

COOLEY LLP
KATHLEEN R. HARTNETT (314267)
(khartnett@cooley.com)
3 Embarcadero Center
20th Floor
San Francisco, CA 94111-4004
Telephone: (415) 693-2000
Facsimile: (415) 693 2222

ARNOLD & PORTER KAYE SCHOLER LLP
DOUGLAS A. WINTHROP (183532)
(Douglas.Winthrop@arnoldporter.com)
JOSEPH FARRIS (263405)
(Joseph.Farris@arnoldporter.com)
Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
Telephone: (415) 471-3100
Facsimile: (415) 471-3400

Attorneys for Defendant
ANTHROPIC PBC

(Additional Counsel on Next Page)

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

ANDREA BARTZ, CHARLES GRAEBER,
and KIRK WALLACE JOHNSON,

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

MORRISON & FOERSTER LLP
DARALYN J. DURIE (169825)
(Ddurie@mofo.com)
425 Market Street
San Francisco, CA 94105-2482
Telephone: (415) 268-7000
Facsimile: (415) 268-7522

MORRISON & FOERSTER LLP
WHITNEY R. O'BYRNE (325698)
(WOByrne@mofo.com)
707 Wilshire Boulevard, Suite 6000
Los Angeles, CA 90017-3543
Telephone: (213) 892-5200
Facsimile: (213) 892-5454

Attorneys for Defendant
ANTHROPIC PBC

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

Action Filed: August 19, 2024

**DEFENDANT ANTHROPIC PBC'S NOTICE
OF FILING OF RULE 23(F) PETITION**

Judge: Honorable William H. Alsup

1 COOLEY LLP
2 EPHRAIM MCDOWELL (*pro hac vice*)
(emcdowell@cooley.com)
3 ALEXANDER J. KASNER (310637)
(akasner@cooley.com)
4 1299 Pennsylvania Avenue NW
Suite 700
Washington, DC 20004-2400
5 Telephone: (202) 842-7800
Facsimile: (202) 842-7899

6
7 ARNOLD PORTER & KAYE SCHOLER LLP
JESSICA L. GILLOTTE (333517)
(Jessica.Gillotte@arnoldporter.com)
8 ESTAYVAINE BRAGG (341400)
(Estayvaine.Bragg@arnoldporter.com)
9 Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
10 Telephone: (415) 471-3100
Facsimile: (415) 471-3400

11 LEX LUMINA LLP
12 MARK LEMLEY (155830)
(mlemley@lex-lumina.com)
13 700 S. Flower Street, Suite 1000
Los Angeles, CA 90017
14 Telephone: (213) 600-6063

15 Attorneys for Defendant
16 ANTHROPIC PBC
17
18
19
20
21
22
23
24
25
26
27
28

MORRISON & FOERSTER LLP
RAMEY W. FISHER (334228)
(RamseyFisher@mofo.com)
JACKSON LANE (351633)
(jlane@mofo.com)
425 Market Street
San Francisco, CA 94105-2482
Telephone: (415) 268-7000
Facsimile: (415) 268-7522

MORRISON & FOERSTER LLP
FITZ BECKWITH COLLINGS (*pro hac vice*)
(fcollings@mofo.com)
MARY PRENDERGAST (272737)
(MPrendergast@mofo.com)
ADITYA VIJAY KAMDAR (324567)
(AKamdar@mofo.com)
2100 L Street, N.W.
Washington, DC 20037
Telephone: (202) 887-1500
Facsimile: (202) 887-0763

NOTICE

Defendant Anthropic PBC (“Anthropic”) gives notice that—in accordance with its statement in its July 24, 2025 Motion for Stay (ECF No. 272, at 1), that it would be filing a Rule 23(f) petition with the U.S. Court of Appeals for the Ninth Circuit—Anthropic timely filed its Rule 23(f) Petition on July 31, 2025. Attached hereto as Exhibit A is Anthropic’s Rule 23(f) Petition.

Dated: August 1, 2025

COOLEY LLP

By: /s/ Kathleen R. Hartnett
Kathleen R. Hartnett

*Attorneys for Defendant
Anthropic PBC*

CERTIFICATE OF SERVICE

I, Kathleen Hartnett, am the ECF user whose identification and password are being used to file the foregoing Defendant Anthropic PBC's Notice of Filing of Rule 23(f) Petition.

Dated: August 1, 2025

By: /s/ Kathleen R. Hartnett

Kathleen R. Hartnett

EXHIBIT A

No. _____

**IN THE UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT**

**ANDREA BARTZ, CHARLES GRAEBER,
and KIRK WALLACE JOHNSON, Plaintiffs,**
v.

ANTHROPIC PBC, Defendant.

On Petition for Permission to Appeal from the United States District Court
for the Northern District of California,
Case No. 3:24-cv-05417-WHA
The Honorable William H. Alsup

**PETITION FOR PERMISSION TO APPEAL
UNDER RULE 23(F)**

Kathleen R. Hartnett
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
Telephone: (415) 693-2000
(khartnett@cooley.com)

Ephraim A. McDowell
Alexander J. Kasner
Elias S. Kim
COOLEY LLP
1299 Pennsylvania Avenue, N.W.
Washington, DC 20004-2400
Telephone: (202) 842-7800

Daralyn J. Durie
MORRISON & FOERSTER LLP
425 Market Street
San Francisco, CA 94105-2482
Telephone: (415) 268-7000

Douglas A. Winthrop
ARNOLD & PORTER KAYE
SCHOLER LLP
Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
Telephone: (415) 471-3100

Allison Wettstein O'Neill
COOLEY LLP
10265 Science Center Drive
San Diego, CA 92121-1117
Telephone: (858) 550-6000

Attorneys for Petitioner Anthropic PBC

CORPORATE DISCLOSURE STATEMENT

Pursuant to Fed. R. App. P. 26.1, the undersigned certifies that the following entities own 10% or more of the equity of Petitioner Anthropic PBC: (1) Google LLC, which is a wholly owned subsidiary of Alphabet, Inc., and (2) Amazon Web Services, Inc., which is a wholly owned subsidiary of Amazon.com.

/s/ Kathleen R. Hartnett

Kathleen R. Hartnett
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
Telephone: +1 415 693 2000
khartnett@cooley.com

*Attorneys for Petitioner
ANTHROPIC PBC*

TABLE OF CONTENTS

	Page
QUESTIONS PRESENTED.....	1
INTRODUCTION	1
STATEMENT OF THE CASE.....	4
REASONS FOR GRANTING THE PETITION.....	9
I. THE CERTIFICATION ORDER IS MANIFESTLY ERRONEOUS IN MULTIPLE RESPECTS THAT IMPLICATE FUNDAMENTAL ISSUES OF CLASS-ACTION LAW.....	9
A. The Court Manifestly Erred in Certifying a Sweeping Copyright Class Where Individualized Issues Will Predominate.....	10
1. Courts Rarely Certify Copyright Classes, Which Typically Raise Numerous Individualized Issues.....	10
2. The Record Here Shows Why Copyright Class Actions Are Rare and Why Certification Was Manifest Error	11
3. The District Court’s Unprecedented Notice Scheme Underscores the Need for Review	15
B. The Court Manifestly Erred in Certifying a Class Despite Anthropic’s Fair-Use Defense.....	17
1. The Class-Certification Order Ignores Anthropic’s Individualized Fair-Use Defenses.....	18
2. To the Extent the District Court Resolved Fair Use Against Anthropic at Class Certification, It Manifestly Erred.....	20
II. THE CLASS-CERTIFICATION ORDER PLACES INORDINATE PRESSURE ON ANTHROPIC TO SETTLE NOTWITHSTANDING THE MERITS.....	21
CERTIFICATE OF COMPLIANCE.....	26

TABLE OF CONTENTS
(continued)

Page

APPENDIX

District Court Order on Class Certification [ECF No. 244]	Exhibit A
District Court Order on Fair Use [ECF No. 231]	Exhibit B
Declaration of Krishna Rao in Support of Defendant’s Petition	Exhibit C
for Permission to Appeal under Rule 23(f)	

TABLE OF AUTHORITIES

	Page
Cases	
<i>Amgen Inc. v. Conn. Ret. Plans & Trust Funds</i> , 568 U.S. 455 (2013).....	3, 20
<i>Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023).....	18
<i>Apple Inc. v. Psystar Corp.</i> , 673 F. Supp. 2d 926 (N.D. Cal. 2009).....	11
<i>Aquarian Found., Inc. v. Lowndes</i> , 127 F.4th 817 (9th Cir. 2025)	10, 13
<i>Authors Guild, Inc. v. Google Inc.</i> , 721 F.3d 132 (2d Cir. 2013)	21
<i>Chamberlan v. Ford Motor Co.</i> , 402 F.3d 952 (9th Cir. 2005)	4, 9, 21
<i>Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.</i> , 499 U.S. 340 (1991).....	10
<i>HarperCollins Publishers LLC v. Open Road Integrated Media, LLP</i> , 7 F. Supp. 3d 363 (S.D.N.Y. 2014)	14
<i>Kadrey v. Meta Platforms, Inc.</i> , No. 23-CV-03417, 2025 WL 1752484 (N.D. Cal. June 25, 2025)	<i>passim</i>
<i>Keiler v. Harlequin Enters. Ltd.</i> , 751 F.3d 64 (2d Cir. 2014)	14
<i>Kihn v. Bill Graham Archives LLC</i> , No. 20-17397, 2022 WL 18935 (9th Cir. Jan. 3, 2022).....	23
<i>Lindsey v. Normet</i> , 405 U.S. 56 (1972).....	16
<i>Microsoft Corp. v. Baker</i> , 582 U.S. 23 (2017).....	9

TABLE OF AUTHORITIES **(continued)**

	Page
<i>In re Napster, Inc. Copyright Litig.</i> , No. C MDL-00-1369 MHP, 2005 WL 1287611 (N.D. Cal. June 1, 2005)	14, 15
<i>Nat’l ATM Council, Inc. v. Visa Inc.</i> , No. 21-7109, 2023 WL 4743013 (D.C. Cir. July 25, 2023).....	23
<i>O’Connor v. Boeing N. Am., Inc.</i> , 184 F.R.D. 311 (C.D. Cal. 1998).....	17
<i>Penguin Grp. (USA) Inc. v. Steinbeck</i> , 537 F.3d 193 (2d Cir. 2008)	14
<i>In re Rail Freight Fuel Surcharge Antitrust Litig.</i> , 725 F.3d 244 (D.C. Cir. 2013).....	23
<i>Random House, Inc. v. Rosetta Books LLC</i> , 283 F.3d 490 (2d Cir. 2002)	14
<i>Schneider v. YouTube, LLC</i> , 674 F. Supp. 3d 704 (N.D. Cal. 2023).....	2, 11, 15
<i>Skidmore v. Led Zeppelin</i> , 952 F.3d 1051 (9th Cir. 2020) (en banc)	11, 13
<i>Van v. LLR, Inc.</i> , 61 F.4th 1053 (9th Cir. 2023)	10
<i>Wal-Mart Stores, Inc. v. Dukes</i> , 564 U.S. 338 (2011).....	3, 15, 16
Statutes	
17 U.S.C. § 107	5
17 U.S.C. § 410(c)	11
17 U.S.C. § 411(a)	10
17 U.S.C. § 504(c)	21

TABLE OF AUTHORITIES **(continued)**

	Page
28 U.S.C. § 1292(b)	4, 7, 8, 23
 Other Authorities	
Daily Journal, <i>IP Lawyers Clash Over Generative AI at 9th Circuit Conference</i> (July 25, 2025), https://www.dailyjournal.com/article/386694-ip-lawyers-clash-over-generative-ai-at-9th-circuit-conference	1
Federal Rule of Civil Procedure 23	2, 3, 4, 9, 20, 23
Melville B. Nimmer & David Nimmer, <i>Nimmer on Copyright</i> § 12.11[C] (1997)	11, 13

QUESTIONS PRESENTED

1. Did the district court manifestly err regarding a fundamental issue of class-action law by certifying a copyright class action with millions of putative class members whose claims present individualized inquiries such as ownership, registration, validity, copying, and damages?

2. Did the district court manifestly err regarding a fundamental issue of class-action law by using its class-certification ruling to opine on the merits of Anthropic's fair-use defense in an effort to minimize individualized issues?

3. Does the death-knell doctrine justify review where Anthropic faces hundreds of billions of dollars in potential damages liability at trial in four months?

INTRODUCTION

This petition seeks review of the unprecedented and erroneous certification of possibly the largest copyright class action ever—involving up to seven million potential claimants, whose works span a century of publishing history—in a case against the generative AI (GenAI) company Anthropic PBC. Several other pending putative class actions likewise challenge companies' use of allegedly copyrighted books to train large language models (LLMs). These cases raise important fair-use questions that will establish the ground rules for the industry moving forward.¹

¹ The important and timely nature of these issues is reflected by a recent debate at the Ninth Circuit Judicial Conference concerning this and other similar cases. *See Daily Journal, IP Lawyers Clash Over Generative AI at 9th Circuit Conference* (July

Despite the complexities and stakes, the district court here rushed to certify a massive class threatening hundreds of billions of dollars in statutory damages, without any reliable means to identify class members or adjudicate their claims. The certification decision also impermissibly purported to resolve critical merits issues. And it did so by glossing over the individualized issues that would pervade a fair trial. This Court should allow immediate appeal of class certification on multiple, independent bases.

First, the district court manifestly erred regarding a fundamental issue of class-action law by certifying a copyright class of massive size and complexity, despite the need for individualized inquiries into issues such as ownership, registration, validity, copying, and damages. “[C]opyright claims are poor candidates for class-action treatment” because they “turn[] ‘upon facts which are particular to that single claim of infringement’” and thus generally fail Rule 23(b)(3)’s predominance requirement. *Schneider v. YouTube, LLC*, 674 F. Supp. 3d 704, 717 (N.D. Cal. 2023) (citation omitted). The court disregarded that principle and the record evidence in asserting that this case “exemplifies the classic litigation that should be certified as a representative action.” Ex. A, ECF 244 (“CC Order”) at 1. It also downplayed the individualized issues based only on “the district judge’s

25, 2025), <https://www.dailyjournal.com/article/386694-ip-lawyers-clash-over-generative-ai-at-9th-circuit-conference>.

judgment from experience (of fifty years).” *Id.* at 25. That is not the “rigorous analysis” that Rule 23 requires. *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 351 (2011).

Second, the district court manifestly erred regarding a fundamental issue of class-action law by purporting to resolve Anthropic’s fair-use defense against it at class certification to avoid the individualized issues presented. In an earlier order, the court granted Anthropic partial summary judgment based on fair-use regarding “training copies and the print-to-digital converted copies”; expressly left open Anthropic’s fair-use defense for trial based on the so-called “pirated copies”; and entered no judgment against Anthropic. Ex. B, ECF 231 (“SJ Order”) at 31. Yet, the class-certification order asserts that “Anthropic *is liable*” for infringing Plaintiffs’ works by downloading books found in certain datasets “regardless of whether they were used for training” LLMs. CC Order at 24-25 (emphasis added). That is wrong as a matter of copyright law, as another district court recently held. *See Kadrey v. Meta Platforms, Inc.*, No. 23-CV-03417, 2025 WL 1752484, at *6, *12 (N.D. Cal. June 25, 2025) (Chhabria, J.). In any event, the court manifestly erred by “engag[ing] in [a] free-ranging merits inquir[y] at the certification stage,” in violation of bedrock Rule 23 principles. *Amgen Inc. v. Conn. Ret. Plans & Trust Funds*, 568 U.S. 455, 466 (2013).

Finally, the district court’s class-certification grant may “sound[] the death

knell of the litigation for [Anthropic] because the grant [would] put considerable pressure on [it] to settle independent of the merits of the plaintiffs' claims." *Chamberlan v. Ford Motor Co.*, 402 F.3d 952, 958 (9th Cir. 2005). Given the size of the class and risk of statutory damages up to \$150,000 per work, Anthropic faces hundreds of billions of dollars in potential liability—a tremendous amount, particularly for an emerging company. Making matters worse, the court is proceeding at warp speed toward a December 1, 2025 trial and otherwise has pressed for settlement. And because Anthropic's motion for interlocutory appeal of the court's erroneous summary-judgment order (which conflicts with *Kadrey*) rests entirely within the court's discretion, *see* 28 U.S.C. § 1292(b), this settlement pressure may ultimately deny Anthropic an appeal of either the merits or class certification absent permission to appeal under Rule 23(f).

One district court's errors should not be allowed to decide the fate of a transformational GenAI company like Anthropic or so heavily influence the future of the GenAI industry generally. This Court can and should intervene now.

STATEMENT OF THE CASE

1. Anthropic is a leading GenAI company that has developed a family of LLMs called Claude. ECF 119-5 ¶ 6. LLMs are text-based GenAI models trained to develop a functional understanding of how language works and then generate completely new text. *Id.* Claude is a versatile LLM that assists with various tasks,

including writing computer code, drafting professional emails, and analyzing business data. *Id.* ¶¶ 20-22. People worldwide use Claude for everything from curriculum planning to animal conservation efforts to drug development. *Id.* ¶¶ 23-28. Developing LLMs requires a voluminous and diverse set of data that allows LLMs to generalize beyond specific inputs and produce creative and useful outputs. *Id.* ¶¶ 38, 40. Thus, to develop its LLMs, Anthropic collected a massive body of data—trillions of “tokens,” each of which roughly represents one English word—from wide-ranging sources. *Id.* ¶¶ 39, 45-53. Among other things, when Anthropic began LLM development in 2021, it downloaded certain books datasets available only on the Internet, such as “LibGen,” “Books3,” and “PiLiMi.” *Id.* ¶¶ 48-50.

2. Plaintiffs filed this putative class action in August 2024, alleging that Anthropic infringed their copyrights by using their books without authorization as part of Claude’s training corpus. ECF 1 ¶¶ 68-74; ECF 70 ¶¶ 72-78. This case is one of the many recent putative class actions filed by authors or creators against GenAI companies, including OpenAI, Meta, Google, Microsoft, NVIDIA, Databricks, Stability AI, DeviantArt, Midjourney, and Runway AI. Several of these cases were filed long before this one, but this case has proceeded so rapidly that it is the first to reach class certification and is presently set for trial on December 1, 2025—almost a year ahead of *Plaintiffs’* initial proposal. *See* ECF 37.

