

DOUGLAS A. WINTHROP (Bar No. 183532)
Douglas.Winthrop@arnoldporter.com
JOSEPH FARRIS (Bar No. 263405)
Joseph.Farris@arnoldporter.com
JESSICA L. GILLOTTE (Bar No. 333517)
Jessica.Gillotte@arnoldporter.com
ESTAYVAINE BRAGG (Bar No. 341400)
Estayvaine.Bragg@arnoldporter.com
ARNOLD & PORTER KAYE SCHOLER LLP
Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
Telephone: (415) 471-3100
Facsimile: (415) 471-3400

MARK LEMLEY (Bar No. 155830)
mlemley@lex-lumina.com
LEX LUMINA LLP
700 S. Flower Street, Suite 1000
Los Angeles, CA 90017
Telephone: (213) 600-6063

Attorneys for Defendant ANTHROPIC PBC

ANGEL T. NAKAMURA (Bar No. 205396)
Angel.Nakamura@arnoldporter.com
OSCAR RAMALLO (Bar No. 241487)
Oscar.Ramallo@arnoldporter.com
ALLYSON MYERS (Bar No. 342038)
Ally.Myers@arnoldporter.com
ARNOLD & PORTER KAYE SCHOLER LLP
777 South Figueroa Street, 44th Floor
Los Angeles, CA 90017-5844
Telephone: (213) 243-4000
Facsimile: (213) 243-4199

JOSEPH R. WETZEL (Bar No. 238008)
joe.wetzel@lw.com
ANDREW M. GASS (Bar No. 259694)
andrew.gass@lw.com
LATHAM & WATKINS LLP
505 Montgomery Street, Suite 2000
San Francisco, CA 94111
Telephone: (415) 391-0600
Facsimile: (415) 395-8095

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

ANDREA BARTZ, ANDREA BARTZ, INC.,
CHARLES GRAEBER, KIRK WALLACE
JOHNSON, and MJ + KJ, INC., individually and on
behalf of others similarly situated,

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

Case No. 3:24-CV-05417-WHA

Action Filed: August 19, 2024

**ANTHROPIC PBC'S OPPOSITION TO
PLAINTIFFS' MOTION FOR CLASS
CERTIFICATION**

Date: May 15, 2025

Time: 8:00 a.m.

Place: Courtroom 12, 19th Floor

Judge: Honorable William H. Alsup

TABLE OF CONTENTS**Page**

INTRODUCTION AND SUMMARY OF ARGUMENT	1
FACTS	2
A. The proposed classes.....	2
B. The class members and their works.	3
1. The Plaintiffs.....	3
2. Andrea Bartz and Andrea Bartz, Inc.	3
3. Charles Graeber.....	4
4. Kirk Wallace Johnson and MJ + KJ, Inc.	5
5. Other authors in the putative classes.....	6
C. The [REDACTED] Class datasets.	7
D. The Internet Books Class datasets.	8
[REDACTED]	8
[REDACTED]	8
[REDACTED]	9
E. Identifying books in the datasets.....	9
ARGUMENT	10
I. THE COURT SHOULD NOT CERTIFY PLAINTIFFS' PROPOSED RULE 23(b)(3) DAMAGES CLASSES BECAUSE THEY FAIL TO ESTABLISH COMMONALITY, PREDOMINANCE, OR ADEQUACY.	10
A. Individualized issues of copyright ownership and validity, infringement, and remedies predominate.	10
1. Copyright registrations are insufficient evidence of ownership.	11
2. Book-by-book evidence required to prove ownership and validity predominates. ...	13
3. The proposed classes are not ascertainable because neither the rightsholders nor the relevant books can be readily identified.	17
a. Metadata is not sufficient to ascertain the books in the Internet Books Datasets.....	17
b. [REDACTED] are not sufficient to identify the books in Anthropic's [REDACTED] or [REDACTED]	18
c. Dr. Zhao's hash method is not sufficient to identify specific books in the Internet Books Datasets.	19
4. Infringement requires individualized proof.	19
5. Plaintiffs fail to provide any class-wide model for determining actual damages and disgorgement.	21
6. Statutory damages suffer the same flaws as actual damages.	22
B. The class action vehicle is not superior.	23
C. Plaintiffs are not adequate class representatives because many rightsholders use LLMs and do not object to the use of their works in training.	23



II. THE COURT SHOULD NOT CERTIFY INJUNCTIVE RELIEF CLASSES BECAUSE
PLAINTIFFS SEEK TO MAXIMIZE MONETARY RECOVERY.24

III. THE COURT SHOULD NOT CERTIFY SUBCLASSES OR “ISSUE CLASSES”
THAT MERELY REPLICATE THE SUBCLASSES24

CONCLUSION25

TABLE OF AUTHORITIES

Page(s)Cases

<i>Alaska Stock, LLC v. Houghton Mifflin Harcourt Publ'g Co.</i> , 747 F.3d 673 (9th Cir. 2014).....	22
<i>Am. Express Co. v. Italian Colors Rest.</i> , 570 U.S. 228 (2013).....	10
<i>Antonick v. Elec. Arts, Inc.</i> , 841 F.3d 1062 (9th Cir. 2016).....	20
<i>Aquarian Found., Inc. v. Lowndes</i> , 127 F.4th 814 (9th Cir. 2025)	11
<i>Bateman v. Am. Multi-Cinema, Inc.</i> , 623 F.3d 708 (9th Cir. 2010).....	23
<i>Betts v. Reliable Collection Agency, Ltd.</i> , 659 F.2d 1000 (9th Cir. 1981).....	25
<i>Comcast Corp. v. Behrend</i> , 569 U.S. 27 (2013).....	10
<i>Danjaq LLC v. Sony Corp.</i> , 263 F.3d 942 (9th Cir. 2001).....	15
<i>Daubert v. Merrell Dow Pharms., Inc.</i> , 509 U.S. 579 (1993).....	18, 21
<i>DRK Photo v. McGraw-Hill Glob. Educ. Holdings, LLC</i> , 870 F.3d 978 (9th Cir. 2017).....	13, 14
<i>Fahmy v. Jay-Z</i> , 908 F.3d 383 (9th Cir. 2018).....	17
<i>Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., Inc.</i> , 499 U.S. 340 (1991).....	11
<i>Flo & Eddie, Inc. v. Sirius XM Radio, Inc.</i> , 2015 WL 4776932 (C.D. Cal. May 27, 2015)	12
<i>Football Ass'n Premier League Ltd. v. YouTube, Inc.</i> , 297 F.R.D. 64 (S.D.N.Y. 2013)	11, 17, 23
<i>Hanagami v. Epic Games, Inc.</i> , 85 F.4th 931 (9th Cir. 2023)	20

