the Class are entitled to statutory damages, actual damages,  
9 restitution of profits, and other remedies provided by law.

10  
11 **THIRD CAUSE OF ACTION**

12 **DIGITAL MILLENNIUM COPYRIGHT ACT – REMOVAL OF COPYRIGHT**  
13 **MANAGEMENT INFORMATION**  
14 **17 U.S.C. § 1202(B)**

15 65. Plaintiffs incorporate by reference all allegations of the preceding paragraphs as  
16 though fully set forth herein.

17 66. Plaintiffs bring this claim on behalf of themselves and on behalf of the Class  
18 against Defendants.

19 67. Plaintiffs included one or more forms of copyright-management information  
20 (“CMI”) in each of the Infringed Works, including: copyright notice, title and other identifying  
21 information, or the name or other identifying information about the owners of each book, terms  
22 and conditions of use, and identifying numbers or symbols referring to CMI.

23 68. Without the authority of Plaintiffs and the Class, Meta copied the Infringed  
24 Works and used them as training data for the LLaMA language models. By design, the training  
25 process does not preserve any CMI. Therefore, Meta intentionally removed CMI from the  
26 Infringed Works in violation of 17 U.S.C. § 1202(b)(1).

27 69. Without the authority of Plaintiffs and the Class, Defendant created derivative  
28 works based on the Infringed Works. By distributing these works without their CMI, Meta

1 violated 17 U.S.C. § 1202(b)(3).

2 70. By falsely claiming that it has sole copyright in the LLaMA language models—  
3 which it cannot, because the LLaMA language models are infringing derivative works—Meta  
4 violated 17 U.S.C. § 1202(a)(1).

5 71. Meta knew or had reasonable grounds to know that this removal of CMI would  
6 facilitate copyright infringement by concealing the fact that every output from the LLaMA  
7 language models is an infringing derivative work, synthesized entirely from expressive  
8 information found in the training data.

9 72. Plaintiffs and the Class have been injured by Meta’s removal of CMI. Plaintiffs  
10 and the Class are entitled to statutory damages, actual damages, restitution of profits, and other  
11 remedies provided by law.

12 **FOURTH CAUSE OF ACTION**

13 **VIOLATIONS OF THE CALIFORNIA UNFAIR COMPETITION LAW**  
14 **CAL. BUS. & PROF. CODE §§ 17200, *ET SEQ.***

15 73. Plaintiffs and the Class incorporate by reference each preceding and succeeding  
16 paragraph as though fully set forth at length herein.

17 74. Plaintiffs bring this claim on behalf of themselves and on behalf of the Class  
18 against Defendants.

19 75. Defendant has engaged in unlawful business practices, including violating  
20 Plaintiffs’ and the Class’s rights under the DMCA, and using the Infringed Works to train  
21 LLaMA without Plaintiffs’ or the Class’s authorization.

22 76. The unlawful business practices described herein violate California Business and  
23 Professions Code section 17200 et seq. because that conduct is otherwise unlawful by violating  
24 the DMCA.

25 77. The unlawful business practices described herein violate California Business and  
26 Professions Code section 17200 et seq. because they are unfair, immoral, unethical, oppressive,  
27 unscrupulous or injurious to consumers, because, among other reasons, Defendant used  
28

1 Plaintiffs’ protected works to train LLaMA for Defendant’s own gain without Plaintiffs’ and the  
2 Class’s authorization.

3 78. The unlawful business practices described herein violate California Business and  
4 Professions Code section 17200 et seq. as fraudulent because consumers are likely to be  
5 deceived because, among other reasons, Meta caused LLaMA’s output to be emitted without  
6 any credit to Plaintiffs’ or the Class whose Infringed Works comprise LLaMA’s training dataset.

7 79. Plaintiffs and the Class have been injured by Meta’s removal of CMI. Plaintiffs  
8 and the Class are entitled to statutory damages, actual damages, restitution of profits, and other  
9 remedies provided by law.