Anthropic moved for summary judgment, arguing fair use under Section 107

of the Copyright Act. ECF 119-7. Plaintiffs simultaneously moved to certify, *inter alia*, a “Pirated Books Class” of authors whose books Anthropic downloaded from Books3, LibGen, or PiLiMi. ECF 121-3. Although the district judge generally bars parties from settlement negotiations before class certification, ECF 8 at 5-6, he changed that practice here because, as stated in a written order, he intends to “take inactive status before the end of the year.” ECF 210.

3. The district court granted-in-part and denied-in-part Anthropic’s summary-judgment motion. The court found that using Plaintiffs’ works “to train specific LLMs [was] justified as a fair use” because “[t]he technology at issue was among the most transformative many of us will see in our lifetimes.” SJ Order at 30 (emphasis omitted). But the court concluded *sua sponte* that the record, taken most favorably to Plaintiffs, indicated that Anthropic may have used Plaintiffs’ works for another purpose—“creat[ing] a central, general-purpose library.” *Id.* at 19. Plaintiffs had not asserted any such “library” theory, and yet the court analyzed Anthropic’s supposed “retain[ing]” of “copies in its central library” as an independent “use.” *Id.* at 14. The court “denie[d] summary judgment for Anthropic that the pirated library copies must be treated as training copies.” *Id.* at 31. It noted Plaintiffs “did not move for summary judgment,” and ordered “a trial on the pirated copies used to create Anthropic’s central library.” *Id.*

Two days later, a different court in the same district reached a divergent fair-

use conclusion. *See Kadrey*, 2025 WL 1752484, at *1. As here, *Kadrey* concerned whether LLM training on “pirated” datasets—including “Books3” and “LibGen”—constituted fair use. *Id.* at *6. Yet unlike here, Judge Chhabria found that fair-use doctrine protected Meta from liability. *Id.* at *23.

Anthropic has moved the district court to certify its summary-judgment order for immediate appeal under 28 U.S.C. § 1292(b), or, alternatively, for reconsideration. ECF 241. That motion will be heard on August 28, 2025.

4. On July 17, 2025, the district court granted class certification as to “[a]ll beneficial or legal copyright owners of the exclusive right to reproduce copies of any book in the versions of LibGen or PiLiMi downloaded by Anthropic,” which includes the owners of up to seven million books. CC Order at 3, 31.

In finding predominance, the court repeatedly cited its “judgment and experience” as the basis for deeming individualized issues “manageable.” *Id.* at 12, 25. And, although the summary-judgment order had determined that using books data to “train specific LLMs” was fair and left open fair use regarding the “pirated copies” for trial, SJ Order at 30-31 (emphasis omitted), the class-certification decision states “that Anthropic *is liable* for all pirated copies regardless of whether they were used for training.” CC Order at 24-25 (emphasis added). Without elaboration, the court stated that “[t]o the extent fair use were to recur as an issue at trial, there is no reason to believe it would not be susceptible to treatment by common

evidence and common methods.” *Id.* at 26.

Finally, the district court devised an unprecedented notice and claims-made process. It ordered Plaintiffs to produce by September 1, 2025 a list of every copyrighted work they plan to assert at trial, along with registration status, ownership information, and other elements, to be subject to “*Daubert* and defendant’s exacting scrutiny.” *Id.* at 16-17 & n. 3. Presumably after Anthropic’s scrutiny and challenge, the court will approve a final list, after which Plaintiffs’ counsel must provide notice to all authors, publishers, and claimants whose works were downloaded, and to all major regional and university presses. *Id.* at 29-30. The court never explained how this notice and opt-out process could feasibly occur before the December 1 trial.

If trial results in liability against Anthropic, then to recover a class member must make a claim stating under oath that he owns the copyright and has served a copy of the claim form on “anyone with whom [he] has entered an agreement respecting the copyright and to any other publisher, author, or person he suspects has any copyright interest of any type in the work.” *Id.* Thus, claimants must act counter to their self-interest by identifying competing claimants to the same work post-trial, after which a court or a jury will decide competing claims. *Id.*

5. Given its then-forthcoming Rule 23(f) petition and pending Section 1292(b) motion, on July 24, 2025, Anthropic moved for a stay in district court. The district court ordered expedited briefing of the stay motion, which it has not yet ruled

on. As Anthropic informed the district court, absent a district court stay, Anthropic will move expeditiously for a stay from this Court.

REASONS FOR GRANTING THE PETITION

Rule 23(f) gives this Court “unfettered discretion” to allow immediate review of a class-certification order “on the basis of *any* consideration.” *Microsoft Corp. v. Baker*, 582 U.S. 23, 31-33 (2017). This Court has recognized three independent circumstances in which review is “most appropriate,” *Chamberlan*, 402 F.3d at 959, all of which are present here. First, the decision “presents an unsettled and fundamental issue of law relating to class actions . . . that is likely to evade end-of-the-case review,” *id.*, because it certifies a sprawling class in which individual issues of ownership and validity, among others, will predominate. Second, the decision is “manifestly erroneous,” *id.*, for that same reason, and because it inappropriately resolves merits issues at class certification. Third, the decision “sounds the death knell of the litigation” for reasons “independent of the merits of the underlying claims,” *id.* at 958-59, due to the significant settlement pressure it places on Anthropic. Review should be granted.

I. THE CERTIFICATION ORDER IS MANIFESTLY ERRONEOUS IN MULTIPLE RESPECTS THAT IMPLICATE FUNDAMENTAL ISSUES OF CLASS-ACTION LAW

No court has ever certified a copyright class action of this magnitude—and for good reason. Rule 23(b)(3) requires that “questions of law or fact common to class members predominate over any questions affecting only individual members.”

When “the discovery and trial process must assess thousands of claims one claim at a time,” individualized issues will predominate. *Van v. LLR, Inc.*, 61 F.4th 1053, 1067 n.11 (9th Cir. 2023). Here, the class members’ claims will require *millions* of fact-specific inquiries into individualized issues, yet the district court summarily asserted that “this is the classic case for certification.” CC Order at 15. Because the certification order rests on multiple manifest errors implicating fundamental class-action issues, the Court should grant immediate review.

A. The Court Manifestly Erred in Certifying a Sweeping Copyright Class Where Individualized Issues Will Predominate

The class-certification order fails to grapple with the individualized issues making copyright-infringement claims by disparate rightsholders generally unsuitable for class treatment, and it compounded that error with novel and problematic notice and claims procedures. Review is needed to correct these manifest errors and clarify the framework applicable to many similar pending cases.

1. Courts Rarely Certify Copyright Classes, Which Typically Raise Numerous Individualized Issues

A plaintiff claiming copyright infringement must prove: (1) “ownership of a valid copyright”; (2) “copying of constituent elements of the work that are original”; and (3) registration with the Copyright Office. *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 361 (1991); *see* 17 U.S.C. § 411(a). Notably, a registration certificate listing a plaintiff is not “definitive” evidence of ownership or validity, but

creates only a rebuttable presumption. *Aquarian Found., Inc. v. Lowndes*, 127 F.4th 817, 819 (9th Cir. 2025); *see* 17 U.S.C. § 410(c). Indeed, because registration certificates often do not reflect subsequent assignments, establishing ownership may require individualized examination of publication agreements, option agreements, and terminations of rights. *See* Melville B. Nimmer & David Nimmer, *Nimmer on Copyright* § 12.11[C] (1997). A plaintiff also must show direct copying or unlawful appropriation of her work’s protected elements. *Skidmore v. Led Zeppelin*, 952 F.3d 1051, 1064 (9th Cir. 2020) (en banc). Assessing copyright damages is also fact-intensive, even if a plaintiff seeks statutory damages. *See Apple Inc. v. Psystar Corp.*, 673 F. Supp. 2d 926, 928 (N.D. Cal. 2009).

Accordingly, “‘copyright claims are poor candidates for class-action treatment,’ and for good reason.” *Schneider*, 674 F. Supp. 3d at 717 (quoting *Football Ass’n Premier League Ltd. v. YouTube, Inc.*, 297 F.R.D. 64, 65 (S.D.N.Y. 2013)). Because “[e]very copyright claim turns ‘upon facts which are particular to that single claim of infringement, and separate from all other claims,’” individualized issues generally predominate. *Id.* (citation omitted).

2. The Record Here Shows Why Copyright Class Actions Are Rare and Why Certification Was Manifest Error

The district court held that Plaintiffs can proceed under the theory that Anthropic infringed their works by downloading millions of books from LibGen and

PiLiMi. SJ Order at 31. Anthropic vigorously disputes that theory, *see* ECF 241, but even under it, individualized issues would predominate.

a. Each class member would need to make numerous showings under the district court’s framework. First, the class member would need to identify a copyrighted book and present evidence that Anthropic downloaded a file containing more than a de minimis portion of the book’s copyrighted text. She would then need to show that she is a legal or beneficial owner of the copyright, and that the copyright is registered and valid. Anthropic would then be entitled to rebut her *prima facie* case with evidence obtained in discovery. If the class member establishes liability, she would still need to establish the proper amount of damages.

The named Plaintiffs here reflect the individualized and fact-specific nature of the class members’ claims. While Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson were the three initial named plaintiffs, *see* Compl. ¶¶ 12-14, discovery revealed that Bartz and Johnson were neither legal nor beneficial owners of the copyrights they asserted. Plaintiffs thus added Andrea Bartz, Inc. and MJ + KJ, Inc. (the loan-out companies that own the asserted works) as named plaintiffs. FAC ¶¶ 12-16. At class certification, the district court scrutinized each Plaintiff’s claimed ownership interest, concluding that Bartz and Johnson could not be part of the class. CC Order at 6-8, 13-14. The court explained, for instance, that although the original registration certificate for one of Johnson’s works named only Johnson

as “author and claimant,” Johnson had later assigned all ownership interests in that work to other entities. *See id.* at 7-8. Indeed, because a registration certificate captures only a snapshot in time, it will frequently fail to list entities that later become beneficial or legal owners. And the fact that the registration certificates of the works asserted by *the named plaintiffs* did not always list those works’ current owners vividly illustrates that registration certificates are not “definitive” evidence of ownership (or validity) and that establishing ownership will require individualized inquiry, including of subsequent licensing agreements. *Aquarian*, 127 F.4th at 819; *see Nimmer* § 12.11[C].

Similarly individualized inquiries will pervade other aspects of class members’ claims. For example, even if a class member shows that Anthropic downloaded a file labeled with the title of a work that has a valid copyright she owns, she still must show that the file contains the text of her work. *Skidmore*, 952 F.3d at 1064. That will not always be straightforward. For instance, LibGen’s metadata was “crowdsourced over many years” and had “missing and erroneous information,” and PiLiMi’s metadata “is of unknown provenance and contains an incalculable amount of erroneous information.” ECF 147-30 at 9.

b. The district court summarily brushed aside ownership, copying, and other individualized issues based on speculation and improper burden-shifting. For example, the court hypothesized that “from experience (of fifty years) . . . there will

be very few competing claimants for the same work.” CC Order at 25. The court also speculated that “authors and their publishers have ongoing business relationships and they will work out whatever differences (if any) they have over how to divide recovery.” *Id.* at 25-26. But, in reality, authors and publishers often litigate the scope of licensing rights. *See, e.g., Random House, Inc. v. Rosetta Books LLC*, 283 F.3d 490 (2d Cir. 2002); *Penguin Grp. (USA) Inc. v. Steinbeck*, 537 F.3d 193 (2d Cir. 2008); *Keiler v. Harlequin Enters. Ltd.*, 751 F.3d 64 (2d Cir. 2014); *HarperCollins Publishers LLC v. Open Road Integrated Media, LLP*, 7 F. Supp. 3d 363 (S.D.N.Y. 2014).

The court also improperly placed the burden on Anthropic to *disprove* infringement, faulting *Anthropic* for failing to identify enough examples of files containing partial books or empty records. CC Order at 18 (“Anthropic’s expert still failed to provide a single example of a purportedly partial book.”). Although Anthropic did not purport to offer an exhaustive list of such examples and many more may exist, the court erred by imposing the burden on Anthropic at all.

c. To support its conclusory assertion that “[t]his civil action exemplifies the classic litigation that should be certified as a representative action,” *id.* at 1, the district court cited only one, inapposite case: *In re Napster, Inc. Copyright Litigation*, No. C MDL-00-1369 MHP, 2005 WL 1287611 (N.D. Cal. June 1, 2005). There, 27,000 music publishers—all working with the same licensing agency—sued

Napster for operating a platform that impermissibly distributed the publishers' songs. *Id.* at *1. In certifying the class, the court emphasized the agency's up-to-date records on the asserted copyrights, which "reduce[d], if not eliminate[d], any complications that may arise from adjudicating the individual aspects of the class members' claims." *Id.* at *9. *Napster* is a "far cry from" cases like this one, where thousands (or millions) of unrelated plaintiffs assert distinct claims. *Schneider*, 674 F. Supp. 3d at 722.

At bottom, the class-certification decision does not reflect the "rigorous analysis" required to determine whether plaintiffs have "affirmatively demonstrate[d]" predominance. *Wal-Mart*, 564 U.S. at 350-51. Such rigorous analysis precludes reliance on a "Trial by Formula"—*i.e.*, allowing plaintiffs to establish liability with statistical modeling—if doing so prevents the defendant from "litigat[ing] its statutory defenses to individual claims." *Id.* at 367. The court violated that principle here by relying on its generalized "experience" to conclude that Plaintiffs could prove their claims collectively and that Anthropic's defenses would be uniform. CC Order at 12. This Court should grant review to correct that manifest error and provide guidance to the many courts facing similar certification requests.

3. The District Court's Unprecedented Notice Scheme Underscores the Need for Review

The court's novel and unsupported process for identifying and notifying class

members reinforces that individualized inquiries predominate and class members cannot be reliably identified. The court mandated that Plaintiffs compile, as purportedly common proof of liability and the basis for class notice, a list of every copyrighted work they plan to assert at trial, along with registration and ownership information. CC Order at 16-17, 30. Yet, recognizing that this list will not capture all interested parties, the court also required *further* notice to potential class members, from other class members, *after* trial and during the claims process. *Id.* Notably, the Plaintiffs did not propose such an unprecedented mechanism.

The court-ordered list of all asserted works is not common proof of liability, but instead at best a compilation of *prima facie* cases for each class member. Yet the court purported to allow Plaintiffs to submit the file “as evidence at trial” of classwide infringement. *Id.* at 17. This one-size-fits-all approach prevents Anthropic from scrutinizing each claim through discovery and “litigat[ing] its statutory defenses to individual claims,” *Wal-Mart*, 564 U.S. at 367, such as a lack of ownership, invalidity, or de minimis copying. If carried out, this approach would raise serious concerns under the Due Process Clause, which “requires that there be an opportunity to present every available defense.” *Lindsey v. Normet*, 405 U.S. 56, 66 (1972). As explained above, the named plaintiffs’ experience powerfully illustrates why Anthropic needs to test and challenge each class member’s claim. *See* pp. 12-13, *supra*.

The district court’s novel notice process also shows that the class is not ascertainable. A class must be “definite enough so that it is administratively feasible for the court to ascertain whether an individual is a member.” *O’Connor v. Boeing N. Am., Inc.*, 184 F.R.D. 311, 319 (C.D. Cal. 1998) (citing 7A Charles A. Wright et al., Federal Practice and Procedure § 1760 (1986)). Yet the court’s process will require class members making claims—*after* class notice and trial have already occurred—to notify any person they “suspect[] has any copyright interest of any type in the work,” so that the person can also make a claim. CC Order at 30. Not only will Anthropic be unable to defend itself at this point—as trial will be over—but the court expects class members to act against self-interest by rigorously identifying competing claimants. And competing claimants receiving notice *after* trial would have had no adequate opportunity to opt-out beforehand. Nor does the court’s plan address the further problem that an original legal owner of a copyright (*i.e.*, the author listed on a registration certificate) who assigned away his rights—and is thus not part of the class—would likely lack incentive to notify *current* rightholders. By outsourcing the work of finding class members to competing claimants (or non-class members) after trial, the court effectively concedes that ascertaining class members now is infeasible.

B. The Court Manifestly Erred in Certifying a Class Despite Anthropic’s Fair-Use Defense

The class-certification order contains two fundamental errors regarding

Anthropic's fair-use defense. First, the court failed to recognize that under the framework articulated in its summary-judgment order, Anthropic's fair-use defense would turn largely on individualized inquiries. Second, to find predominance and certify the class, the class-certification order minimizes and even appears to categorically reject Anthropic's fair-use defense. The apparent use of a class-certification order to resolve critical merits issues contravenes Supreme Court precedent. Both errors justify immediate review.

1. The Class-Certification Order Ignores Anthropic's Individualized Fair-Use Defenses

Courts determine whether a use of a copyrighted work is “fair” by assessing, among other things, “‘whether and to what extent’ the use at issue has a purpose or character different from the original.” *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 510 (2023). Throughout this case, Anthropic has argued that it copied books data—including by downloading datasets from LibGen and PiLiMi—to develop Claude, which is a transformative new product that is nothing like Plaintiffs’ books. ECF 122 at 7, 10-16. Anthropic moved for summary judgment on that basis. *Id.*

The district court then *sua sponte* bifurcated Anthropic's *LLM training* on copyrighted books—which it deemed fair use, SJ Order at 30—from Anthropic's *downloading* of those books. *Id.* at 14. According to the court, the record viewed in Plaintiffs’ favor shows that Anthropic also downloaded books for a different use:

to create a “central library” where books would be stored. *Id.* at 14.² Thus, the Court reserved for trial whether Anthropic in fact had a separate “central library” use for some or all downloaded works, and, if not, whether Anthropic’s use of those works for LLM development rendered its downloading fair. *See id.* at 31. The court indicated that the fair-use inquiry would require examining how quickly Anthropic used a work for training; whether and for how long Anthropic retained the work; and the purpose of retention. *See id.* at 19. Because those facts vary work to work, allowing trial on the novel “central library” theory raised a host of individualized factual issues.