1	<i>In re Napster</i> ,	
2	2005 WL 1287611 (N.D. Cal. June 1, 2005)	12
3	<i>In re SFPP Right-of-Way Claims</i> ,	
4	2017 WL 2378363 (C.D. Cal. May 23, 2017)	15
5	<i>Itar-Tass Russian News Agency v. Russian Kurier, Inc.</i> ,	
6	153 F.3d 82 (2d Cir. 1998).....	17
7	<i>Kim v. Allison</i> ,	
8	87 F.4th 994 (9th Cir. 2023)	23
9	<i>Li v. EFT Holdings, Inc.</i> ,	
10	2015 WL 12681648 (C.D. Cal. Dec. 14, 2015)	23
11	<i>Lytle v. Nutramax Lab'ys, Inc.</i> ,	
12	114 F.4th 1011 (9th Cir. 2024)	21
13	<i>Olean Wholesale Grocery Coop., Inc. v. Bumble Bee Foods LLC</i> ,	
14	31 F.4th 651 (9th Cir. 2022) (en banc)	10
15	<i>Pecover v. Elec. Arts Inc.</i> ,	
16	2010 WL 8742757 (N.D. Cal. Dec. 21, 2010)	24
17	<i>Peer Int'l Corp. v. Pausa Recs., Inc.</i> ,	
18	909 F.2d 1332 (9th Cir. 1990).....	22
19	<i>Schneider v. YouTube, LLC</i> ,	
20	674 F. Supp. 3d 704 (N.D. Cal. 2023)	11, 12, 25
21	<i>Sybersound Records, Inc. v. UAV Corp.</i> ,	
22	517 F.3d 1137 (9th Cir. 2008).....	16
23	<i>UMG Recordings, Inc. v. Uncharted Labs, Inc.</i> ,	
24	2025 WL 1047517 (S.D.N.Y. Apr. 8, 2025).....	19, 20
25	<i>Utopia Ent., Inc. v. Claiborne Par.</i> ,	
26	2006 WL 8435006 (W.D. La. Jan. 10, 2006).....	12
27	<i>Valentino v. Carter-Wallace, Inc.</i> ,	
28	97 F.3d 1227 (9th Cir. 1996).....	25
	<i>Van v. LLR, Inc.</i> ,	
	61 F.4th 1053 (9th Cir. 2023)	10
	<i>Vulcan Golf, LLC v. Google</i> ,	
	254 F.R.D. 521 (N.D. Ill. 2008).....	13
	<i>Waite v. UMG Recordings, Inc.</i> ,	
	450 F. Supp. 3d 430 (S.D.N.Y. 2020).....	3

1	<i>Wal-Mart Stores, Inc. v. Dukes,</i>	
2	564 U.S. 338 (2011).....	10
3	<i>WB Music Corp. v. Rykodisc, Inc.,</i>	
4	1995 WL 631690 (E.D. Pa. Oct. 26, 1995).....	11
5	<i>Xavier v. Philip Morris USA Inc.,</i>	
6	787 F. Supp. 2d 1075 (N.D. Cal. 2011)	15
7	<i>Zinser v. Accufix Rsch. Inst., Inc.,</i>	
8	253 F.3d 1180 (9th Cir. 2001).....	23, 24

Statutes

9	17 U.S.C. § 107(2)-(3)	20
10	17 U.S.C. § 304(a)(1)(C)	15
11	17 U.S.C. § 501(b)	13
12	17 U.S.C. § 504(c)(1)-(2).....	22

Other Authorities

14	Fed. R. Civ. P. 23	12, 25
15	Fed. R. Civ. P. 23(a).....	10
16	Fed. R. Civ. P. 23(b)	10
17	Fed. R. Civ. P. 23(b)(3).....	10, 23
18	Fed. R. Civ. P. 23(c)(4)	25
19	Fed. R. Civ. P. 23(c)(5)	25

INTRODUCTION AND SUMMARY OF ARGUMENT

Plaintiffs ask this Court to certify an unprecedented copyright class action to adjudicate the rights associated with five million or more books, spanning nearly a century of publishing history. The proposed classes would sweep in a medley of “natural persons, estates, literary trusts, and loan-out companies” from all over the world who are the “legal or beneficial” owners of these copyrights. But determining whether even a *single* asserted copyright is valid, and, if so, who or what owns it, is rarely straightforward and is often one of the most highly litigated aspects of any copyright suit.

Plaintiffs alone vividly illustrate the problems presented by their requested certification. This case began as a lawsuit by three individual authors. ECF No. 1. Four months later, apparently in recognition of overlooked ownership gaps, Plaintiffs filed a First Amended Complaint, adding as additional plaintiffs two loan-out companies associated with two of the authors, and it is these loan-out companies, and not the authors, that now claim to own the copyrights in the subject works. ECF No. 70 (“FAC”) ¶¶ 12-16. Then, discovery revealed that they may have overlooked yet another loan-out company with ownership interests in the copyrights asserted here. Discovery also revealed that all Plaintiffs have engaged in assignments or transfers of their copyrights—including some *during* the pendency of this case apparently designed to remedy additional ownership gaps. The prospect of this kind of fact-specific inquiry playing out up to five million times explains why no court, ever, has certified a copyright class even remotely approaching this scale or complexity. This action cannot be litigated in a way that will provide common answers to the foundational issue of whether the millions of purported class members in fact own copyrights nor whether such copyrights are valid and enforceable.

Plaintiffs have also failed to carry their burden to show that individualized issues of copyright infringement and damages do not predominate. They propose no plan for how the Court should determine whether any single book was infringed and what the damages for that particular infringement were. Nor have Plaintiffs demonstrated that the classes are ascertainable: There is no practical or reliable way to identify each of the works at issue, much less to link each of the works back to the correct rightsholders.

And even if all of these issues could be overcome, Plaintiffs are not adequate class

representatives because they have an irreconcilable conflict with the putative classes: Many rightsholders, including academics, researchers, and writers who use Large Language Models (“LLMs”), disagree with the position taken by Plaintiffs and actively use and benefit from LLMs like Claude.

The Court should deny the motion for class certification.

FACTS

A. The proposed classes.

The individual Plaintiffs are book authors who seek to represent classes of the “legal or beneficial owners” of the copyrights in millions of books. Plaintiffs’ Motion for Class Certification (“Mot.”) (ECF No. 121-3) at 1; Expert Declaration of Dr. Ben Y. Zhao (“Zhao Decl.”) (ECF No. 121-4) ¶¶ 20, 24, 33, 39-40. They propose two principal classes: (1) an “Internet Books Class” (referred to by Plaintiffs as a “Pirated Books Class,” though they have not defined what “pirated” means) and (2) a [REDACTED] Mot. at 1. Both classes include “natural persons, estates, literary trusts, and loan-out companies” who are the “legal or beneficial owners” of books that: (a) were registered with the United States Copyright Office within five years of the work’s publication; (b) were registered with the United States Copyright Office before being acquired by Anthropic, or within three months of publication; and (c) are assigned an International Standard Book Number (ISBN) or Amazon Standard Identification Number (“ASIN”). *Id.* For the [REDACTED] rightsholders whose books were [REDACTED] by Anthropic, Plaintiffs further limit the class based on books that were actually “used by Anthropic in LLM training.” *Id.* In contrast, for the Internet Books Class, they define the class based solely on whether the rightsholders’ books were “downloaded by Anthropic” as part of the [REDACTED] dataset or from [REDACTED] [REDACTED] without a further limitation that the books were actually used in LLM training. *Id.*

Plaintiffs also propose two “Alternative” subclasses of the Internet Books Class. *Id.* at 1-2. The only difference between the subclasses and their parent class is that: Alternative Class #1 is limited to rightsholders whose books were [REDACTED]; Alternative Class #2 is limited to rightsholders whose books were [REDACTED] *Id.* In addition to damages and

issues classes, Plaintiffs ask to certify an injunctive relief class despite seeking actual damages, disgorgement, pre- and post-judgment interest, and statutory damages to “maximiz[e] recovery . . . and compensation” for the class. *Id.* at 13, 22-23; ECF No. 70 at 18.