10 **FIFTH CAUSE OF ACTION**

11 **NEGLIGENCE**

12 80. Plaintiffs incorporate by reference the allegations of all foregoing paragraphs as  
13 if they had been set forth in full herein.

14 81. Plaintiffs bring this claim on behalf of themselves and on behalf of the Class  
15 against Defendants.

16 82. Defendant owed a duty of care toward Plaintiffs and the Class based upon  
17 Defendant’s relationship to them. This duty is based upon Defendant’s obligations, custom and  
18 practice, right to control information in its possession, exercise of control over the information  
19 in its possession, authority to control the information in its possession, and the commission of  
20 affirmative acts that result in said harms and losses. Additionally, this duty is based on the  
21 requirements of California Civil Code section 1714, requiring all “persons,” including  
22 Defendant, to act in a reasonable manner toward others.

23 83. Defendant breached its duties by negligently, carelessly, and recklessly  
24 collecting, maintaining and controlling Plaintiffs’ and Class members’ Infringed Works and  
25 engineering, designing, maintaining and controlling systems—including LLaMA—which are  
26 trained on Plaintiffs’ and Class members’ Infringed Works without their authorization.

27 84. Defendant owed Plaintiffs and Class members a duty of care to maintain the  
28 Infringed Works once collected and ingested for training LLaMA.



1 Rules of Civil Procedure, declaring Plaintiffs as the representative of the Class, and Plaintiffs'  
2 counsel as counsel for the Class;

3 B. An order awarding declaratory relief and temporarily and permanently enjoining  
4 Defendant from continuing the unlawful and unfair business practices alleged in this Complaint  
5 and to ensure that all applicable information set forth in 17 U.S.C. § 1203(b)(1) is included when  
6 appropriate;

7 C. An award of statutory and other damages under 17 U.S.C. § 504 for violations of  
8 the copyrights of Plaintiff and the Class by Defendants.

9 D. An award of statutory damages under 17 U.S.C. § 1203(b)(3) and 17 U.S.C. §  
10 1203(c)(3), or in the alternative, an award of actual damages and any additional profits under 17  
11 U.S.C. § 1203(c)(2);

12 E. A declaration that Defendant is financially responsible for all Class notice and  
13 the administration of Class relief;

14 F. An order awarding any applicable statutory and civil penalties;

15 G. An order requiring Defendant to pay both pre- and post-judgment interest on any  
16 amounts awarded;

17 H. An award of costs, expenses, and attorneys' fees as permitted by law; and

18 I. Such other or further relief as the Court may deem appropriate, just, and  
19 equitable.

20 **DEMAND FOR JURY TRIAL**

21 Plaintiffs hereby demand a jury trial for all claims so triable.

22  
23 DATED: October 5, 2023

Respectfully submitted,

24  
25 /s/ Daniel J. Muller

26 DANIEL J. MULLER, SBN 193396

dmuller@venturahersey.com

27 **VENTURA HERSEY & MULLER, LLP**

1506 Hamilton Avenue  
28

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

San Jose, California 95125  
Telephone: (408) 512-3022  
Facsimile: (408) 512-3023  
dmuller@venturahersey.com

*/s/ Bryan L. Clobes*

Bryan L. Clobes (*pro hac vice anticipated*)  
**CAFFERTY CLOBES MERIWETHER  
& SPRENGEL LLP**

205 N. Monroe Street  
Media, PA 19063  
Tel: 215-864-2800  
bclobes@caffertyclobes.com

Alexander J. Sweatman (*pro hac vice anticipated*)

**CAFFERTY CLOBES MERIWETHER  
& SPRENGEL LLP**

135 South LaSalle Street, Suite 3210  
Chicago, IL 60603  
Tel: 312-782-4880  
asweatman@caffertyclobes.com

*Attorneys for Plaintiffs*