By contrast, the class-certification order brushes away those issues, stating that “[t]o the extent fair use were to recur as an issue at trial, there is no reason to believe that it would not be susceptible to treatment by common evidence and common methods.” CC Order at 26. The court cited no support for such unidentified “common evidence” and “common methods.” The court thus manifestly erred in certifying a class despite the individualized issues presented by Anthropic’s fair-use

² Plaintiffs never argued that such a separate use existed, and the court misconstrued the record: Anthropic downloaded books only to develop LLMs, not to create a “general-purpose library.” *See* ECF 241 at 16-19. Anthropic has sought reconsideration and immediate appellate review of the summary-judgment order based on this and other errors. As noted, the district court in *Kadrey* granted summary judgment to Meta on the basis that copying for the ultimate purpose of training—including downloading from some of the same “pirate sites”—is fair. 2025 WL 1752484, at *23.

defense, justifying this Court’s review.

2. To the Extent the District Court Resolved Fair Use Against Anthropic at Class Certification, It Manifestly Erred

Although the summary-judgment order recognizes Anthropic’s right to present its fair-use defense at trial, the class-certification order seems to eliminate that defense. It declares that “Anthropic *is liable* for all pirated copies regardless of whether they were used for training and regardless of whether another copy of the book was bought and scanned.” CC Order at 24-25 (emphasis added). To the extent the court purported to resolve fair use at class certification—and indeed, where Plaintiffs had not moved for summary judgment—it manifestly erred. “Rule 23 grants courts no license to engage in free-ranging merits inquiries at the certification stage.” *Amgen*, 568 U.S. at 466. Instead, “[m]erits questions may be considered to the extent—but only to the extent—that they are relevant to determining whether the Rule 23 prerequisites for class certification are satisfied.” *Id.*

Here, the merits of fair use are not “relevant to determining whether the Rule 23 prerequisites for class certification are satisfied.” *Id.* The question before the district court should simply have been whether individualized inquiries will predominate at trial. To the extent the court purported to resolve the merits of fair use through class certification, it manifestly erred.

The court’s merits assessment also highlights why the class-certification issues should not be litigated until fair use has been resolved. In a similarly

momentous copyright class action involving application of fair use to Google Books, the Second Circuit observed that resolving the fair-use issue would “perhaps moot [its] analysis of many class certification issues.” *Authors Guild, Inc. v. Google Inc.*, 721 F.3d 132, 134-35 (2d Cir. 2013). It therefore vacated a class-certification order and held “the issue of class certification in abeyance” until resolution of fair use. *Id.* This Court could follow the Second Circuit’s approach by, at minimum, granting review and vacating the class-certification order, so that appellate review of the district court’s erroneous fair-use ruling can precede any class-certification decision.

II. THE CLASS-CERTIFICATION ORDER PLACES INORDINATE PRESSURE ON ANTHROPIC TO SETTLE NOTWITHSTANDING THE MERITS

This case also presents a quintessential “death-knell” situation justifying immediate review, as the class-certification order places inordinate pressure on Anthropic to “settle rather than incur the costs of defending a class action and run the risk of potentially ruinous liability.” *Chamberlan*, 402 F.3d at 957 (citation omitted). The coercive pressure of the district court’s order also means that the manifest and important errors identified, *see supra* Section I, may well “evade end-of-case review” absent an immediate appeal. *Chamberlan*, 402 F.3d at 959.

As the class-certification order recognizes, the potential damages award here is tremendous. CC Order at 28-29. The court certified a class with approximately seven million works, and per-work statutory damages can range from \$200 (for “innocent infringement”) to \$150,000 (for “willful infringement”). 17 U.S.C.

§ 504(c). And although the jury should determine willfulness, the court repeatedly indicated its view that Anthropic has “actively submitted that piracy is for an AI company the fair price to pay.” CC Order at 24; *see id.* at 5, 15, 18, 25 (Anthropic “stole” works); *id.* at 3, 18, 28, 31 (referencing Anthropic’s “pirating”). There is thus a liability risk of hundreds of billions of dollars.

A damages award in that range would be devastating for Anthropic—or indeed almost any company. Although Anthropic has raised a total of \$15 billion, it operates at a loss—i.e., it is years from generating any profits. Rao Decl. ¶ 2-4 (attached as Ex. C). It currently expects to generate no more than \$5 billion in revenue in 2025, and to operate at a loss of billions of dollars this year. *Id.* ¶ 3. If Anthropic must wait to appeal the district court’s erroneous orders until after trial, the potential total damages may make it impossible to post an appeal bond. *Id.* ¶ 5.

The district court’s acceleration of this litigation and other unusual orders have intensified the settlement pressure. Although Plaintiffs filed suit long after other similar copyright class actions, the court expedited this case beyond the parties’ proposals to be the first to class certification and trial—by a wide margin. *See* ECF 37; ECF 49. The court also deviated from its typical bar on settlement discussions before class certification, encouraging talks “sooner rather than later if the parties prefer that the undersigned judge manage the class approval process, given that the judge will likely take inactive status before the end of the year.” ECF

210. The court also set a compressed schedule for notice that, among other things, gives Anthropic little time to challenge Plaintiffs' proposed class list. And the court has repeatedly made clear its views that Anthropic is "liable" and guilty of "piracy" and "stealing" before trial has even occurred. These actions have exacerbated the pressure for Anthropic to settle regardless of the merits.

Anthropic's fair-use defense is meritorious and had it been before Judge Chhabria, Anthropic would have prevailed on summary judgment. *See Kadrey*, 2025 WL 1752484, at *6. But Anthropic's pending Section 1292(b) motion rests entirely within the discretion of the same district court that has accelerated this case to hold a trial before taking inactive status. Thus, absent a stay and Rule 23(f) review, Anthropic may have no path to appeal the rejection of its fair-use defense.

The death-knell doctrine exists to address situations exactly like this one. And courts of appeals routinely grant Rule 23(f) petitions in similar circumstances. *See, e.g., Nat'l ATM Council, Inc. v. Visa Inc.*, No. 21-7109, 2023 WL 4743013, at *5 (D.C. Cir. July 25, 2023) (involving potential treble damages that "would exceed [defendants'] annual net income"); *Kihn v. Bill Graham Archives LLC*, No. 20-17397, 2022 WL 18935, at *1, *3 (9th Cir. Jan. 3, 2022); *In re Rail Freight Fuel Surcharge Antitrust Litig.*, 725 F.3d 244, 254 (D.C. Cir. 2013) (involving "pressure to settle posed by the threat to the defendants' market capitalization").

Although the stakes for Anthropic alone could not be higher, the stakes of this

case are higher still because the district court’s framework threatens not only the fate of Anthropic—one of the industry’s smaller and newer companies—but that of the GenAI industry more broadly. One court’s defective rulings should not be permitted to alter the course of this transformative technology.

CONCLUSION

For the foregoing reasons, the Court should grant the petition.

Dated: July 31, 2025

Respectfully submitted,

/s/ Kathleen R. Hartnett

Kathleen R. Hartnett
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
Telephone: +1 415 693 2000
khartnett@cooley.com

Ephraim A. McDowell
Alexander J. Kasner
Elias S. Kim
COOLEY LLP
1299 Pennsylvania Avenue, N.W.
Washington, DC 20004-2400
Telephone: +1 202 842 7800

Allison Wettstein O'Neill
COOLEY LLP
10265 Science Center Drive
San Diego, CA 92121-1117
Telephone: +1 858 550 6000

Daralyn J. Durie
MORRISON & FOERSTER LLP
425 Market Street
San Francisco, CA 94105-2482
Telephone: (415) 268-7000

Douglas A. Winthrop
ARNOLD & PORTER KAYE
SCHOLER LLP
Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
Telephone: (415) 471-3100

Attorneys for Petitioner
ANTHROPIC PBC

CERTIFICATE OF COMPLIANCE

This petition complies with the page-volume limitations set out in Ninth Circuit Rule 5-2(b) and 32-3 because it is 5,599 words. This petition complies with the typeface requirements of Rule 32(a)(5) and the type-style requirements of Rule 32(a)(6) because it has been prepared in proportionally spaced typeface using Microsoft Word in 14-point Times New Roman font.

/s/ Kathleen R. Hartnett
Kathleen R. Hartnett

EXHIBIT A

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

ANDREA BARTZ, ANDREA BARTZ
INC., CHARLES GRAEBER, KIRK
WALLACE JOHNSON, and MJ + KJ INC.,

No. C 24-05417 WHA

Plaintiffs,

v.

**ORDER ON CLASS
CERTIFICATION**

ANTHROPIC PBC,

Defendant.

INTRODUCTION

This civil action exemplifies the classic litigation that should be certified as a representative action, for the entire class stands aggrieved by defendant's downloading of their books from pirate libraries on the internet. It will be straightforward to prove the classwide wrong done. In short order, counsel can prove defendant violated the Copyright Act by doing Napster-style downloading of millions of works. Defendant, however, objects that it will be a litigation nightmare to determine the specific works it downloaded and, as to each, who is the rightful victim of the infringement.

STATEMENT

At issue on class certification is whether it will be so impractical to identify the millions of books downloaded by defendant Anthropic PBC from pirate libraries — let alone to match

1 them to the authors, publishers, and the like who own the copyrights in them — that class
2 certification should be denied.

3 **1. THE PIRATED BOOKS.**

4 In early 2021, a co-founder of Anthropic downloaded 196,640 unauthorized copies of
5 copyrighted books from the pirate library known as *Books3*. This library conveniently
6 packaged for mass download pairs of each book’s extracted text and filename, enabling the
7 books to be readily rebuilt into separate files or reviewed (*see* Opp. Expert Iyyer ¶¶ 20–21, 23–
8 24, 27 n.10, 39).¹ The pirate behind Books3 publicly downplayed any “copyright backlash”
9 from these efforts (*see* Amd. Compl. ¶ 40). Anthropic’s co-founder downloaded the pirated
10 books to avoid the trouble of paying for them, hoping they might prove useful for training
11 large language models (LLMs) or for something else.

12 Later that year, Anthropic wanted still more. But centralized pirate libraries were being
13 shut down by the government. So, in June 2021, Anthropic’s co-founder used the infamous
14 BitTorrent protocol to copy books peer-to-peer from decentralized copies of another pirate
15 library — Library Genesis, or *LibGen* (*see* ¶ 45). He set a computer program to pause
16 torrenting if disk space got “exhausted” (Br. Exh. 19 at -0389775). These five million copies
17 again came from ebooks, this time remaining in file types like .epub and .pdf to be “viewed as
18 eBooks.” They generally had covers including title and authors (*see* ¶¶ 45, 51 n.25, 60).

19 As pirate libraries were still being shuttered, third parties online copied LibGen and
20 improved upon it to create a new pirate library, simply called “Z-Library.” The FBI shut down
21 Z-Library, too (Br. Exh. 38). But by then, Z-Library had itself been copied or “mirrored” to
22 create the Pirate Library Mirror, or PiLiMi. When Anthropic’s co-founder saw PiLiMi was
23 ready for torrenting, he messaged Anthropic coworkers: “[J]ust in time!” One answered:
24 “zlibrary my beloved” (Br. Exh. 20 at -0391318).

25 So, in July 2022, Anthropic torrented at least two million copies of pirated books from
26 distributed copies of *PiLiMi*. These copies were similarly of .epub, .pdf, and .txt files of digital
27

28 ¹ All citations to paragraphs in this order are from the declaration of Anthropic’s Expert Mohit Iyyer, unless noted.

books (§ 46). Because LibGen and PiLiMi had a common lineage, they had many books in common, too. Appreciating this overlap, about one month before this download, Anthropic engineers had compared Anthropic’s five million copies torrented from LibGen against PiLiMi’s seven million copies available to torrent, and torrented only the two million copies “that [we]re not in Lib[G]en,” as the co-founder explained (Br. Exh. 20 at -0391318).

Plaintiffs thus allege that up to seven million copies of books were stolen by Anthropic in this Napster-like way from all three pirate libraries. Although neither side has systematically tallied the overlap in the books copied from each pirate library, all agree there remains obvious overlap at least in popular titles obtained from each (*see* Br. 9; Opp. 23 (estimating damages from five million unique works); Reply 6 n.8; §§ 93–102 & n.35–38 (explaining that 65, 70, and 78 of 82 popularly listed titles tested appeared in each set)). Anthropic retained all these copies of pirated books in a central library or general data area “forever” (*see, e.g.*, Br. Exh. 12 at -0144509).

Starting in 2024, instead of pirating more books, Anthropic purchased and scanned books. To do so, Anthropic’s chief executive officer and head of training first discussed the “cost to buy all the books in [L]ib[G]en” and the “probability for any given item we can buy it” (Reply Exh. 43 at -0227005). Then its purchasers populated a database with complete metadata for books the company wished to but had not yet purchased (Reply Exh. 45 at -0365971).

2. THE MEANS TO IDENTIFY BOOKS.

Anthropic itself needed to identify the works it copied. It did so using common metadata and methods including hashing. There were several reasons:

First, early on, Anthropic realized it had pirated multiple copies of some books into its central library and that using multiple copies of one text to train an LLM caused the model to memorize the text and perform less well. *Second*, Anthropic held in its central library everything from textbooks to trade fiction. It wanted to understand how the breadth and depth of categories selected for recopying to train LLMs affected performance, if at all. *Third*, other parties sometimes asked Anthropic to remove “discrete corpora” of books (*see* Br. Exh. 1 at 7–

8; *see also* Br. Expert Zhao ¶ 72). *Fourth*, Anthropic was concerned that its completed AI service would regurgitate copyrighted training content, so Anthropic needed to match its LLMs’ output against known works and bar such output (Reply Exh. 46 at -0533421). *Fifth*, once Anthropic stopped using pirate-sourced copies of books to train LLMs, its models suffered a “performance hit.” It wanted to “recover” from this by finding books at least as good as the “old-books quality,” meaning the pirated ones (*see* MSJ Opp. Exh. 20 at -0391439). This prompted an effort to identify books to purchase (but purchase only once) (*see also* Br. Exh. 12 at -0144507 (“deduping”); Reply Exh. 45 at -0365971). *And*, sometimes its researchers “want[ed] to search for a book” (Reply Exh. 45 at -0365941).

So, Anthropic obtained catalogs of metadata corresponding to the books it torrented from LibGen and PiLiMi, and some metadata for Books3. In the catalogs — as Anthropic’s co-founder pointed out to colleagues when reviewing one — were ISBNs (Br. Exh. 19 at -0389774). An *ISBN* is an Industry Standard Book Number for published books issued by a private organization since 1970. It can be included on copyright registrations. Each ISBN is a string of numbers — having substrings defined to flag each book’s language, region, publisher, and unique title (at the edition level). Anthropic used ISBNs to identify books quickly and to unpack from them “intelligence about editions and languages” (Reply Exh. 45 at -0366000). Similarly, an *ASIN*, or Amazon Standard Identification Number, is issued by the online retailer to identify any item sold there, including ebooks or self-published books. Anthropic’s metadata sometimes used ASIN. A *hash value* is another critical piece of metadata contained in those catalogs. Anthropic used hash values, too.

In hashing, a user starts with a source input: either a raw file, the text extracted from that file, a paragraph of any text, or so on. This source input is passed through a *hash function* or algorithm to produce a unique *hash value*, or short alphanumeric signature, which can be used directly or saved for later in a database alongside information known about that input. This can be repeated for many source inputs. Then, a test input is passed through the same algorithm. If its resulting hash matches one of the source hashes, then the test input matches that source input. The algorithm’s *threshold* for a match can be adjusted to recognize perfect

matches or to tolerate fuzzy ones. The hash function is sensitive to differences in the content, but also sensitive to differences in formatting or file-type to the extent those are passed through without normalization — all are just ones and zeroes to it. Finally, such algorithms can be complemented by other programming to do things in sequence and at scale, such as proceeding file by file, extracting and normalizing text from each file to divide into many sub-inputs to hash separately, and comparing each file’s resulting set of hashes alone and in combination against those of other files, and so on (Br. Expert Zhao ¶¶ 53–67; *e.g.*, Br. Exh. 3 at 4).

* * *

In our litigation, both sides used methods like those above to identify works by named plaintiffs that Anthropic had downloaded from pirate libraries. Their ISBNs were furthermore used to find certificates of copyright registration, and thereby to identify relevant plaintiffs as at least beneficial owners of the exclusive right to copy books at issue — a process plaintiffs propose can be followed for all class members.

3. THE MEANS TO IDENTIFY COPYRIGHT REGISTRATIONS, PRESUMPTIVE RIGHTSHOLDERS, AND TRANSFERS.

Plaintiffs must further show, in brief, that they either own the copyrights in their works outright, or else exchanged the exclusive right to reproduce copies of their book for royalties in return. Our named plaintiffs are three writers and two so-called “loan-out” companies used for marketing books — one of those loan-outs being fully owned by one writer, the other co-owned by another and his spouse (“Writers” and “Loan-Outs”). This order finds that all except Author Bartz and Author Johnson are at least beneficial owners of books stolen from the two websites.² Only works in the pirate-sourced books are discussed here. Note well that

² From Books3, Anthropic downloaded works by our writers (with at least beneficial owners):

- Andrea Bartz’s *The Lost Night* (Andrea Bartz, Inc.);
- Charles Graeber’s *The Good Nurse* and *The Breakthrough* (Charles Graeber); and,
- Kirk Wallace Johnson’s *To Be A Friend Is Fatal* and *The Feather Thief* (MJ + KJ, Inc.).

From LibGen, Anthropic torrented:

- Andrea Bartz’s *The Lost Night* and *The Herd* (Andrea Bartz, Inc.);
- Charles Graeber’s *The Good Nurse* (two) and *The Breakthrough* (Charles Graeber); and,
- Kirk Wallace Johnson’s *The Feather Thief* (MJ + KJ, Inc.).

From PiLiMi, Anthropic torrented:

- Andrea Bartz’s *The Lost Night* and *The Herd* (Andrea Bartz, Inc.);

an author starts with the legal title to the entire copyright (or else his employer), and if in exchange for royalties he transfers the legal title to the copyright or to the exclusive right to make copies he nevertheless remains the beneficial owner of that right.

A. WRITER CHARLES GRAEBER.

Writer Charles Graeber penned two non-fiction books at issue: *The Good Nurse: A True Story of Medicine, Madness, and Murder*, and *The Breakthrough: Immunotherapy and the Race to Cure Cancer*.