B. The class members and their works.

1. The Plaintiffs.

In the Complaint, filed on August 19, 2024, there were three Plaintiffs: the book authors Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson. On December 4, 2024, Plaintiffs filed the First Amended Complaint to add two “loan-out corporation[s]”: Andrea Bartz, Inc. and MJ + KJ, Inc. (for Kirk Wallace Johnson). FAC ¶¶ 15-16. A “loan-out company” is typically a wholly owned “personal services company formed by individuals looking to minimize their tax burdens.” Michael Lovitz, *Loan-Out Companies: Unintended Consequences for Creators?*, 35-FALL Del. Law 16, 16 (2017). It typically employs the creator and is, therefore, the legal author and initial owner of the creator’s works under the work for hire doctrine. *See Waite v. UMG Recordings, Inc.*, 450 F. Supp. 3d 430, 441-42 (S.D.N.Y. 2020).

Plaintiffs propose that copyright ownership for the classes could be proven by copyright registrations alone, but, as discussed below, the history of this case and discovery have revealed that each Plaintiff has a unique and complex ownership story that cannot be demonstrated in that way. Nor could “affidavit[s]” (Mot. at 16) supply the necessary missing ownership information. Even these Plaintiffs have demonstrated an inability to understand the complexities of their own copyrights and lack records sufficient to fill gaps in their knowledge.

2. Andrea Bartz and Andrea Bartz, Inc.

Plaintiffs Andrea Bartz and her loan-out company, Andrea Bartz, Inc. (“Bartz Inc.”), allege Anthropic infringed the copyrights in four novels published by Penguin Random House: *The Lost Night: A Novel*, *The Herd*, *We Were Never Here*, and *The Spare Room*. FAC ¶¶ 58, 61. Bartz Inc. is, and always has been, the registered owner of the copyrights in all four works. Exs. 1-4.¹ Exs. 5-7.

¹ All exhibits are attached to the Declaration of Douglas A. Winthrop in Support of Defendant Anthropic PBC’s Opposition to Plaintiffs’ Motion for Class Certification (“Winthrop Decl.”).

[REDACTED]

1 [REDACTED]

2 [REDACTED] Ex. 8, Deposition of Andrea Bartz (“Bartz Dep.”)

3 162:22-24; Ex. 9; Ex. 10. All advances and royalties for all four books have been paid to Bartz Inc.

4 *E.g.*, Bartz Dep. 195:1-17; 213:21-214:4.

5 [REDACTED]

6 [REDACTED]

7 [REDACTED] Ex. 11. However, Bartz

8 testified that she ultimately realized she was “incorrect in asking [her publisher] to change [the

9 copyright registrations].” Bartz Dep. 148:20-152:22. [REDACTED]

10 [REDACTED]

11 [REDACTED] Ex. 12. At her deposition, Bartz testified that there was

12 no amendment to her contracts as a result of this confusion; but when shown the amendment, Bartz

13 testified that her memory was “refreshed” that there was in fact an amendment to the contract. Bartz

14 Dep. 152:18-154:15.

15 [REDACTED]

16 [REDACTED]

17 [REDACTED] Ex. 13. She testified that she did this in

18 connection with a film option to Netflix to “clear out any possibility of confusion” about the rights.

19 Bartz Dep. 159:3-162:5. [REDACTED]

20 [REDACTED] Ex. 14. When asked if she believed that

21 she retained any right to her copyrights after these assignments, Bartz testified that she was “not

22 positive” (Bartz Dep. 159:3-162:5) and noted several times that she was “not a lawyer” (*id.* at

23 164:19-166:3).

24 3. Charles Graeber

25 Plaintiff Charles Graeber alleges that Anthropic infringed the copyrights in two non-fiction

26 books published by Hachette Books Group: *The Good Nurse: A True Story of Medicine, Madness,*

27 *and Murder* (a true crime account of the life of a serial killer) and *The Breakthrough:*

28 *Immunotherapy and the Race to Cure Cancer* (about the history of immunotherapy). FAC ¶ 59.

1 Graeber is the registered owner of the copyrights in both books. Exs. 15-16; *see also* Exs. 17-
2 18. [REDACTED] Ex. 19, Deposition of Charles A. Graeber
3 (“Graeber Dep.”) 125:10-12; 136:14-137:15.

4 [REDACTED]
5 [REDACTED]
6 [REDACTED]
7 [REDACTED]
8 [REDACTED]
9 [REDACTED]
10 [REDACTED]
11 [REDACTED]
12 *Id.* at 162:20-164:5.
13 [REDACTED]
14 [REDACTED]
15 [REDACTED]
16 [REDACTED]
17 [REDACTED]
18 [REDACTED]
19 [REDACTED]
20 [REDACTED]
21 [REDACTED]
22 [REDACTED]
23 164:24-165:22.

Graeber Dep.

24 **4. Kirk Wallace Johnson and MJ + KJ, Inc.**

25 Plaintiffs Kirk Wallace Johnson and his loan-out company, MJ + KJ, Inc. (“MJ + KJ”),
26 allege Anthropic infringed three non-fiction books: *To Be A Friend Is Fatal* (a memoir about the Iraq
27 war), *The Feather Thief* (a true crime account of a famous thief), and *The Fishermen and the Dragon*
28 (a historical account of an ecological disaster). FAC ¶¶ 60, 62.

Johnson was the original registered owner of the copyright for *To Be a Friend Is Fatal*, whereas MJ + KJ was the original registered owner of the copyrights for the other two books as Johnson's "employer." Exs. 23-25.

Ex. 26, Deposition of Kirk W. Johnson ("Johnson Dep.") 88:16-90:23.

Exs. 27-29.

Ex. 30.

Exs. 31-32.

Ex. 33, § 3.

Id. Ex. A.

On January 28, 2025, during this litigation,

(*id.* at 133:20-134:13).

5. Other authors in the putative classes.

The putative classes are broad and heterogeneous, consisting of hundreds of thousands to millions of rightsholders whose books span all genres and were published in many countries over approximately the last century. While it is impossible to identify the books in the datasets with any precision, it is certain that each class includes rightsholders of: (i) innumerable genres of books, including fiction, textbooks, encyclopedias, history, biography, language arts, philosophy, poetry, technology, and much more; (ii) books authored by citizens of other countries; and (iii) books published across the full span of the past century. Declaration of Mohit Iyyer ("Iyyer Decl.") ¶¶ 91-

102; Ex. 35 at 3.

The class members' books may have been published by any of the 16,000+ publishers in the world, with a patchwork of publishing agreements (not to mention agreements related to film rights and other uses) allocating the rights in those works.²

Exs. 5-7; 17-18; 30-32. *See also* Ex. 36, International Thriller Writers Survey on Artificial Intelligence ("ITW Survey"), at 10-11 (less than 3% of survey respondents said that their publishing contracts contained any wording specific to artificial intelligence).³

The rightsholders in the classes undoubtedly reflect a wide spectrum of views about AI technology and the use of their books by LLMs. Indeed, many of the rightsholders will not be the original authors at all—but will own the rights via assignment, such as academic institutions, media companies, non-author members of loan-out companies (such as Johnson's spouse), estates, and literary trusts. Moreover, as addressed below, even among actual authors, views vary significantly depending on the uses and benefits these individuals receive from AI technology.

C. The [REDACTED] Class datasets.

Declaration of Tom Turvey ("Turvey Decl.") (ECF No. 123-2)

¶¶ 20-27.

Id. ¶ 23.