Certificates of registration from the United States Copyright Office show the works were registered effective as of 2013 and 2018, the same years first published. (Anthropic’s downloading began in 2021.) Each named Writer Graeber as “Author” and “Claimant” (the latter meaning the holder of legal title in the entire copyright at registration), while various agreements were separately recorded (Opp. Exhs. 15–16 (certs.); Br. Expert Zhao ¶ 66 (downloads); MSJ Opp. Exh. 6 at 4 (download dates)).

Respecting *The Good Nurse*, in 2007, a publisher agreed to pay royalties to Writer Graeber in exchange for the “exclusive” “right to reproduce” the work in English including in “print, audio, and electronic” formats (Opp. Exh. 17 at -0207). The parties agreed as to how to prosecute infringement (*id.* at -0216). New York law governed (*id.* at -0219). Respecting *The Breakthrough*, substantially similar terms were agreed (Opp. Exh. 18).

Also respecting *The Good Nurse*, film production companies purchased successive options dated in 2013, 2017, and 2019. Each was a period during which its holder could, if milestones were met along the way to making a film, exercise the option to then obtain certain exclusive rights in exchange for additional compensation and royalties to the author (*e.g.*, Opp. Exh. 22 at -04661, -04666, -04667). Our parties treat the last as having been exercised (*cf.* Opp. Exh. 19 at 155–56; Opp. Exh. 22 at -04623 ¶ 2). Writer Graeber reserved any rights he held “to publish and distribute in all languages throughout the world printed versions” of the

• Charles Graeber’s *The Breakthrough* (Charles Graeber); and,
 • Kirk Wallace Johnson’s *The Feather Thief* (two copies) (MJ + KJ, Inc.).
 Other books were identified in scanned but not pirated books (Br. Expert Zhao ¶¶ 66, 68).

book “whether hardcover or softcover” or “electronic or digital” (Opp. Exh. 22 at -04653). California law governed. (A term for royalties is purportedly defined in an attachment not in our record, if ever attached; this order need not parse that issue because the agreement did not purport to transfer the relevant rights.)

Again respecting *The Good Nurse*, Author Graeber’s solely owned loan-out, Having & Selling LLC, granted other publishers rights to prepare translations and publish the results (Opp. Exh. 20 (Italian); Opp. Exh. 21 (Czech); Opp. Exh. 19 at 164; Reply Graeber ¶ 3).

Writer Graeber is a named plaintiff; his loan-out is not.

B. WRITER KIRK WALLACE JOHNSON AND LOAN-OUT MJ + KJ INC.

Writer Kirk Wallace Johnson put two non-fiction titles at issue: *To Be A Friend Is Fatal: The Fight to Save the Iraqis America Left Behind*, and *The Feather Thief: Beauty, Obsession, and the Natural History Heist of the Century*. Loan-Out MJ + KJ, Inc. is co-owned by him and his spouse.

Copyright certificates show the works were registered effective as of 2013 and 2018, the same years as first published. On the first, Writer Johnson was named author and claimant. On the second, Loan-Out MJ + KJ was named author and claimant, and stating Writer Johnson had worked for hire. Other agreements were recorded (*see* Opp. Exhs. 23, 25 (certs.); Br. Expert Zhao ¶ 66 (downloads); MSJ Opp. Exh. 6 at 4 (download dates)).

Respecting *To Be A Friend Is Fatal*, in 2012, a publisher agreed to pay royalties to Writer Johnson in exchange for “the exclusive right to publish the Work” worldwide including in “electronic text” (Opp. Exh. 30 at -04875–76, -04895–96). They agreed as to how to prosecute infringement (*id.* at -04895). New York law governed (*id.* at -04898).

Respecting *The Feather Thief*, in 2015, another publisher agreed to pay royalties to Loan-Out MJ + KJ in exchange for the “exclusive right to print, publish, and sell the Work” including in electronic copies (Opp. Exh. 31 at -05520–21; *see id.* at -05524–25, -05529). The agreement bound the parties to “transfer and permit the recordation of such copyright ownership [if needed to] permit [one or the other] to bring [an infringement] action in his own

or its own name” (*id.* at -05526). New York law again governed (*id.* at -05528).

Respecting *The Feather Thief*, Loan-Out MJ + KJ granted a movie studio an option to buy “all right[s] . . . other than the reserved rights,” those reserved rights including for “[p]ublishing” including in “e-read” formats (Opp. Exh. 33 at -088, -090).

In an agreement recorded at the Copyright Office this year, Writer Johnson assigned his copyright interests in *To Be A Friend Is Fatal* to Loan-Out MJ + KJ (*see* Opp. Exh. 34; Br. Exh. 30 at -04708).

Writer Johnson and Loan-Out MJ + KJ are named plaintiffs (the spouse is not).

C. WRITER ANDREA BARTZ AND LOAN-OUT ANDREA BARTZ INC.

Writer Andrea Bartz penned two novels at issue: *The Lost Night: A Novel*, and *The Herd*. Loan-Out Andrea Bartz, Inc. is fully owned by her.

Copyright certificates show the works were registered effective as of 2019 and 2020, the same years published. Each named Writer Bartz as author and Loan-Out Bartz as claimant, with further agreements recorded (Opp. Exhs. 1, 4).

Respecting *The Lost Night*, in 2017, a publisher agreed to pay royalties to Loan-Out Bartz in exchange for “the right to publish . . . any and all editions and/or formats of the Work” — “[e]xclusively” (Opp. Exh. 7 at -05290; *see id.* at -05297 (ebooks)). The parties agreed as to how to prosecute infringement (*id.* at -05294). New York law governed (*id.* at -05303). Respecting *The Herd*, in 2018, Loan-Out Bartz entered a substantially similar agreement with the same publisher (Opp. Exh. 5).

For both books, companies purchased options to buy further rights. If exercised, the portion of one agreement in our record states that Loan-Out Bartz was to receive royalties and reserve “print and audio publication rights” (Opp. Exh. 10 at -023). The options for both were returned unexercised (MSJ Br. Exh. 2, Writer Bartz Dep. Tr. 180–81).

Recently, Writer Bartz agreed to transfer to Loan-Out Bartz any residual rights she held in these works (Opp. Exh. 14). Writer Bartz and Loan-Out Bartz are named plaintiffs.

* * *

In August 2024, Writers brought this putative class action complaining that Anthropic had infringed their federal copyrights by amassing copyrighted books from the internet without paying even for the first copy, and further by not obtaining licenses to reproduce further copies for LLM training (Compl. ¶¶ 1–3, 64, 67; *see* Amd. Compl. ¶¶ 1–3, 68, 71). In December 2024, they amended their complaint to include the Loan-Outs.

4. THIS MOTION AND SUMMARY OF CLASS CERTIFICATION.

Now, plaintiffs move to certify classes of copyright owners in a “Pirated Books Class” and in a “Scanned Books Class.” For the former, they propose two alternatives: The Books3 Pirated Books Class and the LibGen & PiLiMi Pirated Books Class (Br. 1–2).

Named plaintiffs bear the burden of affirmatively demonstrating their compliance with Rule 23. *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 350 (2011). The district court may exercise its discretion to modify a class definition under Rule 23(c). *See Gen. Tel. Co. of Sw. v. Falcon*, 457 U.S. 147, 160 (1982). For the same reason, modifying a subclass — even creating one — is “within the district court’s broad power under [the rules] to adopt procedural innovations to facilitate management of the class action.” *Am. Timber & Trading Co. v. First Nat’l Bank of Oregon*, 690 F.2d 781, 786–87 (9th Cir. 1982). To the extent certification turns on expert methods, those methods are subject to a “limited” *Daubert* review. *Lytle v. Nutramax Lab’ys, Inc.*, 114 F.4th 1011, 1025, 1031 (9th Cir. 2024). And, because it is “in the best position to consider the most fair and efficient procedure for conducting any given litigation,” the district court finally exercises discretion “in weighing the correct mix of factors” to make its decision. *Bateman v. Am. Multi-Cinema, Inc.*, 623 F.3d 708, 712 (9th Cir. 2010) (quotation omitted) (vacating denial of certification). A decision to certify the class commands “noticeably more deference” than a denial. *Abdullah v. U.S. Sec. Assocs., Inc.*, 731 F.3d 952, 956 (9th Cir. 2013) (quotation omitted).

This order certifies only a LibGen & PiLiMi Pirated Books Class. This class is limited to actual or beneficial owners of timely registered copyrights in ISBN/ASIN-bearing books downloaded by Anthropic from these two pirate libraries. And, it is further limited to actual or

1 beneficial owners of the copyright interest under Section 106(1), *i.e.*, the right to make
2 reproductions. The class is not limited to authors or author-like entities, however. A key point
3 is to cover everyone who owns the specific copyright interest in play, the right to make copies,
4 either as the actual or as the beneficial owner.

5 A rigorous notice process will be required. Notice will be by first-class mail and email to
6 the author, publisher, and copyright owner listed on the copyright certificate for each work
7 recorded. Notice will also be sent by first-class mail and email to all trade and university
8 publishers in the United States. Notice will also be published at least once in a trade journal.
9 And, as a step in the claims process, notice also will be served by the class claimant himself on
10 all others he attests may claim a copyright interest in the work to notify them that he is
11 claiming it — including all those with whom he has contracted concerning the copyright and/or
12 publishing. If disputes arise over ownership, which will be unlikely, the district court or as
13 needed a jury will resolve them. There will only be one recovery, if any, per single work.

14 This order denies certification for a Books3 Pirated Books Class. The books downloaded
15 as part of Books3 came with fewer fields of metadata, less routinely complete files of content,
16 and as a result the identification of titles and authors would be too problematic. This order also
17 denies certification for a Scanned Books Class.

18 Finally, it does not escape notice that some copyrighted works were copied more than
19 once from across the various sources. Recovery of statutory damages under the Copyright Act
20 is limited to the statutory maximum per work in a given infringement action. As a result, a
21 copyright owner whose work was copied from LibGen and from PiLiMi will be allowed but
22 one recovery.

23 ANALYSIS

24 “Under Rule 23, a class action may be maintained if the four prerequisites of Rule 23(a)
25 are met, and the action meets one of the three kinds of actions listed in Rule 23(b).” *Van v.*
26 *LLR, Inc.*, 61 F.4th 1053, 1062 (9th Cir. 2023).

27 The following class definition will be approved:
28

All beneficial or legal copyright owners of the exclusive right to reproduce copies of any book in the versions of LibGen or PiLiMi downloaded by Anthropic. “Book” refers to any work possessing an ISBN or ASIN which was registered with the United States Copyright Office within five years of the work’s publication and which was registered with the United States Copyright Office before being downloaded by Anthropic, or within three months of publication. Excluded are the directors, officers and employees of Anthropic, personnel of federal agencies, and district court personnel.

1. CRITERIA UNDER RULE 23(a).

Under Rule 23(a), the court must find a proposed class meets (1) numerosity, (2) commonality, (3) typicality, and (4) adequacy as to class representatives and counsel. Here, Anthropic does not contest numerosity or typicality, and treats commonality later within predominance. But Anthropic does not concede adequacy of class representatives (Opp. 23).

A. ADEQUACY OF CLASS REPRESENTATIVES.

(i) Conflict with Owners of Other Copyrights?

Anthropic first contends author- and author-affiliated entities do not adequately represent owners of other copyrights because “academics, researchers, and writers who use [LLMs] disagree with the position taken by plaintiffs and actively use and benefit from LLMs like Claude” (Opp. 2). This is unpersuasive. Whoever would like Anthropic to take their books for free can achieve that result even if the remaining class prevails: They can opt out and license their works to Anthropic or to all the world for nothing. *E.g., Effects Assocs., Inc. v. Cohen*, 908 F.2d 555, 559 (9th Cir. 1990) (individual); *see also Matamoros v. Starbucks Corp.*, 699 F.3d 129, 138–39 (1st Cir. 2012) (class).

And, because the class is limited to those whose books were timely registered, all are entitled to receive statutory damages. There is no conflict between one cohort with that choice and another confined to actual damages — and if one class member thinks her compensable loss is uniquely compelling, she can opt out. *See In re Literary Works in Elec. Databases Copyright Litig.*, 654 F.3d 242, 254–55 & n.6 (2d Cir. 2011).

(ii) ***Conflict with Co-Owners or Rightful Claimants of the Same Works?***

Anthropic also suggests conflicts may emerge between beneficial owners and legal owners, or among co-owners, with beneficial owners better represented by named plaintiffs. A beneficial owner of the right to make copies is someone like an author who receives royalties from any publisher's revenues or recoveries from the right to make copies. Yes, the legal owner might be the publisher but the author has a definite stake in the royalties, so the author has standing to sue. And, each stands to benefit from the copyright enforcement at the core of our case however they then divide the benefit. Also, given the terms on which certification is granted herein, there is no risk of double jeopardy of defendant because there will be only one recovery per work. *See, e.g., EMI Christian Music Grp., Inc. v. MP3tunes, LLC*, 844 F.3d 79, 94–95 (2d Cir. 2016). Therefore, the class definition will bind both the legal and beneficial owners of the pirated books.

It is theoretically possible that a dispute could arise, as Anthropic posits, as to the true copyright owner(s) entitled to recover. For example, an author might claim the entire award only to have the publisher claim separately that it should have received it (and possibly sue Anthropic all over again). In the judgment and experience of the district judge, this is unlikely to occur except in the rarest of instances. This is because authors and publishers are in business together and will work out the best way to recover.

But if a dispute does occur as a part of our action, the following procedures will prevent it from becoming a problem. *First*, publishers and the book world at large will receive a class notice. That circular will put them on notice of the claims and let them know that all owners of the right to make copies will be bound. *Second*, all class claimants in the claims process will be required to serve notice that they are claiming the award to all others associated with the book. For example, an author would be required to serve notice to the publisher that the author has submitted his claim, so that the publisher will have an opportunity to claim a share from him, or else to submit a competing claim in our action, in which case the district judge or a jury

will decide ownership. *Third*, in the very unlikely event such ruinous conflicts become overwhelming, the district court can decertify the class.

(iii) Are Named Plaintiffs Legal or Beneficial Owners?

Anthropic also argues that Writer Bartz and Writer Johnson are inadequate representatives for any class: Neither is a legal owner of any copyrighted work at issue (they have no exclusive rights). And, neither is a beneficial owner (as during this litigation they transferred any and all their interests to their loan-out corporations and not in exchange for royalties) (Opp. 13–15).

This order agrees neither is a legal owner. It is a closer question whether either is a beneficial one. Our court of appeals has called the author who grants exclusive rights in exchange for royalties the “classic example” — but has expressly left the door ajar to others. *DRK Photo v. McGraw-Hill Glob. Educ. Holdings, LLC*, 870 F.3d 978, 988 (9th Cir. 2017). At bottom, beneficial ownership is a “standing doctrine” to enable suit. *BMI v. Hirsch*, 104 F.3d 1163, 1166 (9th Cir. 1997); *Yount v. Acuff Rose-Opryland*, 103 F.3d 830, 834 (9th Cir. 1996). And, the author who transfers her copyright to her loan-out still stands to benefit from the exclusive right by dividends or enhanced valuation. Indeed, the same would be true even if her corporation were itself receiving only royalties. For reasons like these, other district courts have held such authors to be beneficial owners. *See Drake v. Malouf*, No. 99-CV-0315-G, 1999 WL 1007642, at *4 (N.D. Tex. 1999) (Judge Joe Fish); *cf. Wildlife Internationale, Inc. v. Clements*, 591 F. Supp. 1542, 1546 (S.D. Ohio 1984) (Judge Carl Rubin). *But see Roberts v. Gordy*, 359 F. Supp. 3d 1231, 1249 (S.D. Fla. 2019) (Judge Kathleen Williams) (declining to follow).

But there is no reason to follow those courts, because this problem is readily resolved regardless. If the author controls the loan-out, she can bring the suit in its name, or effect a transfer from one hand to the other hand by “at anytime simply memorializ[ing] that intent in writing.” *Jules Jordan Video, Inc. v. 144942 Canada Inc.*, 617 F.3d 1146, 1156 (9th Cir. 2010). And, for the same reason, defendant can hardly complain of any purported defects as between an author and that author’s loan-out in the first place: Where one of them plainly has

standing, and there is no disagreement as between the two, “it would be unusual and unwarranted to permit a third-party infringer to invoke [the written transfers requirement] to avoid suit for copyright infringement.” *See id.* at 1157 (cleaned up).

Because Writer Bartz and Writer Johnson have made the choice to transfer all rights to their loan-outs, they are no longer adequate class representatives. But this situation nevertheless shows why ambiguities between authors and loan-outs will not metastasize.

Loan-Out Bartz and Loan-Out MJ + KJ are on our record adequate representatives. Loan-Out Bartz is the registered claimant of the books penned by Writer Bartz. In exchange for royalties, Loan-Out Bartz granted an exclusive license to reproduce copies. Yes, production companies bought options. But they returned them unexercised. Much the same is true as to Loan-Out MJ + KJ. Specifically as to *The Feather Thief*, Writer Johnson prepared the work for hire for Loan-Out MJ + KJ, which then granted an exclusive license to reproduce copies in exchange for royalties. Separate film options did not purport to transfer those rights a second time.

As for our other named plaintiff, Author Graeber was himself the registered claimant and granted an exclusive license to reproduce copies of his books at issue in exchange for royalties. Yes, he transferred some rights to a filmmaker, but the agreement expressly did not transfer any rights to publish his book — consistent with the above. Nor are rights to prepare Czech or Italian translations and publish the results at issue here.

Loan-Out Bartz, Loan-Out MJ + KJ, and Writer Graeber are adequate class representatives — specifically for the LibGen & PiLiMi Pirated Books Class.

B. ADEQUACY OF CLASS COUNSEL.

The proposed co-lead law firms of Lieff Cabraser Heimann & Bernstein, LLP and Susman Godfrey LLP are experienced in class actions, including copyright (Br. Attorneys Geman & Nelson ¶¶ 30, 34, 46, 47). They represent related parties in co-pending litigation against other AI companies (*id.* ¶¶ 24–26). They do “not hesitate to go to trial” (*id.* ¶ 53; *see id.* ¶ 30). They will fulfill their duty to their class clients. *See Ellis v. Costco Wholesale Corp.*, 657 F.3d 970, 985 (9th Cir. 2011).

Adequacy is met. And, if the commonality prong is also met — as incorporated into the predominance inquiry below — then all of Rule 23(a) is met.

2. CRITERIA UNDER RULE 23(b)(3).

Under Rule 23(b)(3), the court must “find[] that the questions of law or fact common to class members predominate over any questions affecting only individual members, and that a class action is superior to other available methods for fairly and efficiently adjudicating the controversy.” The rule does not impose a threshold requirement that a class be ascertainable, although similar questions may arise to the extent implicated by other requirements. *Briseno v. ConAgra Foods, Inc.*, 844 F.3d 1121, 1124 nn.3–4 (9th Cir. 2017).

A. PREDOMINANCE.

“First, we identify which questions are central to the plaintiffs’ claim. Second, we determine which of these questions are common to the class and which present individualized issues. Third, we analyze whether the common questions predominate” *DZ Rsrv. v. Meta Platforms, Inc.*, 96 F.4th 1223, 1233 (9th Cir. 2024).

This is the classic case for certification. Anthropic downloaded millions of book files. This can and should be proven on a classwide basis. A list of works stolen in this downloading can be constructed using the methods set forth in the statement above. Instead of millions of separate lawsuits with millions of juries, we will have a single proceeding before a single jury, *Napster* style.

(i) Copying Original Elements?

(a) LibGen?

There is no question Anthropic torrented five million files from LibGen. Instead, Anthropic contends that infringement of plaintiffs’ books cannot be shown with classwide evidence because without individualized investigation no one can be sure what it even copied, whether they were the books they said they were, or someone else’s books, or something else like empty files. Plaintiffs propose relying on the catalog of bibliographic metadata listing the five million books — a catalog that Anthropic separately torrented and adopted to account for

what it copied — with titles to be confirmed as needed by common hashing methods and occasional manual review.

Plaintiffs point specifically to a LibGen-provided book catalog. As Anthropic’s co-founder waited for the LibGen-sourced books to torrent, he separately downloaded and started exploring the LibGen-sourced book catalog purporting to describe those books (*see* Br. Exh. 19). Each row represented a book and each column a different type of metadata respecting it (¶ 83). The co-founder pointed out to his team two types of metadata in particular: “Identifier,” and “MD5.” As he restated them, “Identifier is the ISBN,” and “MD5 corresponds to [each file’s] hash” (Br. Exh. 19 at -0389774). Anthropic’s expert later further explained what the MD5 hash means: “An md5 hash is . . . used in the LibGen database to uniquely identify books[, and] is generated by passing the raw binary contents of a file through MD5, a cryptographic hash function” (¶ 51 n.23 (emphasis omitted)). If the same book file were passed through the same MD5 hash function, the very same hash value would result as that listed in the LibGen book catalog (*see* ¶¶ 51, 83). Other metadata on the same row as each book’s ISBN and MD5 hash were the book’s “Title,” “Author,” “Edition,” “Publisher,” and “Year,” and so on (Br. Exh. 19 at -0389774). This evidence could be offered to a jury to support the contention that Anthropic adopted and relied upon this list of the books it copied, and thereby admitted the list of books it copied.

Even Anthropic’s own expert stated in his declaration that a list of errors found in LibGen over the span of several years has lengthened by about one post for every 400 books — suggesting an error rate of less than one percent (*compare* ¶ 49, *with* ¶ 45). Such a per work list could come in via an expert at trial. Defendant would be free to prove up the exceptions and errors.

It remains, however, plaintiffs’ burden to produce such a per work list in the form of a CD with a table having columns for title, author, publisher, copyright registration number, copyright-registered author (frequently the author), copyright-registered claimant (frequently

the author or author’s loan-out), and ISBN and/or ASIN.³ Class notice will be mailed to those listed (once is enough per recipient). The class will be limited to the books on the list for which there is complete information — meaning actual and beneficial owners of an exclusive copyright to reproduce any book on the list. The list will also be offered as evidence at trial. Plaintiffs will be allowed reasonable discovery and time to prepare the list. They must omit from the list any book for which the pirate site provided only an empty file. They must correct the title on any book copied from the pirate site for which the pirate catalog provided the wrong title. Although this order certifies a class, it does so on the condition that plaintiffs prepare and file a comprehensive, per work list and do so by **NOON ON SEPTEMBER 1, 2025**.

Anthropic makes various objections. None persuade. What counts is not whether plaintiffs will prevail at trial, but whether they will be able to mount common evidence and methods in support of their claims. Anthropic raises challenges to these forms of proof that are themselves best addressed through common evidence and methods.

First, Anthropic submits the sworn declaration of one employee who swears that “Anthropic did not create this metadata” and that “Anthropic has never relied on this metadata to identify specific books in its LibGen Dataset” (Opp. Employee Tucker ¶ 16). The jury will hear that testimony and its cross-examination as common evidence as to every class member. The challenge is itself what could be raised in any one class member’s case, *e.g.*, *Range Rd. Music, Inc. v. E. Coast Foods, Inc.*, 668 F.3d 1148, 1154 (9th Cir. 2012), and here it can be done common to all at once.

Second, Anthropic argues that sometimes its infringement was *de minimis* because, whatever its expectations, it in fact obtained partial books due to errors in torrenting. One might suppose this criticism would be readily substantiated and then at our trial prompt manual work-to-work comparison. Not so — or not on this record. After using computer-aided methods to rapidly identify thousands of purportedly partial books downloaded as part of

³ The table may also include anything else plaintiffs deem important, including the hash values or set of hash values computed for the work according to the plaintiffs’ methods, which must withstand *Daubert* and defendant’s exacting scrutiny.

Books3 (¶¶ 87–88), and after trawling through 13,733 complaints collected over several years by internet users of the LibGen library (¶ 49), Anthropic’s expert still failed to provide a single example of a purportedly partial book in the LibGen-sourced books acquired by Anthropic. Moreover, stealing a page of a copyrighted work is still a violation.

Relatedly, Anthropic submits that sometimes the pirate librarians’ extractions or its own torrenting got botched so it got no book at all (¶¶ 59–60). He points to ten examples in five million files that omitted any book (¶ 89). Then he elsewhere explains that whether a book available for torrenting was in fact torrented exactly as expected can be easily checked by comparing the canonical MD5 hash values available in the LibGen book catalog against the computed MD5 hash values of the files as in fact torrented by Anthropic (*see* ¶ 51 & n.23; *cf.* ¶ 72). This is a solvable problem that will not overwhelm the trial.

Finally, Anthropic argues that some metadata is mismatched, such that instead of pirating one book wholesale, it pirated another. To be clear, for each changeling so far discovered, Anthropic’s expert readily found its spitting image in an “Actual Content Match” (¶¶ 51–52). So the problem is not whether our jury will be subjected to detailed infringement analysis. These books, too, were *in toto* infringed, but apparently switched. Instead, the problem is whether rare changelings can be identified using common methods so that, among other things, the right copyright owner(s) can be compensated. Neither expert sized up how common this problem might be. When asked to estimate if all errors in the LibGen metadata amounted to more than one percent, Anthropic’s expert demurred. This failure to substantiate the rate of otherwise speculative errors — particularly for the party possessing all the information — cuts against defendant and not the putative class at this point. *Lytle*, 114 F.4th at 1030. Regardless, this issue, too, cries out for common methods to solve, and sooner than trial. There are several.

For one thing, there are the corroborative sources of common metadata. Recall that Anthropic engineers chose *not* to torrent from the PiLiMi library those books already torrented from the LibGen library, and yet that the metadata for the books *not* torrented remains in the PiLiMi book catalog. And, recall that Anthropic’s chief executive officer and its head of training talked about purchasing books to *replace* the LibGen-sourced books, and populated a

1 database listing books to purchase. The first created a list of cross-checked metadata, the
2 second a fresh list of metadata that may be coextensive. The three sets together may catch the
3 changelings.

4 For another thing, there is the common hashing method that plaintiffs’ Expert Ben Zhao
5 proposed to spot and solve problems with the metadata — and to stand as its own, common
6 method for identifying and indeed proving books were duplicated in their entirety or at least in
7 generous verbatim chunks. *E.g., Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S.
8 539, 548–49 (1985) (“300 [to] 400” words of a fact-intensive book); *Thomson Reuters Enter.*
9 *Centre GmbH v. Ross Intell. Inc.*, 765 F. Supp. 3d 382, 391, 395–96 (D. Del. 2025) (Judge
10 Stephanos Bibas) (paragraphs of headnotes), *appeal docketed*, No. 25-8018 (3d Cir. Apr. 14,
11 2025). The specific implementation would compute multiple hashes across a class claimant’s
12 work and compare the hashes both individually and in bunches against those from Anthropic’s
13 LibGen book files. Expert Zhao identified named plaintiffs’ own books in LibGen this way.⁴

14 Anthropic’s Expert Mohit Iyyer identified no fundamental flaw in plaintiffs’ Expert
15 Zhao’s specific implementation of the decades’-old class of hashing methods. And, Expert
16 Iyyer identified no instance where the hashing method as implemented in fact identified a book
17 that Anthropic lacked (false positive). Instead, the expert identified several adjustments all
18 agree could be made during implementation to further improve results (for instance, improving
19 the way text is normalized before being passed through the hash function). The expert’s other
20 main point was that it was not clear how high the threshold should be set to identify a match,
21 whether exact or fuzzy. That is a matter of debate the experts can have before the jury. And,
22 where the jury cannot conclude whether there was virtual duplication via the method, they can
23

24 ⁴ Expert Zhao outlined an algorithm starting with a known book’s full text, splitting its sentences,
25 dropping any sentences under ten words in length to avoid matching on common phrases, and
26 hashing those longer sentences — then performing the same on an unknown book. The resulting
27 long-sentence hashes from the two books are compared. If not a perfect match, additional
28 programming compares in essence whether the known-book hashes that do appear among the
unknown-book hashes do so in a bunch or else equally spread out across the length of the
unknown book — helping to parse whether the unknown book might simply be a different book
discussing and quoting the first. If any doubt remains, a human can review and confirm (*see*
Br. Expert Zhao ¶¶ 53–67, Exh. D).

1 do so via direct comparison of the copies, with source copies for this comparison to be
2 obtained by class counsel where required for this purpose. It is this order’s “predictive
3 judgment” that this method will “bear fruit.” *Lytle*, 114 F.4th at 1031.

4 **(b) PiLiMi?**

5 There also is no question Anthropic downloaded via torrenting two million files from
6 PiLiMi. Again, Anthropic contends that infringement of plaintiffs’ books cannot be shown
7 with classwide evidence or common methods. Plaintiffs propose relying on the catalog of
8 bibliographic metadata unique to PiLiMi, which again Anthropic torrented, as well as on the
9 common hashing method and occasional manual review.

10 The above points apply equally well here, especially given the intertwined origins of the
11 LibGen and PiLiMi pirate libraries. When cross-checking the works Anthropic did not want to
12 torrent a second time, it did so against PiLiMi’s bibliographic catalog, not yet possessing
13 PiLiMi’s book files. It then used metadata from that catalog to make further dissections of
14 what it did download. Anthropic’s Expert Iyyer has explained that the PiLiMi book catalog is
15 substantially similar to the LibGen book catalog — except that it still contains the five million
16 metadata entries for the books it decided not to copy. When asked to estimate whether the
17 errors in this metadata amounted to more than one percent, Expert Iyyer declined. Regardless,
18 all the challenges raised above, if raised again here, can be solved in similar fashion.

19 * * *

20 Because the quality of common evidence and common methods for identifying these two
21 intertwined sources is similar — indeed, sometimes the same metadata and always the same
22 methods will be common to both — there is no reason why those whose works were
23 copyrighted in one, the other, or both cannot be efficiently considered in one trial.
24 Furthermore, an analysis of the other issues below would bear the same result, including for
25 damages, where the maximum statutory amount will be per work rather than per copy in this
26 one infringement action.

(c) **Books3?**

Again, there is no question Anthropic downloaded 196,640 records in the Books3 collection. Again, Anthropic makes the same arguments. This time they are better received.

When Anthropic’s co-founder downloaded Books3, he did not download and explore a separate book catalog. The only metadata available for Books3 was the metadata intrinsic to the books’ filenames. At their best, the filenames strung together two or three chunks of metadata, like “Title – Author” (*see* ¶¶ 22–23). There is no evidence Anthropic relied on this metadata in the same way it relied on the other sets of metadata — no doubt because this metadata was and is less comprehensive and noticeably less accurate.

True, some short filenames permitted identification. For example, to verify the accuracy of the first book in the set Anthropic downloaded, with filename “Lindbergh – A. Scott.Berg.epub.txt” (¶ 23), Anthropic’s expert “manually compared the opening paragraphs of text [for that book] with the opening paragraphs of text in the Google Books preview of *Lindbergh* by A. Scott Berg and found that the texts appeared to be from the same book” (¶ 25; *see* ¶ 82). But, as Anthropic uncovered in its business, a title-author combination (absent ISBN) can point to multiple editions of the same book, each a different copyrighted work — such as multiple copyrighted versions of abridged or annotated classics. And, if there is only a title present, or only an author, the problem compounds.

Not all Books3’s filenames contained even that much. Many were private mnemonics (“math”), not ready-to-parse metadata (“Mathematics, Substance and Surmise: Views on the Meaning and Ontology of Mathematics” by Ernest Davis and Philip J. Davis) (¶ 28). And, there were mismatched filenames. Anthropic’s expert did investigate ten such books and “[i]nferred correct title and author” (¶ 28). But plaintiffs’ expert proposed no convincing way to accelerate this process, partly because of the further problems described next.

The book content was also spotty. For 1,284 of 196,640 files, there was no content present — resulting in no book in need of identification and thus solving itself (¶¶ 37–38). But similarly obvious errors (for 442 files) or other issues (for at least 3,312 files) resulted in partial content for another 3,750 files (*see* ¶¶ 40, 87), these files having the sparse metadata

particular to this set.

In all these examples, the point is not merely the portion or number of issues, but the combination of apparently less reliable metadata and less reliable content. Plaintiffs have not made clear how they can solve the chicken-or-egg problem in this instance. Yes, there may yet be ways to speed up identification. Our record suggests Anthropic may have found one (*see* Reply Exh. 47 at -0389524). But plaintiffs have had plenty of time to do so, too — whether through expert methods or discovery of common evidence. They still have not. This delay prejudices the owners of the other copyrighted works in LibGen and PiLiMi, who already have comprehensive metadata, more complete book content, and a clear path to prosecute.

Because this problem is so fundamental it moots the remaining inquiry, the certification of the Books3 Pirated Books Class is **DENIED**.

(d) Scanned Books?

The summary judgment order on fair use revealed that the path to recovery for scanning purchased print books and storing their digital replacements peters out. Same goes for further copying used for training LLMs. Although there is evidence of other copying for other uses, this evidence has not been presented with the requisite specificity and uniformity to warrant class certification of a scanned books class specifically — and meanwhile the same points about prejudicing all the other claimants while we wait apply here. Its certification is **DENIED**.

This order will focus on the LibGen & PiLiMi Pirated Books Class from here forward.

(ii) Of a Valid, Copyrighted Work?

Next, plaintiffs must prove that what was duplicated was in fact a valid copyrighted work. This matching process can again be shown in the main through common evidence and methods.

The class is limited to books for which an ISBN or ASIN exists. From the metadata described above, as well as through commercial metadata associated with ISBNs or ASINs, copyright registrations associated with these works can be identified. For every certificate of registration presented by named plaintiffs, for example, there is an associated ISBN (*e.g.*, Br. Exh. 30). And, because the classes are limited to works registered within five years of first

1 publication, these certificates and the facts therein are presumed valid. 17 U.S.C. § 410(c).

2 Anthropic speculates that some registrations might be hard to find or invalid (Opp. 11–
3 12). It fails to substantiate either concern.

4 Yes, some works share even title and author — as do multiple editions. But Anthropic’s
5 own records indicate ISBNs and ASINs are typically unique to each book at the edition level,
6 and so reliably can be linked to a copyright. And, not only do the likely choices of pirate
7 librarians limit the scope of hard-to-find or out-of-print works copied in the first place, but the
8 limitation to works by ISBN and ASIN — the one codified in 1970, the other in the 1990s —
9 further limits the likelihood we will encounter copyright-registered works for which their
10 corresponding registration, if any, cannot quickly be found. (Some self-published authors
11 might use an identifier like ASIN but choose not to register at the Copyright Office at all.)

12 And, yes, although the copyright registrations and facts therein are presumptively valid,
13 the Copyright Office does not verify every one. Anthropic is entitled to rebut the validity
14 (Opp. 11 n.5). But any “inadvertent mistakes on registration certificates [would] not invalidate
15 a copyright and thus [would] not bar infringement actions, unless the alleged infringer ha[d]
16 relied to its detriment on the mistake.” *Jules Jordan Video*, 617 F.3d at 1156 (cleaned up).
17 Nothing in our record hints at Anthropic having relied on anything in a copyright certificate.
18 So, if an author’s loan-out were named instead of the author, or vice-versa, such a mistake
19 would not stall our litigation.

20 **(iii) Causing Damages?**

21 Plaintiffs next must prove damages. They propose theories for proving statutory
22 damages and actual damages. This, too, can be shown through common evidence and
23 methods.

24 The LibGen & PiLiMi Pirated Books Class is limited to works that were registered prior
25 to infringement or within three months of first publication. (The date of effective registration
26 and publication are stated on the certificates of registration and, for reasons above, presumed
27 true.) Because class members’ works will meet this scope (by class definition), they will be
28 entitled to elect statutory damages rather than actual damages. 17 U.S.C. §§ 412, 504(c)(1).

1 To the extent a class member believes her actual damages will surpass the statutory ones, she
2 can intervene to prove it or opt out.

3 In order to determine a statutory award, a jury must first make findings as to the
4 infringer's mental state (e.g., willfulness), then consider factors including actual damages and
5 deterrence to set an award within a range of values allowed for that mental state — bottoming
6 out at \$200 per work for the reasonable user who believed his use fair. *See id.* § 504(c)(2);
7 *Desire, LLC v. Manna Textiles, Inc.*, 986 F.3d 1253, 1271–72 (9th Cir. 2021).