Id. ¶ 25.

Id. ¶ 26.

² Shahool Al Bari, *Global Book Publishing – Market Research Report (2014-2029)*, IBIS World, available at <https://www.ibisworld.com/global/industry/global-book-publishing/572/> (July 2024).

³ *See also* The Authors Guild, *Survey Reveals 90 Percent of Writers Believe Authors Should Be Compensated for the Use of Their Books in Training Generative AI* ("Authors Guild Survey"), May 15, 2023, at 3 ("67 percent of writers surveyed said they were not sure whether their publishing contracts . . . include permissions or grant of rights to use their work for any AI-related purposes.").

[REDACTED]

1 [REDACTED]

2 [REDACTED]

3 [REDACTED] Zhao Decl. ¶¶ 39, 40.

4 **D. The Internet Books Class datasets.**

5 **1.** [REDACTED]

6 Before it started using [REDACTED], Anthropic used a books dataset it obtained from the
7 internet known as [REDACTED] in the training corpus of some of its commercial LLMs. The dataset was

8 [REDACTED]

9 [REDACTED]

10 [REDACTED] Declaration of Jared Kaplan (“Kaplan Decl.”) (ECF No. 119-

11 5) ¶¶ 48-53. [REDACTED]

Id. ¶¶

12 49-50.

13 [REDACTED]

14 [REDACTED]

15 [REDACTED]

16 [REDACTED]

17 [REDACTED]

18 [REDACTED] Iyyer Decl. ¶ 20. [REDACTED] is

19 the smallest of the datasets at issue for purposes of class certification; according to Plaintiffs, it is

20 comprised of at most around [REDACTED] books. Zhao Decl. ¶ 20. Other than highly unreliable

21 [REDACTED], there is no metadata identifying the contents of Anthropic’s [REDACTED]

22 [REDACTED]. Iyyer Decl. ¶¶ 22-25. Anthropic’s version of [REDACTED] is also incomplete, due to extraction

23 errors and failed downloads; indeed, more than [REDACTED] records that Dr. Zhao describes as possible

24 “books” in Anthropic’s version of [REDACTED] contain less than one page of text. *Id.* ¶ 87.

25 **2.** [REDACTED]

26 Early in the company’s history, Anthropic acquired two different internet book datasets

27

28 ⁴ Available at [REDACTED].

called [REDACTED]. Declaration of Mycal Tucker (“Tucker Decl.”) ¶ 11. According to public reporting, [REDACTED]. Iyyer Decl. ¶ 44. Anthropic used its [REDACTED] in the training corpus of at least one research and development LLM, which was not released commercially. Tucker Decl. ¶ 17. Anthropic has limited metadata associated with its [REDACTED] that purports to identify the books within it, but that metadata was not created or verified by Anthropic. Rather, like [REDACTED] the metadata was crowdsourced over many years and, therefore, contains an incalculable amount of missing and erroneous information. Iyyer Decl. ¶ 47.

3. [REDACTED]

Anthropic did not use its [REDACTED] in the training corpus for an LLM that completed pretraining (*i.e.*, there was never a completely trained LLM created using [REDACTED] that served any purpose). Tucker Decl. ¶ 20. As with [REDACTED], Anthropic has metadata associated with its [REDACTED] that purports to identify the books within it, but that metadata is of unknown provenance and contains an incalculable amount of erroneous information. Iyyer Decl. ¶ 47. For instance, less than 30% of [REDACTED] metadata entries are associated with an actual book file. *Id.* ¶ 53.

E. Identifying books in the datasets.

It is undisputed that neither Anthropic nor Plaintiffs have a complete list of putative class members or even the titles of each book implicated by their proposed classes. To demonstrate that their classes nonetheless are ascertainable, Plaintiffs rely on the declaration of their expert, Dr. Zhao, who asserts that he can identify “which books are present in Anthropic’s datasets” by: (i) consulting their metadata, (ii) using [REDACTED], and/or (iii) using a hash matching algorithm and human visual comparison. Zhao Decl. ¶¶ 42, 50-67. We address each proposed approach below.

ARGUMENT

Plaintiffs fail to show they are entitled to class treatment—“an exception to the usual rule that litigation is conducted by and on behalf of the individual named parties only.” *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 348 (2011) (citation omitted). They bear the burden of proving the requirements of Rule 23(a) are satisfied as well as the requirements for one kind of class under Rule 23(b). *Olean Wholesale Grocery Coop., Inc. v. Bumble Bee Foods LLC*, 31 F.4th 651, 664-65 (9th Cir. 2022) (en banc). In practice, these “stringent requirements” will “exclude most claims.” *Am. Express Co. v. Italian Colors Rest.*, 570 U.S. 228, 234 (2013).

Plaintiffs’ sprawling and heterogenous proposed classes fail each of the commonality, predominance, and adequacy requirements of Rules 23(a) and Rule 23(b)(3). Their asserted subclasses and issue classes are merely subsets of the Internet Books Class and, thus, have the same defects. And, their proposed injunctive relief class is inappropriate in an action primarily seeking monetary relief.

I. THE COURT SHOULD NOT CERTIFY PLAINTIFFS’ PROPOSED RULE 23(b)(3) DAMAGES CLASSES BECAUSE THEY FAIL TO ESTABLISH COMMONALITY, PREDOMINANCE, OR ADEQUACY.

A. Individualized issues of copyright ownership and validity, infringement, and remedies predominate.

Plaintiffs face overwhelming problems with commonality and predominance. As to commonality, which presents a lower bar, Plaintiffs do not carry their burden of establishing that there are “common *answers*” to the central issues of ownership and validity, infringement, and remedies. *Dukes*, 564 U.S. at 350 (citation omitted). They also are unable to clear the “even more demanding” predominance hurdle of Rule 23(b)(3). *Comcast Corp. v. Behrend*, 569 U.S. 27, 34 (2013). Rather, for due process to be served, their claims would require individualized discovery and potentially millions of mini-trials on copyright ownership and validity, infringement, and remedies. That bars certification. *Van v. LLR, Inc.*, 61 F.4th 1053, 1067 n.11 (9th Cir. 2023) (individualized issues overcome predominance when “the discovery and trial process must assess thousands of claims one claim at a time”).

1. Copyright registrations are insufficient evidence of ownership.

Establishing ownership of a valid copyright is an “essential predicate” for each copyright plaintiffs’ claim. *Schneider v. YouTube, LLC*, 674 F. Supp. 3d 704, 718 (N.D. Cal. 2023) (citing *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 361 (1991)). Plaintiffs’ suggestion that the Court can adjudicate the ownership of copyrights in millions of books based on copyright registrations alone as “common proof” (Mot. at 15-17) is reason enough to deny class certification. There is no precedent for certifying a class on such thin evidence of ownership, and for good reason: Copyright registrations are not “definitive” evidence of ownership of a valid copyright. *Aquarian Found., Inc. v. Lowndes*, 127 F.4th 814, 819 (9th Cir. 2025). Copyright registrations are not typically vetted for accuracy⁵ and, even if initially accurate, only reflect a snapshot of ownership rights at the time of registration. As Plaintiffs’ own experience shows, they will frequently not reflect subsequent assignments. Relying on copyright registrations alone to prove ownership and validity would amount to a violation of due process for both Anthropic and potential plaintiffs who might be erroneously excluded due to inaccurate records.

Courts have repeatedly recognized in the class action context that determining copyright ownership and licensing requires a “highly individualized” inquiry. *Schneider*, 674 F. Supp. 3d at 722. Indeed, another court faced with the prospect of a much smaller copyright class action characterized it as presenting a “Frankenstein monster” of individualized issues. *Football Ass’n Premier League Ltd. v. YouTube, Inc.*, 297 F.R.D. 64, 65 (S.D.N.Y. 2013) (citation omitted). For this reason, “copyright claims are poor candidates for class-action treatment.” *Schneider*, 674 F. Supp. 3d at 717 (citation omitted); *see also WB Music Corp. v. Rykodisc, Inc.*, 1995 WL 631690, at *6 (E.D. Pa. Oct. 26, 1995) (copyright validity is “particular to ... one case and no other”). This is true with classes orders of magnitude smaller than the ones proposed here. *See Football Ass’n*, 297 F.R.D. at 65 (“hundreds” of members) (citation omitted); *Schneider*, 674 F. Supp. 3d at 726 (“one thousand

⁵ Compendium of U.S. Copyright Office Practices (“Compendium”) § 309.2 (Jan. 2021) (the Office “generally will accept the facts stated in the application and other registration materials, unless they are implausible or conflict with information in the registration materials, the Office’s records, or other [generally known] information”).

members”); *Utopia Ent., Inc. v. Claiborne Par.*, 2006 WL 8435006, at *1 (W.D. La. Jan. 10, 2006) (“over 3,000” works).

Plaintiffs rely on only two cases as examples of certified classes “involving allegations of largescale infringement.” Mot. at 10, 15-16. Their analogies fail: Neither case relied solely on copyright registrations to resolve issues of ownership and validity.

In *Flo & Eddie, Inc. v. Sirius XM Radio, Inc.*, the court did not rely on copyright registrations *at all*. 2015 WL 4776932 (C.D. Cal. May 27, 2015). Instead, it certified a class of only 273 copyright owners based on three different pre-existing third-party commercial databases tracking pre-1972 hit songs played on SiriusXM. *Id.* at *8, 11. As the court held, the commercial value of these songs created obvious market incentives for accurate ownership tracking across those databases. *Id.* at *7. *See also* Ex. 37, U.S. Copyright Office, Report on Orphan Works (“2006 Orphan Works Report”) 31-32 (Jan. 2006) (noting problems with identifying copyright owners “appear to be less frequent in areas where rights may be administered collectively” via third-party resources, as “in the music industry”). Here, no accurate commercial database of rightsholders exists like the ones in *Flo & Eddie*.

The class certified in *In re Napster* was limited to 27,000 pre-identified music-publisher principals who worked with a *single* “common licensing and collection agent,” the Harry Fox Agency. 2005 WL 1287611, *5 (N.D. Cal. June 1, 2005). Moreover, the court specifically found that the active involvement of the Harry Fox Agency would “reduce, if not eliminate, any complications that may arise from adjudicating the individual aspects of the class members’ claims.” *Id.* at *9. As *Schneider* noted in distinguishing both *In re Napster* and *Flo & Eddie*, such “streamlined circumstances” are a “far cry” from a situation in which the court would be required to sort out ownership rights of millions of unidentified copyright owners. 674 F. Supp. 3d at 722. In sum, no court has ever held copyright registrations alone are sufficient evidence of ownership and validity to meet Rule 23’s predominance requirement, and the dissimilar circumstances presented by *Flo & Eddie* and *Napster* do not support certification here.

2. **Book-by-book evidence required to prove ownership and validity predominates.**

Each class member will need to provide evidence they are the “legal or beneficial owner” of the exclusive right to the copyright in their book for LLM training purposes. 17 U.S.C. § 501(b). This necessarily implicates myriad individual book-by-book factual issues, which will predominate. *See Vulcan Golf, LLC v. Google*, 254 F.R.D. 521, 528, 537 (N.D. Ill. 2008) (“[T]he possibility of hundreds if not thousands of individual hearings related to [registered trademarks] ownership” renders class certification improper, since “multiple time-consuming inquiries regarding ownership” will be necessary).

To determine “legal ownership” of the copyright for each book will require, at the very least, examination of the circumstances in which the book was authored and the terms of any contractual agreement allocating rights for the book, including publication agreements, option agreements, assignments, and/or termination of rights. Even when a copyright holder transfers legal ownership, the inquiry does not end, as the classes also include “beneficial owner[s],” often “an author who ha[s] parted with legal title to the copyright in exchange for percentage royalties based on sales or license fees.” *DRK Photo v. McGraw-Hill Glob. Educ. Holdings, LLC*, 870 F.3d 978, 988 (9th Cir. 2017) (internal citation omitted). But the question of who receives royalties often does not determine ownership, as “an author who receives royalties for a work created under a work-for-hire agreement, and thus who never had ownership of the work, is not a beneficial owner.” *Id.* Thus, resolving legal and beneficial ownership also requires determining whether the book was created as a work for hire. Even though they presumably were carefully selected and vetted, Plaintiffs themselves perfectly illustrate that deciding who is a legal or beneficial owner is an individualized question that cannot be answered through class-wide proof:

Plaintiffs Andrea Bartz and Bartz Inc. Despite alleging otherwise (FAC ¶ 58), Bartz appears to be neither the legal nor beneficial owner of *any* of her works.

Ex. 11.

[REDACTED] DRK Photo, 870 F.3d at 988.

[REDACTED] Ex. 14. There are, at this time, five purported class members before the Court. It is telling that one of them appears to have no legal or beneficial ownership in her works.

Charles Graeber (and Non-Named Plaintiff Having & Selling LLC). [REDACTED]

Graeber Dep. 125:10-12; 137:13-15.

E.g., Exs. 20-21.

Ex. 22 at BARTZ000004627.

Graeber

Dep. 165:20-22. Therefore, there are unresolved questions about whether Graeber retains legal and beneficial ownership of the copyright in *The Good Nurse*. Notably, 22nd Street is not a plaintiff in this case.

Plaintiffs Kirk Wallace Johnson and MJ + KJ. [REDACTED]

Ex. 33 § 3, Ex. A.

Ex. 34. Therefore, Johnson, like Bartz, appears to have no legal or beneficial ownership in his works.

In sum, just three individual authors, selected by sophisticated law firms, present a morass of

ownership issues raising thorny factual and legal questions. Plaintiffs argue that the need to navigate and resolve such complex issues “could be streamlined by affidavit.” Mot. at 16. Not so.

E.g., Bartz Dep. 162:22-24 (testifying that she optioned all four of her books); Winthrop Decl. ¶ 3; Graeber Dep. 164:24-165:22

Allowing putative class members to simply assert ownership by affidavit would deprive Anthropic of the ability to obtain discovery showing flaws in self-serving ownership claims and test such claims through cross-examination. *Xavier v. Philip Morris USA Inc.*, 787 F. Supp. 2d 1075, 1090 (N.D. Cal. 2011). Moreover, Anthropic has a constitutional right to have a jury and/or Article III judge—not a claims administrator—decide critical factual and legal issues of ownership. *Danjaq LLC v. Sony Corp.*, 263 F.3d 942, 962 (9th Cir. 2001); *see also In re SFPP Right-of-Way Claims*, 2017 WL 2378363, at *15 (C.D. Cal. May 23, 2017).

Resolving these kinds of contractual complications would be just the first step in addressing copyright ownership. The need to assess the individual ownership questions described below also predominates over any common issues.