8 Anthropic has introduced no evidence from which it might be supposed that its mental
9 state differed as to its torrenting of any of these books. To the contrary, it has actively
10 submitted that piracy is for an AI company the fair price to pay. Thus, common evidence can
11 be taken as to the mental state involved in the copying that is the crux of plaintiffs' contentions
12 for these classes. Questions of deterrence would also be general in nature. And, in the
13 alternative, Anthropic has suggested that *it already made recompense* by paying the full price
14 for each work when it later purchased print copies of the same titles. Assuming for the
15 moment that this argument is admissible, it is hard to see how this would not also be evidence
16 in common across the broad category of those for whom it is true.

17 As for any embedded assessment of actual damages, common evidence and methods
18 again suffice. Anthropic proffered in its motion for summary judgment on fair use that there
19 was no world in which it would have been able to negotiate individual deals. Indeed, our
20 entire record reflects as much — not only for Anthropic, but for every other AI company to
21 have purchased books. Books were in every case purchased according to either a blanket
22 price, a tiered price based on subject matter or the like, or else according to a mass-market
23 price. Debating which of these options is appropriate, and the resulting price to be applied to
24 each book here, is susceptible to common evidence and methods. Indeed, for many if not all
25 the works, one piece of common evidence already exists as to what Anthropic would have
26 paid: the price it later paid to purchase any of the same books (assuming admissible) (*see*
27 Br. Exhs. 12, 22).

28 The summary judgment order makes clear that Anthropic is liable for all pirated copies

1 regardless of whether they were used for training and regardless of whether another copy of the
2 book was bought and scanned. To the extent a jury thinks these differences should affect the
3 amount of statutory damages the points can be made on a categorical basis rather than an
4 individual one.

5 **(iv) Plaintiffs Entitled to Recover?**

6 To recover, a claimant will have to submit a claim under oath stating that it is the owner
7 of the copyright interest infringed and stating the title, author, publisher and ISBN and/or
8 ASIN for the claimed work and further stating that the claimant serve notice to the publisher or
9 any other person with an interest to permit the person to contest the claim. In the district
10 judge's experience and judgment, very few disputes over ownership will arise but if they do
11 they can be resolved by the district judge on a summary judgment or by a jury at trial (or by
12 excluding the work altogether from the class).

13 Anthropic asserts that until ownership is proven all putative class members lack standing,
14 so its due process rights will be violated to the extent it is made to answer claims that a
15 copyright was infringed. But named plaintiffs plainly have standing and the class will be
16 defined in a way to bind all actual and beneficial owners of every work it stole, so this
17 objection falls away.

18 Anthropic also argues that managing ownership inquiries at scale will necessarily entail
19 "Trial by Formula" (Supp. Definitions Opp. 2 (citing *Wal-Mart Stores*, 564 U.S. at 367)). A
20 "Trial by Formula" classically involves sampling what happened to some class members,
21 assuming the same happened to all, and not allowing rebuttal. That is not the plan. Plaintiffs
22 shall present, among other things, a list of books with their authors, publishers, and
23 ISBN/ASIN identifiers that Anthropic downloaded. This list will be prepared (or should be)
24 with sufficient care to satisfy *Daubert*, no doubt relying in part on defendant's own records.
25 The class definition is limited to actual and beneficial owners of the copyrights involved. In
26 the district judge's judgment from experience (of fifty years), to repeat, there will be very few
27 competing claimants for the same work. This is because authors and their publishers have
28 ongoing business relationships and they will work out whatever differences (if any) they have

1 over how to divide the recovery.

2 The core claim of the class here is infringement of the right to reproduce copies under
3 Section 106(1) by way of episodes of mass downloading of timely registered copyrighted
4 books. All the claimants can be made out through sworn declarations subject to challenge, in
5 some cases.

6 (v) ***Fair Use?***

7 To the extent fair use were to recur as an issue at trial, there is no reason to believe it
8 would not be susceptible to treatment by common evidence and common methods.

9 * * *

10 At bottom, the major elements of this case concern a common course of conduct that can
11 be established with common evidence and methods. The ownership inquiry is where Anthropic
12 puts its focus — but its objections have been in the main addressed by plaintiffs’ concession
13 that they seek to enforce rights under Section 106(1) and that the focus of the LibGen &
14 PiLiMi Pirated Books Class is on classic reproductive copying by piracy. “We do not permit a
15 defendant to support its invocation of individualized issues with mere speculation.” *Van*, 61
16 F.4th at 1068. And, Anthropic’s option to rebut otherwise completely sufficient evidence by
17 itself “does not preclude the predominance of common questions.” *Cf. Blackie v. Barrack*, 524
18 F.2d 891, 906 n.22 (1975) (early fraud-on-the-market case where defendant enjoyed right to
19 rebut presumption but was unlikely to exercise it or to succeed in doing so). “What a district
20 court may not do is to assume, *arguendo*, that problems will arise, and decline to certify the
21 class on the basis of a mere potentiality that may or may not be realized.” *United Steel*
22 *Workers v. Conoco-Phillips Co.*, 593 F.3d 802, 810 (9th Cir. 2010).

23 This is how Judge Marilyn Patel, one of the great jurists of this district court, assessed
24 predominance in the *Napster* case, which is very close to our case:

25 [Although] ownership, registration, and actual damages ultimately
26 require[] a work-by-work inquiry, viewing these determinations as
27 purely “individual issues” ignores the fact that the claims of every
28 member of the class are uniformly premised upon the uploading or
downloading of a copyrighted work by Napster users. This shared
factual predicate in turn gives rise to a host of common legal issues
concerning Bertelsmann’s involvement in the operation of the

Napster network. There can be no serious dispute that these issues are sufficiently “significant” to warrant adjudication of the parties’ dispute on a representative rather than individual basis, nor is there any question that considerations of judicial economy heavily favor litigating these common issues once, as part of a single class action, rather than rehashing the same questions of law and fact in each of what could likely amount to thousands of individual lawsuits.

In re Napster, Inc. Copyright Litig., No. C MDL-00-1369 MHP, 2005 WL 1287611, at *7 (N.D. Cal. June 1, 2005) (citations omitted); *see also Tyson Foods, Inc. v. Bouaphakeo*, 577 U.S. 442, 453 (2016).

Predominance does not require that every issue be provable by a single method of proof. It does not rule out that one group may involve extra issues. It does require that the broad and differing nuances among the class be treatable by categorical proof as to each nuance. If the nuances are so varied and so many as to be unmanageable by a jury, then predominance is lacking. But if the differences are manageable, then predominance is satisfied. So here.

Common issues and evidence heavily *predominate*.

B. SUPERIORITY.

This inquiry “at [its] very core” assesses whether the claims could not be resolved but for a class action. *Amchem Prods, Inc. v. Windsor*, 521 U.S. 591, 617 (1997). Superiority is met where there is insufficient “incentive for any individual to bring a solo action prosecuting his or her rights” — and therefore insufficient opportunity for any defendant to conclusively vindicate its rights in turn — while by contrast a class action would “aggregat[e] the relatively paltry potential recoveries into something worth someone’s (usually an attorney’s) labor.” *Ibid.* (quoting *Mace v. Van Ru Credit Corp.*, 109 F.3d 338, 344 (7th Cir. 1997)). We answer this question by considering the interests:

(1) of the judicial system, (2) of the potential class members, (3) of the [named] plaintiff[s], (4) of the attorneys for the litigants, (5) of the public at large and (6) of the defendant. The listing is not necessarily in order of importance of the respective interests. Superiority must also be looked at from the point of view [7] of the issues [and legal regime itself].

Bateman, 623 F.3d at 713 (quotation omitted); *see also Fed. R. Civ. P. 23(b)(3)*.

Here, if not brought as a class action, there will likely be no action at all. And, in any case, only a class action that concentrates plaintiffs’ interests and defendant’s resources into one confrontation will present the claims and defenses with a quality commensurate with the issues’ import. Named plaintiffs already have undertaken extensive discovery and litigation, and have expressed their preference for a class action. For good reason: Defendant is formidable. Few if any potential class members alone

could muster comparable resources [to those of defendant], nor is there any guarantee [any alone] could find counsel willing to work pro bono or on a contingency basis. Pooling these burdens among a group of [plaintiffs] helps more [] bring claims. Proceeding as a class action would also maximize efficiency, particularly in light of the considerable discovery the parties have already undertaken.

See Doe v. Mindgeek USA Inc., 702 F. Supp. 3d 937, 952 (C.D. Cal. 2023) (Judge Cormac Carney). Meanwhile, defendant now sits under a cloud of copyright liability — needlessly, assuming it is correct on the facts and on the law — and cannot likely dispel that cloud except by class action. Other stakeholders and the public at large stand to benefit from the clarity only a robustly litigated class action could provide — from authors who wish to know how copies of their works are being used in fact, to other AI companies that would prefer to pirate rather than pay for those works for whatever uses they might have, to people who might not engage in one activity or the other without more clarity.

In opposition, Anthropic offers decisions finding a class action mechanism inappropriate where, for instance, harm was caused “in different individuals in different states by different doctors.” *Zinser v. Accufix Rsch. Inst., Inc.*, 253 F.3d 1180, 1192 (9th Cir.) (quoting favorably *Haley v. Medtronic, Inc.*, 169 F.R.D. 643, 654 (C.D. Cal. 1996) (Judge William Rea)), *amd’d*, 273 F.3d 1266 (9th Cir. 2001). Here, different copyright owners had their copyrighted works downloaded from one of three pirate libraries by one company and kept forever.

Anthropic next argues that “[t]he prospect of ruinous statutory damages — \$150,000 times [five] million books — only compounds the significant due process implications of certification here” (Opp. 23). But our court of appeals has squarely held that it is improper to consider whether or not “defendant’s potential liability [to the class was] ‘enormous and

completely out of proportion to any harm suffered by [the named] Plaintiff.” *Bateman*, 623 F.3d at 712–13 (quoting district court). “If the size of a defendant’s potential liability alone was a sufficient reason to deny class certification, [in other words], the very purpose of Rule 23(b)(3) — ‘to allow integration of numerous small individual claims into a single powerful unit’ — would be substantially undermined.” *Id.* at 722 (quoting *Blackie*, 524 F.2d at 899, and further citing a long line of cases in accord). Here, there is a “legislative decision to authorize awards as high as [\$150,000 per work], combined with multiple violations of the statute.” *Ibid.* (cleaned up). That decision is enforceable, and enforceable most efficiently here by class action.

* * *

There is no serious prospect that these claims can be addressed through individual actions. A denial of a motion to certify the class would amount to a concession that copyright owners’ credible allegations of infringement will go unchecked by courts so long as a copyist allegedly violates the Copyright Act not a little but a whole lot. “[T]o do so would eliminate the class action deterrent for those who engage in complicated and imaginative rather than straightforward schemes” — or, in this instance, straightforward piracy but at massive scale. *Cf. Blackie*, 524 F.2d at 903 n.19 (re Rule 10b-5). It cannot be true that as the scope of copying grows, so does the impunity.

Proceeding as a class is superior. The LibGen & PiLiMi Pirated Books Class thus meets Rule 23(b)(3). This order need not reach the alternative basis for certifying the class, under Rule 23(b)(2) (nor consider certifying issues under Rule 23(c)(4)). The class also meets Rule 23(a), so may be certified.

3. NOTICE TO POTENTIAL CLASS MEMBERS.

Under Rule 23(c)(2)(B), a court must direct to class members the best notice that is practicable under the circumstances. That proposed class notice shall provide all the usual language *and the following*:

The notice shall state that any class member shall be bound by the outcome and judgment and shall not (unless it timely opts out) be allowed to relitigate any claim certified for class

1 treatment against Anthropic. Plaintiffs' counsel shall provide notice to all authors, publishers,
2 copyright-registered authors (if different), and copyright-registered claimants (if different)
3 whose works were downloaded, and to all major, regional, and university presses. Notice shall
4 be by first-class mail, by email, and by publication in at least one national trade journal, all
5 costs of notice to be borne by plaintiffs (but recoverable as costs upon judgment). To assist, by
6 **NOON ON AUGUST 1, 2025**, defendant shall provide the titles, authors, publishers, ISBNs
7 and/or ASINs data to plaintiffs' counsel to the extent such information about the two pirated
8 sets is in its possession or that of its attorneys. This is a court-ordered interrogatory and must
9 be honored.

10 The class notice shall further state that upon any judgment favoring the class, only those
11 who submit satisfactory claims shall recover. A satisfactory claim form shall identify the class
12 claimant and the works claimed. It shall state under oath that the claimant owns the copyright
13 to make copies of the book or receives royalties in exchange for having granted to another the
14 exclusive right to make copies of the book. And, it shall further state under oath that the
15 claimant has served a copy of its claim form to anyone with whom it has entered an agreement
16 respecting the copyright and to any other publisher, author, or person he suspects has any
17 copyright interest of any type in the work (so that such persons can submit competing claims if
18 they really think they own the right to make copies). No payment shall be made until **70 DAYS**
19 have passed after mailing of such notice to others. If a competing claim is in fact made, then
20 the district judge or jury if needed will decide ownership.

21 This will be a claims-made process and the benefit to the class will be determined, in
22 part, by the number of works infringed for which a rightful owner is found, a factor which may
23 influence the amount of any attorney's fees award.

24 By **NOON ON AUGUST 15, 2025**, both sides shall submit a single agreed-on proposed
25 form of notice for the judge's review as well as a plan for distribution of notice. Remember,
26 notice will not be sent until the list of works involved (complete with title, author, publisher,
27 copyright-registered author (if different), copyright-registered claimant (if different),
28 copyright-registration number, and ISBN and/or ASIN) is submitted by plaintiffs' counsel.

CONCLUSION

The motion for class certification is **GRANTED** as to one of the alternative pirated books classes, and otherwise denied, as follows:

1. The following LibGen & PiLiMi Pirated Books Class is CERTIFIED:

All beneficial or legal copyright owners of the exclusive right to reproduce copies of any book in the versions of LibGen or PiLiMi downloaded by Anthropic. “Book” refers to any work possessing an ISBN or ASIN which was registered with the United States Copyright Office within five years of the work’s publication and which was registered with the United States Copyright Office before being downloaded by Anthropic, or within three months of publication. Excluded are the directors, officers and employees of Anthropic, personnel of federal agencies, and district court personnel.

Because this order grants certification to this class under Rule 23(b)(3), it declines to consider alternative bases for certification proposed by plaintiffs under Rule 23(b)(2) or Rule 23(c)(4). Defendant is compelled to produce related information described above by noon on August 1, 2025. This certification is contingent upon plaintiffs’ satisfactory filing of the list described above by noon on September 1, 2025.

2. Certification of a Book3 Pirated Books Class is DENIED, as is certification of a Scanned Books Class.

3. As Representative Plaintiffs for the LibGen & PiLiMi Pirated Books Class, the following person and entities are **DESIGNATED**: Andrea Bartz, Inc., Charles Graeber, and MJ + KJ Inc.

4. As LibGen & PiLiMi Pirated Books Class Counsel, the following law firms are **APPOINTED**: Lieff Cabraser Heimann & Bernstein, LLP and Susman Godfrey LLP.

5. Counsel from both sides shall submit a joint proposal for the form of class notice and for the dissemination of class notice to conform with the standards set forth above, by noon on August 15, 2025.

IT IS SO ORDERED.

Dated: July 17, 2025.

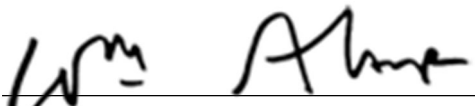

WILLIAM ALSUP
UNITED STATES DISTRICT JUDGE

EXHIBIT B

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

ANDREA BARTZ, CHARLES GRAEBER,
and KIRK WALLACE JOHNSON,

No. C 24-05417 WHA

Plaintiffs,

v.

ORDER ON FAIR USE

ANTHROPIC PBC,

Defendant.

INTRODUCTION

An artificial intelligence firm downloaded for free millions of copyrighted books in digital form from pirate sites on the internet. The firm also purchased copyrighted books (some overlapping with those acquired from the pirate sites), tore off the bindings, scanned every page, and stored them in digitized, searchable files. All the foregoing was done to amass a central library of “all the books in the world” to retain “forever.” From this central library, the AI firm selected various sets and subsets of digitized books to train various large language models under development to power its AI services. Some of these books were written by plaintiff authors, who now sue for copyright infringement. On summary judgment, the issue is the extent to which any of the uses of the works in question qualify as “fair uses” under Section 107 of the Copyright Act.

STATEMENT

Defendant Anthropic PBC is an AI software firm founded by former OpenAI employees in January 2021. Its core offering is an AI software service called Claude. When a user

prompts Claude with text, Claude quickly responds with text — mimicking human reading and writing. Claude can do so because Anthropic trained Claude — or rather trained large language models or LLMs underlying various versions of Claude — using books and other texts selected from a central library Anthropic had assembled. Claude was first released publicly in March 2023. Seven successive versions of Claude have been released since. Users may ask Claude some questions for free. Demanding users and corporate clients pay to use Claude, generating over one billion dollars in annual revenue (Opp. Exh. 18).

Plaintiffs Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson are authors of books that Anthropic copied from pirated and purchased sources. Anthropic assembled these copies into a central library of its own, copied further various sets and subsets of those library copies to include in various “data mixes,” and used these mixes to train various LLMs. Anthropic kept the library copies in place as a permanent, general-purpose resource even after deciding it would not use certain copies to train LLMs or would never use them again to do so. All of Anthropic’s copying was without plaintiffs’ authorization.

Author Bartz wrote four novels Anthropic copied and used: *The Lost Night: A Novel*, *The Herd*, *We Were Never Here*, and *The Spare Room*. Author Graeber wrote two non-fiction books likewise at issue: *The Good Nurse: A True Story of Medicine, Madness, and Murder*, and *The Breakthrough: Immunotherapy and the Race to Cure Cancer*. And, Author Johnson penned three non-fiction books also copied and used: *To Be A Friend Is Fatal: The Fight to Save the Iraqis America Left Behind*, *The Feather Thief: Beauty, Obsession, and the Natural History Heist of the Century*, and *The Fishermen and the Dragon: Fear, Greed, and a Fight for Justice on the Gulf Coast*. Plaintiffs Bartz Inc. and MJ + KJ Inc. are corporate entities that Author Bartz and Author Johnson respectively set up to market their works. Between them, these five plaintiffs (“Authors”) own all the copyrights in the above-listed works.