Public Domain Works. Across the proposed classes there no doubt are countless rightsholders whose books have fallen into the public domain for a wide variety of reasons that would require individualized inquiry to resolve. For reference, Cornell University has compiled a chart illustrating the many ways that the copyright for books can lapse. Ex. 38. As just one example, by now, a reasonable number of the rightsholders in books published before 1964 will have passed away, and their works may have fallen into the public domain if their registrations were not properly renewed. *Id.* It will require litigation of individualized issues to make these determinations. And this is far from a simple process: It requires determining if the author timely renewed the copyright and an investigation into who among the author’s relatives were still living at the time of renewal. *See* 17 U.S.C. § 304(a)(1)(C); Nimmer on Copyright § 9.05.

Orphan Works. Because Plaintiffs’ asserted classes are so expansive, there is no question

[REDACTED]

that a significant number of the books at issue are “orphan works,” *i.e.*, “a copyrighted work [whose rightsholder] cannot be identified.” Ex. 37, 2006 Orphan Works Report 1. As noted in a study by the U.S. Copyright Office, orphan works have been a pervasive problem that has plagued the copyright system for years. *Id.* at 21-22 (citing Carnegie Mellon University Libraries study finding 22% of book publishers could not be located—let alone the individual rightsholders in those books). Another Copyright Office study found that, despite “mass digitization” efforts such as Google Books, the problem of orphan works has persisted. Ex. 39, U.S. Copyright Office, Orphan Works and Mass Digitization (“2015 Orphan Works Report”) 2 (June 2015) (“The Office concludes, as it did previously, that the orphan works problem is widespread and significant.”). Per the findings cited there, “[s]tudies of library collections of printed, published books and similar works estimate that between 17% and 25% of published works and as much as 70% of specialized collections are orphan works.” *Id.* at 38.

Works For Hire (Including Academic Books). The classes include a large number of rightsholders whose works were written as works for hire, such as books written for loan-out companies and academic books. *See* Iyyer Decl. ¶ 93; *see also* Turvey Decl. ¶ 24 [REDACTED]. These books pose significant individualized issues related to copyright ownership. For example, the copyright in a book authored by a professor who is an employee of a university may, under some circumstances, belong to the university. However, this is not necessarily the case because copyright ownership policies for academic books vary from university to university and tend to change over time. *See* Robert C. Denicola, *Copyright and Open Access: Reconsidering University Ownership of Faculty Research*, 85 Neb. L. Rev. 351, 380 (2006).⁶ Therefore, books written in academic or other work for hire contexts all present individualized issues to untangle.

Foreign Rightsholders. There likely are numerous foreign rightsholders in the classes. Iyyer Decl. ¶ 102. For them, determining ownership and the effectiveness of any transfer of rights likely

⁶ In addition, rightsholders lack standing to sue for infringement in scenarios where rights were transferred in a putative exclusive license grant from a non-exclusive rightsholder, such as a joint author. *Sybersound Records, Inc. v. UAV Corp.*, 517 F.3d 1137, 1146 (9th Cir. 2008).

[REDACTED]

will implicate complex and individualized choice of law analyses and/or the application of foreign law. *See, e.g., Itar-Tass Russian News Agency v. Russian Kurier, Inc.*, 153 F.3d 82, 92 (2d Cir. 1998); *Fahmy v. Jay-Z*, 908 F.3d 383, 392 (9th Cir. 2018). *See, e.g., Football Ass’n*, 297 F.R.D. at 67 (“Questions of [copyright] title, assignment ... are better handled in the jurisdictions (often foreign) in which they arise, rather than thousands of miles away.”).

The issues discussed above represent only a sliver of the complications that could arise over the adjudication of rights in millions of books. Even if only 0.5% of the possible 5 million books in [REDACTED] required discovery and trial to determine copyright ownership—which is an unreasonably optimistic assumption given what we have seen with just a handful of carefully selected plaintiffs—that would still require more than 25,000 trials about works in [REDACTED] alone. These ownership issues swamp any other common issues and should preclude certification.

3. The proposed classes are not ascertainable because neither the rightsholders nor the relevant books can be readily identified.

The Court should also deny class certification because, as discussed in detail above, Plaintiffs present no viable method to ascertain who *owns* the copyrights to the books at issue. But to identify the rightsholders, one first needs to identify the books at issue, and Plaintiffs do not present a reliable or practical way to do this either: None of the methods proposed by Plaintiffs’ expert objectively or reliably identify the relevant books, much less their owners.

a. Metadata is not sufficient to ascertain the books in the Internet Books Datasets.

Dr. Zhao first contends he can identify books by “consult[ing] the metadata” associated with [REDACTED], [REDACTED], and [REDACTED] (Zhao Decl. ¶¶ 43-50), but the metadata from these datasets is incomplete and/or unreliable. Iyyer Decl. ¶¶ 22-28, 44-53.

[REDACTED] The contents of Anthropic’s [REDACTED] have file name labels, but there is no other metadata. Tucker Decl. ¶ 13; Iyyer Decl. ¶¶ 22-24.

[REDACTED] Iyyer Decl. ¶ 20. They are riddled with incomplete or incorrect information that cannot reliably identify the books in the dataset.

1 *Id.* ¶¶ 25-28.

2 [REDACTED] Anthropic's [REDACTED] are associated with
3 additional metadata, but because the books in the [REDACTED] are uploaded by users, the dataset
4 suffers all the same quality issues as the [REDACTED] and is unreliable for the same reasons.
5 Iyyer Decl. ¶ 47.

6 Despite these known problems with the metadata, Dr. Zhao [REDACTED]
7 [REDACTED] Ex. 40, Deposition of Ben Zhao
8 ("Zhao Dep.") 32:14-33:19 ([REDACTED]), 36:17-24 ([REDACTED]). Instead, [REDACTED]
9 [REDACTED] Zhao Dep. 33:10-19, 98:21-24. But Anthropic produced the [REDACTED]
10 and [REDACTED] metadata in its possession to comply with its discovery obligations; Anthropic has never
11 used that metadata to identify specific books in those datasets. Tucker Decl. ¶ 16. As such, Dr.
12 Zhao's suggestion that the class could be ascertained by "consulting the metadata" should be
13 rejected. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

14
15 **b. [REDACTED] are not sufficient to identify the books in
Anthropic's [REDACTED] or [REDACTED].**

16 Next, Dr. Zhao posits that the Court could rely on [REDACTED]
17 [REDACTED] even
18 though Dr. Zhao himself later conceded he *did not rely* on them. *Compare* Zhao Decl. ¶ 50 with
19 Zhao Dep. 112:4-113:3. This method is unreliable too because, again, Dr. Zhao did nothing to
20 validate the accuracy of these [REDACTED], which were created by people unknown to him with unknown
21 qualifications. *Id.* 49:19-50:21; 57:9-12. Indeed, [REDACTED]

22 [REDACTED]
23 Iyyer Decl. ¶ 30.
24 [REDACTED]
25 [REDACTED]
26 [REDACTED]

27 *Id.* ¶¶ 33, 58.
28 [REDACTED]

c. Dr. Zhao's hash method is not sufficient to identify specific books in the Internet Books Datasets.

Implicitly acknowledging neither metadata nor [REDACTED] provide an accurate means to identify the works, Dr. Zhao provides an "alternative" hash matching method. [REDACTED]

[REDACTED] Zhao Decl. ¶¶ 52-57; Iyyer Decl. ¶ 68. But this method is neither feasible nor reliable.

Implementing Dr. Zhao's hash method would require first obtaining a digitized copy of *all* of the books that *may or may not* be in the datasets. Iyyer Decl. ¶¶ 64-65. In practice, this would mean acquiring accurate digitized copies of essentially all the books registered for U.S. copyrights to lay the foundation for an admissible infringement analysis. In fact, another district court recently required the production of deposit copies—not just any purportedly authoritative copy—to prove "the contours of what was registered" in the context of another generative AI copyright case, rejecting the use of "audio fingerprinting technology" to identify works in suit. *UMG Recordings, Inc. v. Uncharted Labs, Inc.*, 2025 WL 1047517, at *1 (S.D.N.Y. Apr. 8, 2025). Dr. Zhao does not dispute that [REDACTED]

[REDACTED] Zhao Dep. 71:17-72:5, 74:3-76:24.⁷

4. Infringement requires individualized proof.

Even if it were possible to identify all the books in the Internet Books Datasets and who owns their copyrights, other individual issues predominate, including proving the infringement of specific books, and, if the Court were to find Anthropic's use of copyrighted training data generally

⁷ As detailed in the Iyyer Declaration, the hash method suffers numerous other methodological flaws and subjective criteria that make it unreliable at determining whether even a single book is in or out of any dataset. Iyyer Decl. ¶¶ 67-79. And [REDACTED] Zhao Dep. 84:3-85:15, 86:14-88:17.

is not fair use, whether it constitutes fair use for some books.

Infringement. As just discussed, Plaintiffs cannot meet their burden of showing ascertainability because their proposed methodology for doing so depends on the impossible task of obtaining authoritative digital copies of every book under copyright. For the same reason, Plaintiffs lack any methodology to prove infringement of any particular book. Copyright infringement *requires* an actual comparison of the work subject to copyright and the allegedly infringing material. *Antonick v. Elec. Arts, Inc.*, 841 F.3d 1062, 1066 (9th Cir. 2016) (affirming JMOL for defendant when plaintiff did not introduce evidence of the source code allegedly copied; collecting cases for well-established proposition that “[p]roof of the infringement claim consists of the works alleged to be infringed” (internal quotation omitted)). Sweeping, evidence-free statements that Anthropic “copied books” is not sufficient to prove that, in any given case, Anthropic used a particular book in training that is substantially similar to *the actual work* on which a class member obtained a registered copyright based on the deposit copy. *Accord UMG Recordings*, 2025 WL 1047517, at *1. As just one example of the issues that will arise proving infringement, Anthropic’s expert found *thousands* of records in the datasets that appeared to be extraction errors or other instances where very little text appeared as a possible “book.” Iyyer Decl. ¶¶ 35-40, 87-90. Whether such copying of small amounts of text is *de minimis* or infringing is an intensely factual issue that “does not turn on the mere length of the copied material.” *Hanagami v. Epic Games, Inc.*, 85 F.4th 931, 946 (9th Cir. 2023). A jury will need to assess, book-by-book, whether the amount of text copied is substantial, “both qualitatively and quantitatively,” within the specific “context of each case.” *Id.*

Fair Use. As set out in Anthropic’s motion for summary judgment, there are no material facts in dispute necessary to establish that, as a matter of law, the transformative nature of LLM training, along with the other fair use factors, render Anthropic’s use of Plaintiffs’ books fair use, defeating their claims of infringement. But if this Court disagrees and permits their infringement claims to proceed, Anthropic’s fair use defense may turn on facts specific to each of the individual rightsholders’ books at issue. That is because at least two of the four statutory factors may require individual analyses in this case: “the nature of the copyrighted work” (factor two) and “the amount and substantiality of the portion used” (factor three). 17 U.S.C. §§ 107(2)-(3). Moreover, both of

these factors vary extremely among the books at issue. Iyyer Decl. ¶¶ 87-90 (discussing book fragments), 92 (factual books); Ex. 35 at 3 (categories of books).

5. Plaintiffs fail to provide any class-wide model for determining actual damages and disgorgement.

Class certification should also be denied because Plaintiffs provide no “damages model to demonstrate that damages are susceptible to common proof.” *Lytle v. Nutramax Lab’ys, Inc.*, 114 F.4th 1011, 1024 (9th Cir. 2024). Indeed, they do not even attempt to meet their burden to “chart out a path to obtain all necessary data and demonstrate that the proposed method will be viable as applied to the facts of a given case” and to survive at least a “limited” *Daubert* analysis (*id.* at 1031-32). Rather, they hand-wave the problem by (i) calling disgorgement a “collective remedy” (ignoring how this would be applied book-to-book and model-to-model among the many possible rightsholders) and (ii) arguing that “actual damages” will not “vary book-to-book” (ignoring rafts of evidence to the contrary). *See* Turvey Decl. ¶ 17; Ex. 41, Deposition of Tom Turvey (“Turvey Dep.”) 251:13-252:15; Tucker Decl. ¶¶ 4-10.

Plaintiffs provide no roadmap for how a disgorgement remedy would be crafted where Anthropic used vastly different sets of books to train different types of models. For example, Anthropic’s more recent commercial models were trained [REDACTED]. Kaplan Decl. ¶¶ 49-50. In contrast, Anthropic never trained any commercial model that actually generated revenue using [REDACTED]. And it never used [REDACTED] to train any type of complete model at all.

As their sole proof that damages “do *not* vary book-to-book,” Plaintiffs point to a single HarperCollins license (which they obtained via subpoena but do not attach as evidence) that offered \$5,000 per book. Mot. at 22. But even that license proves [REDACTED]. The license provides that [REDACTED]

Confirming that different kinds of books have different values, HarperCollins' CEO has recently noted his company has not been able to license *any* fiction books at *any* price.⁸

6. Statutory damages suffer the same flaws as actual damages.

The same individualized issues that pervade the actual damages analysis afflict the analysis of statutory damages. Potential statutory damages cover a very broad range of \$200 to \$150,000 per work. 17 U.S.C. §§ 504(c)(1)-(2). In determining the amount of an award, "the court is to be guided by what is just in the particular case" *Peer Int'l Corp. v. Pausa Recs., Inc.*, 909 F.2d 1332, 1336 (9th Cir. 1990) (internal quotation omitted). The amount of the award depends on numerous factors that can only be decided on an individualized basis, including: (i) revenue lost by the copyright holder as a result of the infringement, (ii) the profits earned by the defendant as a result of the infringement, (iii) the need to deter future infringement, (iv) the need to penalize the defendant, (v) the circumstances of infringement, and (vi) whether the infringement was intentional. Ninth Circuit Man. Model Civ. Jury Instrs. § 17.35 (2017).

As a first step, to even reach the threshold issue of entitlement to statutory damages, there would need to be an individualized inquiry into when the copyright was registered with respect to the date of publication and infringement. *Alaska Stock, LLC v. Houghton Mifflin Harcourt Publ'g Co.*, 747 F.3d 673, 678 (9th Cir. 2014) ("Registration prior to infringement or, if the work is published, within three months of publication, is necessary for an owner to obtain statutory damages and attorneys' fees."). Thus, to award statutory damages, the Court must make all of the individualized assessments that go into an actual damages or disgorgement analysis—*e.g.*, plaintiffs' losses (which presumably would vary dramatically based on the license fee the book could command, assuming a license were required here),⁹ and defendant's gain—and then layer that

⁸ See The Open Book Podcast, "Truly Global with Brian Murray," (Jan. 13, 2025), [available at https://openroadintegratedmedia.com/podcast/](https://openroadintegratedmedia.com/podcast/) at 00:22:51 (HarperCollins has "not been unable to reach terms with anybody for fiction. I don't have a partner on the AI platform side who's willing to entertain those talks right now.").