From the start, Anthropic “ha[d] many places from which” it could have purchased books, but it preferred to steal them to avoid “legal/practice/business slog,” as cofounder and chief executive officer Dario Amodei put it (*see* Opp. Exh. 27). So, in January or February 2021, another Anthropic cofounder, Ben Mann, downloaded Books3, an online library of

1 196,640 books that he knew had been assembled from unauthorized copies of copyrighted
2 books — that is, pirated. Anthropic’s next pirated acquisitions involved downloading
3 distributed, reshared copies of other pirate libraries. In June 2021, Mann downloaded in this
4 way at least five million copies of books from Library Genesis, or LibGen, which he knew had
5 been pirated. And, in July 2022, Anthropic likewise downloaded at least two million copies of
6 books from the Pirate Library Mirror, or PiLiMi, which Anthropic knew had been pirated
7 (Opp. Exh. 6 at 4; Opp. Expert Zhao ¶¶ 17–29; *see* Class Cert. (“CC”) Opp. Expert Iyyer
8 ¶¶ 45–46). Although what was downloaded and later duplicated from these sources was
9 sometimes referred to as data or datasets, at bottom they contained full-text “ebooks or scans of
10 books” saved in individual files in formats like .pdf, .txt, and .epub (*see, e.g.*, Opp. Exh. 12 at -
11 0391318). For Books3, most filenames identified the book inside. For LibGen and PiLiMi,
12 Anthropic downloaded a separate catalog of bibliographic metadata for each collection, with
13 fields like title, author, and ISBN (*see, e.g., ibid.*; Opp. Exh. 16 -0533972–73). Anthropic
14 thereby pirated over seven million copies of books, including copies of at least two works at
15 issue for each Author.¹

16 As Anthropic trained successive LLMs, it became convinced that using books was the
17 most cost-effective means to achieve a world-class LLM. During this time, however,
18 Anthropic became “not so gung ho about” training on pirated books “for legal reasons” (Opp.
19 Exh. 19). It kept them anyway (*e.g.*, Opp. Exh. 17 at 93–94; CC Opp. Exh. 35 at -0273474).

20 To find a new way to get books, in February 2024, Anthropic hired the former head of
21 partnerships for Google’s book-scanning project, Tom Turvey. He was tasked with obtaining
22 “all the books in the world” while still avoiding as much “legal/practice/business slog” as
23

24 ¹ Specifically, those works were (*see* Opp. Expert Zhao ¶ 36; CC Br. Expert Zhao ¶ 66):
25 • Author Bartz’s *The Herd* (five copies total) (in LibGen and PiLiMi);
26 • Author Bartz’s *The Lost Night* (three copies total) (in Books3, LibGen, and PiLiMi);
27 • Author Graeber’s *The Breakthrough* (four copies) (in Books3, LibGen, and PiLiMi);
28 • Author Graeber’s *The Good Nurse* (five copies total) (in Books3 and LibGen);
• Author Johnson’s *To Be A Friend Is Fatal* (one copy) (in Books3); and
• Author Johnson’s *The Feather Thief* (four copies total) (in Books3, LibGen, PiLiMi).
Some evidence suggests Anthropic downloaded still more copies before culling empty files,
duplicates, and so on to reach the numbers kept in the central library and counted here.

possible (Opp. Exhs. 21, 27). So, in spring 2024, Turvey sent an email or two to major publishers to inquire into licensing books for training AI. Had Turvey kept up those conversations, he might have reached agreements to license copies for AI training from publishers — just as another major technology company soon did with one major publisher (e.g., Opp. Expert Malackowski ¶¶ 50, 64). But Turvey let those conversations wither.

Instead, Turvey and his team emailed major book distributors and retailers about bulk-purchasing their print copies for the AI firm’s “research library” (Opp. Exh. 22 at 145; Opp. Exh. 31 at -035589). Anthropic spent many millions of dollars to purchase millions of print books, often in used condition. Then, its service providers stripped the books from their bindings, cut their pages to size, and scanned the books into digital form — discarding the paper originals. Each print book resulted in a PDF copy containing images of the scanned pages with machine-readable text (including front and back cover scans for softcover books). Anthropic created its own catalog of bibliographic metadata for the books it was acquiring. It acquired copies of millions of books, including of all works at issue for all Authors.²

Anthropic may have copied portions of Authors’ books on other occasions, too — such as while copying book reviews, academic papers, internet blogposts, or the like for its central library. And, Anthropic’s scanning service providers may have copied Authors’ print books along the way to delivering the final digital copies to Anthropic. But neither side here specifically raises legal issues implicated by any such copies. Nor will this order.

From all the above sources, Anthropic created a general “research library” or “generalized data area.” What was this for? As Turvey said, this was a “way of creating information that would be voluminous and that we would use for research,” or otherwise to

² In other words, within the scanned books were one or more copies of the following works:

- Author Bartz’s *The Herd*;
- Author Bartz’s *The Lost Night*;
- Author Bartz’s *We Were Never Here*;
- Author Bartz’s *The Spare Room*;
- Author Graeber’s *The Breakthrough*;
- Author Graeber’s *The Good Nurse*;
- Author Johnson’s *To Be A Friend Is Fatal*;
- Author Johnson’s *The Feather Thief*; and,
- Author Johnson’s *The Fishermen*.

“inform our — our products” (Opp. Exh. 22 at 145–46, 194). The copies were kept in the original “version of the underlying” book files Anthropic had “obtained or created,” that is, pirated or scanned (Opp. Exh. 30 at 3, 4). Anthropic planned to “store everything forever; we might separate out books into categories[, but t]here [wa]s no compelling reason to delete a book” — even if not used for training LLMs. Over time, Anthropic invested in building more tools for searching its “general purpose” library and for accessing books or sets of books for further uses (*see* CC Br. Exh. 12 at -0144509; CC Reply Exh. 45 at -0365931–32, -0365939–42 (reviewing and seeking to improve “[w]hat [] researchers do today if they want to search for a book,” including improving bibliographic metadata and consolidating varied resources)).

One further use was training LLMs. As a preliminary step towards training, engineers browsed books and bibliographic metadata to learn what languages the books were written in, what subjects they concerned, whether they were by famous authors or not, and so on — sometimes by “open[ing] any of the books” and sometimes using software. From the library copies, engineers copied the sets or subsets of books they believed best for training and “iterate[d]” on those selections over time. For instance, two different subsets of print-sourced books were included in “data mixes” for training two different LLMs. Each was just a fraction of all the print-sourced books. Similarly, different sets or “subsets” or “parts of” or “portions” of the collections sourced from Books3, LibGen, and PiLiMi were used to train different LLMs. Anthropic analyzed the consequences of using more books, fewer books, different books. The goal was to improve the “data mix” to improve each LLM and, ultimately, Claude’s performance for paying customers.³

³ (*See, e.g.*, Opp. Exh. 12 at -0391318 (engineers were able to “open any of the books”); CC Reply Exh. 45 at -0365941 (some engineers “want[ed] to search for a book” and get its “scanned book file[]”); Opp. Exh. 30 at 3 (made copies of “each such dataset or portions thereof” for training); Opp. Exh. 6 at 3–4 (trained on “portions of datasets,” with at least two such portions from LibGen and four from PiLiMi); Opp. Expert Zhao ¶¶ 27–28, 30–31 (plus two more from PiLiMi, and at least three from scanned books); CC Opp. Exh. 35 at -0273477–82 (tested subsets of pirated and purchased-and-scanned books to see consequences for training); CC Br. Exh. 12 at -0144508–09 (“iterate[d]” selections from library and “train[ed] new models on the best data”); Br. Expert Kaplan ¶¶ 42–45 (explained goals of improving data mixes); Br. Expert Peterson ¶ 14 (similar)).

Over time, Anthropic came to value most highly for its data mixes books like the ones Authors had written, and it valued them because of the creative expressions they contained. Claude’s customers wanted Claude to write as accurately and as compellingly as Authors. So, it was best to train the LLMs underlying Claude on works just like the ones Authors had written, with well-curated facts, well-organized analyses, and captivating fictional narratives — above all with “good writing” of the kind “an editor would approve of” (Opp. Exh. 3 at -03433). Anthropic could have trained its LLMs without using such books or any books at all. That would have required spending more on, say, staff writers to create competing exemplars of good writing, engineers to revise bad exemplars into better ones, energy bills to power more rounds of training and fine-tuning, and so on. Having canonical texts to draw upon helped (e.g., Opp. Expert Zhao ¶ 81).

Each work selected for training any given LLM was copied in four main ways — and in fact so many times that Anthropic admits it would be impractical even to estimate.

First, each work selected was copied from the central library to create a working copy for the training set.

Second, each work was cleaned to remove a small amount of lower-valued or repeating text (like headers, footers, or page numbers), with a “cleaned” copy resulting. If the same book appeared twice, or if while looking across the entire provisional training set it became clear there was some other reason to cull a book or category, Anthropic had the capability to delete relevant copy(ies) from the set at this step (see CC Br. Expert Zhao ¶¶ 71–72).

Third, each cleaned copy was translated into a “tokenized” copy. Some words were “stemmed” or “lemmatized” into simpler forms (e.g., “studying” to “study”). And, all characters were grouped into short sequences and translated into corresponding number sequences or “tokens” according to an Anthropic-made dictionary. The resulting tokenized copies were then copied repeatedly during training. By one account, this process involved the iterative, trial-and-error discovery of contingent statistical relationships between each word fragment and all other word fragments both within any work and across trillions of word

fragments from other copied books, copied websites, and the like. Other steps in training are not at issue here (*id.* ¶¶ 73–76; *see* Opp. Expert Zhao ¶ 38 & n.6).

Fourth, each fully trained LLM itself retained “compressed” copies of the works it had trained upon, or so Authors contend and this order takes for granted. In essence, each LLM’s mapping of contingent relationships was so complete it mapped or indeed simply “memorized” the works it trained upon almost verbatim. So, if each completed LLM had been asked to recite works it had trained upon, it could have done so (*e.g.*, Opp. Expert Zhao ¶ 74). Further steps refining the LLM are not at issue here.

However, that was as far as the training copies propagated towards the outside world. When each LLM was put into a public-facing version of Claude, it was complemented by other software that filtered user inputs to the LLM and filtered outputs from the LLM back to the user (*id.* ¶¶ 75–77). As a result, Authors do not allege that *any* infringing copy of their works was or would ever be provided to users by the Claude service. Yes, Claude could help less capable writers create works as well-written as Authors’ and competing in the same categories. But Claude created no exact copy, nor any substantial knock-off. Nothing traceable to Authors’ works. Such allegations are simply not part of plaintiffs’ amended complaint, nor in our record.

Neither side puts directly at issue any copies of any works that might have been used for the filtering software. Nor will this order.

In sum, the copies of books pirated or purchased-and-destructively-scanned were placed into a central “research library” or “generalized data area,” sets or subsets were copied again to create training copies for data mixes, the training copies were successively copied to be cleaned, tokenized, and compressed into any given trained LLM, and once trained an LLM did not output through Claude to the public any further copies. Finally, once Anthropic decided a copy of a pirated or scanned book in the library would not be used for training at all or ever again, Anthropic still retained that work as a “hard resource” for other uses or future uses. At least one work from each Author was present in every phase described above.

* * *

In August 2024, the three individual authors brought this putative class action complaining that Anthropic had infringed its federal copyrights by pirating copies for its library and by reproducing them to train its LLMs (Compl. ¶¶ 45–46, 71; *see* Amd. Compl. ¶¶ 47–48, 75). In October 2024, a scheduling order required that any motion for class certification be brought by March 6, 2025 (Dkt. No. 49).

The individual authors soon amended their complaint to include affiliated corporate entities as named plaintiffs, with consent. And, Anthropic chose not to move to dismiss the amended complaint, as it earlier had planned (*see* Dkt. No. 37). Instead, Anthropic moved to allow an early motion for summary judgment on fair use, even before class certification (Dkt. No. 88; *see* Feb. 25, 2025 Tr. 15). Permission was granted.

Anthropic now moves for summary judgment on fair use only. Fair use is a legal question for the judge with underlying fact questions, if any, for the jury. To prevail on summary judgment, Anthropic must rely on undisputed facts and/or factual inferences favoring the opposing side. Anthropic thus bears the burdens of production and persuasion in this motion. *See Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 23–24 (2021); *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 547 n.21 (2023); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 & n.20, 594 (1994); *see also Nissan Fire & Marine Ins. Co. v. Fritz Cos.*, 210 F.3d 1099, 1102–03 (9th Cir. 2000).

Notably, in its motion, Anthropic argues that pirating initial copies of Authors’ books and millions of other books was justified because all those copies were at least reasonably necessary for training LLMs — and yet Anthropic has resisted putting into the record what copies or even sets of copies were in fact used for training LLMs. For example, at oral argument, Anthropic asserted that if a purported fair user had retained pirated copies for uses beyond the fair use, then her piracy would not be excused by the fair use (Tr. 53, 56). But when Authors earlier interrogated Anthropic in discovery about what library copies (the original copies “obtained or created” by Anthropic) Anthropic had recopied for further uses, Anthropic responded that providing information about any copies made for uses beyond training commercially released LLMs would be overbroad, and that it could not count up all its

copying even for LLMs in any case (*e.g.*, Opp Exh. 30 at 3). We know that Anthropic has more information about what it in fact copied for training LLMs (or not). Anthropic earlier *produced* a spreadsheet that showed the composition of various data mixes used for training various LLMs — yet it clawed back that spreadsheet in April (Opp. Fredricks Decl. ¶¶ 2–3). A discovery dispute regarding that spreadsheet remains pending. But Anthropic did not need a court order to offer up what it possessed in support of its motion. All deficiencies must be held against Anthropic and not the other way around.

This is the first substantive order in this case. A contemporaneous motion for class certification remains pending. It proposes one class related to works that were pirated (whether or not used to train LLMs), and a second class related to works that were purchased, scanned, and used in training LLMs. This order follows full briefing, a hearing, and supplemental briefing.

To summarize the analysis that now follows, the use of the books at issue to train Claude and its precursors was exceedingly transformative and was a fair use under Section 107 of the Copyright Act. And, the digitization of the books purchased in print form by Anthropic was also a fair use but not for the same reason as applies to the training copies. Instead, it was a fair use because all Anthropic did was replace the print copies it had purchased for its central library with more convenient space-saving and searchable digital copies for its central library — without adding new copies, creating new works, or redistributing existing copies. However, Anthropic had no entitlement to use pirated copies for its central library. Creating a permanent, general-purpose library was not itself a fair use excusing Anthropic’s piracy.

ANALYSIS

Section 107 of the Copyright Act identifies four factors for determining whether a given use of a copyrighted work is a fair use:

[T]he fair use of a copyrighted work . . . for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

These factors presuppose a “use.” So, at the threshold, a court must decide whether a “copyrighted [work] has been used in multiple ways,” then evaluate each. *Warhol*, 598 U.S. at 533. Uses do not turn on “the subjective intent of the user” but on “an objective inquiry into what use was made, *i.e.*, what the user d[id] with the original work.” *Id.* at 544–45. A “use” should be construed narrowly enough to not “swallow” distinguishable infringing uses, much less categories of exclusive rights *in toto*. *Id.* at 541, 543 n.18, 546–48. Sometimes, the challenged copying involves just one use: In *Perfect 10, Inc. v. Amazon.com, Inc.*, Google visited websites having full-sized images, made only reduced-sized copies, and incorporated those directly into its search engine — the sole use of the thumbnails being as “pointer[s]” to the images themselves. 508 F.3d 1146, 1157, 1160, 1165 (9th Cir. 2007). Sometimes, the copying involves many uses: In the *Google Books* cases, Google borrowed books from libraries, made both full-image and text-only copies, and incorporated different copies into different tools — one use being to reveal information “*about those books*,” another use being to provide the books to print-disabled patrons, and still another being to back up the print books if lost. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 217 (2d Cir. 2015) (quoted); *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 97, 101, 103 (2d Cir. 2014) (other cited uses).

Our parties debate an instructive decision. In *American Geophysical Union v. Texaco Inc.*, Texaco employees used scientific articles in a central library, used copies of them in personal desk libraries, and used selected copies again in the scientific laboratory — the first use paid for, the second infringing, and the third plausibly fair but in fact a rare occurrence. 802 F. Supp. 1, 4–5, 14 (S.D.N.Y. 1992) (Judge Pierre Leval), *aff’d*, 60 F.3d 913, 918–19, 926 (2d Cir. 1994).

Here, our parties contest what use or uses are at issue. Anthropic contends it copied Authors’ books only for *one* use: *Only* to train LLMs. By contrast, Authors contend it did so for at least *two* uses: *First* to build a vast, central library of potentially useful content, and *second* to train specific LLMs using shifting sets and subsets of that content — over time selecting the more well-organized and well-expressed works for training. Authors also complain that the print-to-digital format change was itself an infringement not abridged as a fair use (Opp. 15, 25). Authors do not allege, however, that any LLM outputs infringing upon their works ever reached users of the public-facing Claude service.

This order addresses each of the four factors in turn, pointing out how each applies to the training copies and to the purchased and pirated library copies. It concludes with an integrated analysis.

1. THE PURPOSE AND CHARACTER OF THE USE.

For a given use at issue, the first factor addresses “the purpose and character of th[at] use, including whether [it] is of a commercial nature or is for nonprofit educational purposes.” 17 U.S.C. § 107(1).

A. THE COPIES USED TO TRAIN SPECIFIC LLMs.

All agree that one use at issue was training LLMs to receive text inputs and return text outputs. More specifically, Anthropic used copies of Authors’ copyrighted works to iteratively map statistical relationships between every text-fragment and every sequence of text-fragments so that a completed LLM could receive new text inputs and return new text outputs as if it were a human reading prompts and writing responses. Authors further argue — and this order takes for granted — that such training entailed “memoriz[ing]” works by “compress[ing]” copies of those works into the LLM (Opp. 16–17; *see* Opp. Expert Zhao ¶ 74). The LLMs “memorize[d] A LOT, like A LOT” (Opp. Exh. 35 at -029109). Regardless, the “purpose and character” of using works to train LLMs was transformative — spectacularly so.