⁹ As established in the Tucker and Turvey Declarations, Tucker Decl. ¶¶ 9-10; Turvey Decl. ¶ 17.

analysis with additional subjective factors, such as the defendant’s state of mind and concepts of “what is just.” *Peer Int’l Corp.*, 909 F.2d at 1336. Rather than engage with those factors, Plaintiffs propose the Court should simply steamroll over them (and due process) by applying “basic arithmetic: multiplying the number of copyrighted books . . . by the statutory penalty.” Mot. at 22. The Court should decline that invitation.

B. The class action vehicle is not superior.

Plaintiffs fail to carry their burden to show a class action “is superior to other available methods for fairly and efficiently adjudicating the controversy.” Fed. R. Civ. P. 23(b)(3). Because individual issues predominate, a class action is not superior to individual actions. *See Zinser v. Accufix Rsch. Inst., Inc.*, 253 F.3d 1180, 1192 (9th Cir. 2001) (superiority element is closely linked to predominance).

There is ample incentive for putative class members to pursue individual claims because “the availability of statutory damages [in copyright cases] is designed to give litigation value to each individual case.” *Football Ass’n*, 297 F.R.D. at 66. Moreover, litigating this case as a class action raises significant due process problems for Anthropic and members of the classes that can be avoided by individual lawsuits. It deprives Anthropic of the ability to obtain discovery about individualized issues and to test individualized flaws in those claims. And it would require countless individuals and entities from across the globe to have their claims lumped together and potentially adjudicated without any knowledge that they were part of the class. *Li v. EFT Holdings, Inc.*, 2015 WL 12681648, at *2 (C.D. Cal. Dec. 14, 2015) (denying certification with Chinese class members because of concern that “notice [could] be made on the absent class members consistent with due process”). The prospect of ruinous statutory damages—\$150,000 times 5 million books—only compounds the significant due process implications of certification here. *Bateman v. Am. Multi-Cinema, Inc.*, 623 F.3d 708, 722-23 (9th Cir. 2010).

C. Plaintiffs are not adequate class representatives because many rightsholders use LLMs and do not object to the use of their works in training.

Plaintiffs have alleged they are harmed by the use of their works in LLM training, but, given the massive scope of the classes they seek to represent, this undoubtedly conflicts with the views of

putative class members (many of whom are not even authors). *Kim v. Allison*, 87 F.4th 994, 1000-02 (9th Cir. 2023) (reversing grant of class certification based on conflict with 7,000 members constituting 5% of the class). Claude assists people, including authors, with writing tasks. Kaplan Decl. ¶ 23. Many members of the classes *benefit* from the use of AI tools and do not wish to see them crippled by licensing. *See* Ex. 43¹⁰ at 2 (20% of fiction writers and 25% of non-fiction writers use AI); Ex. 44 (“Authors Alliance NOI Response”) at 8 (noting that many writers use “Claude . . . as part of their creative processes”).¹¹ Academic authors in particular use AI in their professional work, and many of them use books in the same way Anthropic does: to develop LLMs. Ex. 45 at 2.¹² Similarly, Oxford University Press surveyed academic researchers and found that 76% of researchers use AI tools in their research. Ex. 46 at 4.¹³ Plaintiffs cannot adequately represent a class of people when a significant number of people in that same class benefit from the very conduct they claim is unlawful. *See Pecover v. Elec. Arts Inc.*, 2010 WL 8742757, at *9 (N.D. Cal. Dec. 21, 2010) (noting that one type of disqualifying conflict is “if class members benefit from the same acts alleged by the [Plaintiffs] to be harmful to other members of the class”) (internal quotation omitted).

II. THE COURT SHOULD NOT CERTIFY INJUNCTIVE RELIEF CLASSES BECAUSE PLAINTIFFS SEEK TO MAXIMIZE MONETARY RECOVERY.

The Court should deny Plaintiffs’ underdeveloped request to certify injunctive relief classes, which is appropriate only when the “primary relief sought is declaratory or injunctive.” *Zinser*, 253 F.3d at 1195. Plaintiffs admit that their goal is “maximizing recovery . . . and compensation” for the classes (Mot. at 13), thereby defeating their request.

III. THE COURT SHOULD NOT CERTIFY SUBCLASSES OR “ISSUE CLASSES” THAT MERELY REPLICATE THE SUBCLASSES

Plaintiffs alternatively seek to split the Internet Books Class into subclasses or “issue classes”

¹⁰ Ex. 43, The Society of Authors Policy Team, *SoA survey reveals a third of translators and quarter of illustrators losing work to AI* (Apr. 11, 2024), <https://societyofauthors.org/2024/04/11/soa-survey-reveals-a-third-of-translators-and-quarter-of-illustrators-losing-work-to-ai/>.

¹¹ Ex. 44, Authors Alliance, *RE: Policy Study on Artificial Intelligence*, Docket Number 2023-6 (Oct. 30, 2023).

¹² Ex. 45, Project LEND, Response to Notice of Inquiry (NOI) and request for comments (Oct. 30, 2023).

¹³ Ex. 46, Oxford University Press, *Researchers and AI: Survey Findings*.

based on (1) [REDACTED], and (2) [REDACTED] but neither proposed subclass is certifiable. Splitting the massive Internet Books Class into two massive subclasses does nothing to address the intractable problems with the parent class and, under Rule 23(c)(5) “each subclass must independently meet the requirements of Rule 23.” *Betts v. Reliable Collection Agency, Ltd.*, 659 F.2d 1000, 1005 (9th Cir. 1981).

Plaintiffs’ cursory alternative request that the Court certify “issue classes” also fails because they do not identify any “particular issues” for resolution. Fed. R. Civ. P. 23(c)(4). Thus, Plaintiffs do not carry their burden to explain how any issue class would “advance the litigation as a whole.” *Schneider*, 674 F. Supp. 3d at 727. Their contention that they need not show predominance to certify an issue class (Mot. at 24) is wrong. *See Valentino v. Carter-Wallace, Inc.*, 97 F.3d 1227, 1234 (9th Cir. 1996) (reversing class certification because district court erred by not considering whether issue class met predominance requirement as to those issues). Certifying either proposed “issue class” would leave the Court bogged down in adjudicating millions of individual issues of ownership, validity, infringement, and remedies, for which Plaintiffs submit no realistic management plan, and for which no realistic plan could be devised.

CONCLUSION

Courts have consistently recognized that “copyright claims are poor candidates for class-action treatment.” *Schneider*, 674 F. Supp. 3d at 717 (citation omitted). The copyright claims asserted here are no different. The Court should deny the motion for class certification.

Dated: April 17, 2025

By: /s/ Douglas A. Winthrop

ARNOLD & PORTER KAYE SCHOLER LLP

Attorneys for Defendant
ANTHROPIC PBC



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

CERTIFICATE OF SERVICE

I, Douglas A. Winthrop, am the ECF user whose identification and password are being used to file the foregoing **ANTHROPIC PBC’S OPPOSITION TO PLAINTIFFS’ MOTION FOR CLASS CERTIFICATION**.

Dated: April 17, 2025

/s/ Douglas A. Winthrop