To repeat and be clear: Authors do not allege that any LLM output provided to users infringed upon Authors’ works. Our record shows the opposite. Users interacted only with the Claude service, which placed additional software between the user and the underlying LLM to

1 ensure that no infringing output ever reached the users. This was akin to the limits Google
2 imposed on how many snippets of text from any one book could be seen by any one user
3 through its Google Books service, preventing its search tool from devolving into a reading tool.
4 *Google*, 804 F.2d at 222. Here, if the outputs seen by users had been infringing, Authors
5 would have a different case. And, if the outputs were ever to become infringing, Authors
6 could bring such a case. But that is not this case.

7 Instead, Authors challenge only the inputs, not the outputs, of these LLMs. They point to
8 the fully trained LLMs and the Claude service only to shed light on how training itself uses
9 copies of their works and the ways the Claude service could be used to produce still other
10 works that would compete with their works. This order does the same. Authors' arguments
11 that the training use is not transformative are unavailing.

12 *First*, Authors argue that using works to train Claude's underlying LLMs was like using
13 works to train any person to read and write, so Authors should be able to exclude Anthropic
14 from this use (Opp. 16). But Authors cannot rightly exclude anyone from using their works for
15 training or learning as such. Everyone reads texts, too, then writes new texts. They may need
16 to pay for getting their hands on a text in the first instance. But to make anyone pay
17 specifically for the use of a book each time they read it, each time they recall it from memory,
18 each time they later draw upon it when writing new things in new ways would be unthinkable.
19 For centuries, we have read and re-read books. We have admired, memorized, and internalized
20 their sweeping themes, their substantive points, and their stylistic solutions to recurring writing
21 problems.

22 *Second*, to that last point, Authors further argue that the training was intended to
23 memorize their works' creative elements — not just their works' non-protectable ones (Opp.
24 17). But this is the same argument. Again, Anthropic's LLMs have not reproduced to the
25 public a given work's creative elements, nor even one author's identifiable expressive style
26 (assuming *arguendo* that these are even copyrightable). Yes, Claude has outputted grammar,
27 composition, and style that the underlying LLM distilled from thousands of works. But if
28 someone were to read all the modern-day classics because of their exceptional expression,

1 memorize them, and then emulate a blend of their best writing, would that violate the
2 Copyright Act? Of course not. Copyright does not extend to “method[s] of operation,
3 concept[s], [or] principle[s]” “illustrated[] or embodied in [a] work.” 17 U.S.C. § 102(b); *see*,
4 *e.g.*, *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 120–22 (2d Cir. 1930) (Judge Learned
5 Hand) (stage properties and storytelling elements); *Apple Comput., Inc. v. Microsoft Corp.*, 35
6 F.3d 1435, 1445 (9th Cir. 1994) (“user-friendly” design principles and elements); *Swirsky v.*
7 *Carey*, 376 F.3d 841, 848 (9th Cir. 2004) (music theory principles and chord progressions).

8 *Third*, Authors next argue that computers nonetheless should not be allowed to do what
9 people do.

10 Authors cite a decision seeming to say as much (Opp. 16–17). But the judge there twice
11 emphasized while discussing “purpose and character” of the use that what was trained was “not
12 generative AI (AI that writes new content itself).” Rather, what was trained — using a
13 proprietary system for finding court opinions in response to a given legal topic — was a
14 competing AI tool for finding court opinions in response to a given legal topic. That was not
15 transformative. *Thomson Reuters Enter. Centre GmbH v. Ross Intell. Inc.*, 765 F. Supp. 3d
16 382, 398 (D. Del. 2025) (Judge Stephanos Bibas), *appeal docketed*, No. 25-8018 (3d Cir. Apr.
17 14, 2025).

18 A better analogue to our facts would be an AI tool trained — using court opinions, and
19 briefs, law review articles, and the like — to receive legal prompts and respond with fresh legal
20 writing. And, on facts much like those, a different court came out the other way. It found fair
21 use. *White v. W. Pub. Corp.*, 29 F. Supp. 3d 396, 400 (S.D.N.Y. 2014) (Judge Jed Rakoff).

22 The latter use stood sufficiently “orthogonal” to anything that any copyright owner
23 rightly could expect to control. *See Warhol*, 598 U.S. at 538–40. It could thus be freed up for
24 the copyist to use, “promot[ing] the progress of science and the arts, *without* diminishing the
25 incentive to create.” *Id.* at 531 (emphasis added); *see* U.S. CONST. art. I, § 8, cl. 8.

26 In short, the purpose and character of using copyrighted works to train LLMs to generate
27 new text was quintessentially transformative. Like any reader aspiring to be a writer,
28 Anthropic’s LLMs trained upon works not to race ahead and replicate or supplant them — but

1 to turn a hard corner and create something different. If this training process reasonably
2 required making copies within the LLM or otherwise, those copies were engaged in a
3 transformative use.

4 The first factor *favors* fair use for the training copies.

5 ***B. THE COPIES USED TO BUILD A CENTRAL LIBRARY.***

6 But that is not the only use at issue. Recall that Anthropic purchased millions of print
7 books for its central library and pirated millions of digital books for its central library, too. It
8 used specific sets and subsets of books for training specific LLMs. And, it then retained all the
9 copies in its central library for other uses that might arise *even after deciding it would not use*
10 *them to train any LLM (at all or ever again)*. Anthropic seems to believe that because some of
11 the works it copied were sometimes used in training LLMs, Anthropic was entitled to take for
12 free all the works in the world and keep them forever with no further accounting. There is no
13 carveout, however, from the Copyright Act for AI companies.

14 Because the legal issues differ between the library copies Anthropic purchased and
15 pirated, this order takes them in turn.

16 ***(i) The Purchased Library Copies Converted from Print to Digital.***

17 Anthropic *purchased* millions of print copies to “build a research library” (Opp. Exh. 22
18 at 145, 148). It destroyed each print copy while replacing it with a digital copy for use in its
19 library (not for sharing nor sale outside the company). As to these copies, Authors do not
20 complain that Anthropic failed to pay to acquire a library copy. Authors only complain that
21 Anthropic changed each copy’s format from print to digital (*see* Opp. 15, 25 & n.14). On the
22 facts here, that format change itself added no new copies, eased storage and enabled
23 searchability, and was not done for purposes trenching upon the copyright owner’s rightful
24 interests — it was transformative.

25 Anthropic purchased its print copies fair and square. With each purchase came
26 entitlement for Anthropic to “dispose[]” each copy as it saw fit. 17 U.S.C. § 109(a). So,
27 Anthropic was entitled to keep the copies in its central library for all the ordinary uses. Yes,
28

Anthropic changed the format of these library copies from print to digital — giving rise to the issue here.

All agree on the facts of the format change. Anthropic “destructively scan[ned]” the print copies to create the digital ones. Anthropic or its vendors stripped the bindings from the print books, cut the pages to workable dimensions, and scanned those pages — discarding each print copy while creating a digital one in its place. The digital copy was then housed in the “research library” or “generalized data area” in place of the print copy (Opp. Exh. 22 at 145–46, 193–94). Authors do not allege and our record does not show that Anthropic provided its converted digital copies of print books to anyone outside Anthropic.

The parties disagree about the legal consequences of the format change. Was scanning the print copies to create digital replacements transformative? Anthropic argues it was because it was reasonably necessary to training LLMs. Authors argue it was a distinguishable step requiring independent justification.

Here, for reasons narrower than Anthropic offers, the mere format change was a fair use.

Storage and searchability are not creative properties of the copyrighted work itself but physical properties of the frame *around* the work or informational properties *about* the work. *See Texaco*, 802 F. Supp. at 14 (physical), *aff’d*, 60 F.3d at 919; *Google*, 804 F.3d at 225 (informational); *Sony Corp. of Am. v. Universal City Studios, Inc.* (“*Sony Betamax*”), 464 U.S. 417, 447 (1984) (rightful interests). In *Texaco*, the court reasoned that if a purchased scientific journal article had been copied “onto microfilm to conserve space, this might [have been] a persuasive transformative use.” 802 F. Supp. at 14 (Judge Pierre Leval), *aff’d*, 60 F.3d at 919 (reducing “bulk[]” “might suffice to tilt the first fair use factor in favor of *Texaco* if these purposes were dominant”). In *Google Books*, the court reasoned that a print-to-digital change to expose information about the work was transformative. *Google*, 804 F.3d at 225 (Judge Pierre Leval). And, in *Sony Betamax*, the Supreme Court held that making a recording of a television show in order to instead watch it at a later time was copying but did not usurp any rightful interest of the copyright owner. 464 U.S. at 447, 455. Important to the Supreme Court’s reasoning was the expectation that most such copiers would not distribute the

1 permanent copies of the work. Finally, in *A&M Records, Inc. v. Napster, Inc.*, our court of
2 appeals recognized the reasoning just explained, and therefore rejected by contrast a
3 digitization effort that was touted as space-shifting but in fact resulted in the multiplication of
4 copies shared with outsiders through a file-sharing service. 239 F.3d 1004, 1019 (9th Cir.
5 2001), *aff'g in this part* 114 F. Supp. 2d 896, 912–13, 915–16 (N.D. Cal. 2000) (Judge Marilyn
6 Hall Patel) (citing *Sony Betamax* and *Texaco*).

7 Here, every purchased print copy was copied in order to save storage space and to enable
8 searchability as a digital copy. The print original was destroyed. One replaced the other. And,
9 there is no evidence that the new, digital copy was shown, shared, or sold outside the company.
10 This use was even more clearly transformative than those in *Texaco*, *Google*, and *Sony*
11 *Betamax* (where the number of copies went up by at least one), and, of course, more
12 transformative than those uses rejected in *Napster* (where the number went up by “millions” of
13 copies shared for free with others).

14 Yes, Anthropic is a commercial outfit. And, this order takes for granted that Anthropic in
15 fact benefited from the print-to-digital format change — or it would not have gone to all the
16 trouble. But the crux of the first fair use factor’s concern for “commercial” use is in protecting
17 the copyright owners and their entitlements to exploit their copyright as they see fit (or not).
18 *See, e.g., Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985). That
19 the accused is a commercial entity is indicative, not dispositive. That the accused stands to
20 benefit is likewise indicative. But what matters most is whether the format change exploits
21 anything the Copyright Act reserves to the copyright owner. Anthropic already had purchased
22 permanent library copies (print ones). It did not create new copies to share or sell outside.

23 Yes, Authors also might have wished to charge Anthropic more for digital than for print
24 copies. And, this order takes for granted that Authors could have succeeded if Anthropic had
25 been barred from the format change. “But the Constitution’s language [in Clause 8] nowhere
26 suggests that [the copyright owner’s] limited exclusive right should include a right to divide
27 markets or a concomitant right to charge different purchasers different prices for the same
28 book, [merely] say to increase or to maximize gain.” *See Kirtsaeng v. John Wiley & Sons*,

Inc., 568 U.S. 519, 552 (2013); *see also* U.S. CONST. art. I., § 8, cl. 8. Nor does the Copyright Act itself. Section 106 sets out exclusive rights that fair uses under Section 107 abridge. Section 106(1) reserves to the copyright owner the right to make reproductions. But on our facts we face the unusual situation where one copy entirely replaced the another. And, Section 106(2) reserves to the copyright owner the right to make derivative works that add or subtract creative material — as occurs in a “translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, [or] condensation” of a book, 17 U.S.C. § 101 (definitions). For some “other modification[]” of a book to constitute a “derivative work,” it must itself “represent an original work of authorship.” *Ibid.* But on our facts the format was changed but no content was added or subtracted. *See Mirage Editions, Inc. v. Albuquerque A.R.T. Co.*, 856 F.2d 1341, 1342, 1343–44 (9th Cir. 1988) (yes where elements added to create new decorative ceramic).⁴ Section 106(3) further reserves to the copyright owner the right to distribute copies. But again, the replacement copy here was kept in the central library, not distributed. *Cf. Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 176–78 (2d Cir. 2018) (enabling searching for “information about the material” can be transformative use, even if some distribution results); *Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc.*, 964 F.2d 965, 968, 971 (9th Cir. 1992) (using nifty converter to “merely enhance[]” audiovisual displays emitted from purchased videogame cartridge was fair use of those displays partly because no surplus copies of cartridge or displays were ever created).

As a result, Anthropic’s format-change from print library copies to digital library copies was transformative under fair use factor one. Anthropic was entitled to retain a copy of these works in a print format. It retained them instead in a digital format, easing storage and

⁴ Even if print-to-digital format change *did* infringe the right to prepare derivative works, Authors have conceded that “Plaintiffs’ infringement claims are predicated on Anthropic’s unauthorized reproduction (17 U.S.C. § 106(1)); Plaintiffs are not alleging infringement by Anthropic of any right to prepare derivative works (*id.* at § 106(2))” (Dkt. No. 203 at 2 (citations original)). Whether this concession had consequence for copies tokenized and used for training or “compressed” into the trained LLMs is not reached by this order because Anthropic does not rely on Authors’ concession and those copies were here used transformatively.

searchability. And, the further copies made therefrom for purposes of training LLMs were themselves transformative for that further reason, as above.

To be clear, this print-to-digital conversion involved a different and narrower form of transformative use than the broader one advanced by Anthropic. Anthropic argues that the central library use was part and parcel of the LLM training use and therefore transformative. This order disagrees. However, this order holds that the mere conversion of a print book to a digital file to save space and enable searchability was transformative for that reason alone. Therefore, the digital copy should be treated just as if the purchased print copy had been placed in the central library.

In sum, the first fair use factor *favors* fair use for the digital library copies converted from purchased print library copies — but these do not excuse the pirated library copies.

(ii) The Pirated Library Copies.

Before buying books for its central library, Anthropic downloaded over seven million pirated copies of books, paid nothing, and kept these pirated copies in its library even after deciding it would not use them to train its AI (at all or ever again). Authors argue Anthropic should have paid for these pirated library copies (*e.g.*, Tr. 24–25, 65; Opp. 7, 12–13). This order agrees.

The basic problem here was well-stated by Anthropic at oral argument: “You can’t just bless yourself by saying I have a research purpose and, therefore, go and take any textbook you want. That would destroy the academic publishing market if that were the case” (Tr. 53). Of course, the person who purchases the textbook owes no further accounting for keeping the copy. But the person who copies the textbook from a pirate site has infringed already, full stop. This order further rejects Anthropic’s assumption that the use of the copies for a central library can be excused as fair use merely because some will eventually be used to train LLMs.

This order doubts that any accused infringer could ever meet its burden of explaining why downloading source copies from pirate sites *that it could have purchased or otherwise accessed lawfully* was itself reasonably necessary to any subsequent fair use. There is no decision holding or requiring that pirating a book that could have been bought at a bookstore

was reasonably necessary to writing a book review, conducting research on facts in the book, or creating an LLM. Such piracy of otherwise available copies is inherently, irredeemably infringing even if the pirated copies are immediately used for the transformative use and immediately discarded.

But this order need not decide this case on that rule. Anthropic did not use these copies only for training its LLM. Indeed, it retained pirated copies even after deciding it would not use them or copies from them for training its LLMs ever again. They were acquired and retained, as a central library of all the books in the world.

Building a central library of works to be available for any number of further uses was itself the use for which Anthropic acquired these copies. *One further use* was making further copies for training LLMs. But not every book Anthropic pirated was used to train LLMs. And, every pirated library copy was retained even if it was determined it would not be so used. Pirating copies to build a research library without paying for it, and to retain copies should they prove useful for one thing or another, was its own use — and not a transformative one (*see* Tr. 24–25, 35, 65; Opp. 4–10, 12 n.6; CC Br. Exh. 12 at -0144509 (“everything forever”)). *Napster*, 239 F.3d at 1015; *BMG Music v. Gonzalez*, 430 F.3d 888, 890 (7th Cir. 2005).

Anthropic’s briefing contains other reasons why it believes its pirated library copies are irrelevant to our fair use analysis, notwithstanding its own statements at our oral argument.

First, Anthropic accepts in this posture that it acted in bad faith but argues that its bad faith in pirating copies cannot “somehow short-circuit[]” the fair use analysis (Reply 6 (downplaying *Atari Games Corp. v. Nintendo of Am., Inc.*, 975 F.2d 832, 843 (Fed. Cir. 1992) (applying law of Ninth Circuit))). But its bad faith is not the basis for this decision. Each use of a work must be analyzed objectively. *Warhol*, 598 U.S. at 544–45. The objective analysis here shows the initial copies were pirated to create a central, general-purpose library, as a substitute for paid copies to do the same thing. (Of course, if infringement is found, bad faith would matter for determining willfulness. 17 U.S.C. § 504(c)(2).)

Second, Anthropic argues that its goal to put the copies eventually “to a highly transformative use” requires that each copy and use along the way be justified as having a

transformative use, too (Reply 14). But now Anthropic seeks to take the shortcut Anthropic just said cannot be taken. Again, the Supreme Court tasks us with looking past the “subjective intent of the user” to the objective *use* made of each copy. *Warhol*, 598 U.S. at 544–45 (emphasis added). Put another way, what a copyist *says or thinks or feels* matters only to the extent it shows what a copyist in fact *does* with the work. Indeed, the same copy can be used one way, then another, each with a different result. *Id.* at 533. Here, what Anthropic said about its acquisitions at the time — that they were made to “build[] a research library” while avoiding a “huge legal/practice/business slog” — are relevant in this regard. And, Anthropic’s actual use of these pirated copies was to create its central library of texts that, like any university or corporate library, stored the works’ well-organized facts, analyses, and expressive examples for various contingent uses, one being training.⁵

Third, Anthropic argues that *Texaco* — the case involving copies used in a central library, copies used in desk libraries, and copies used in the laboratory — is inapposite. Anthropic argues that the disputed copies in *Texaco* were never used in the laboratory but instead in personal desk libraries for a use “identical to the original purpose and use” of the central library copies, and so not for a transformative use (Reply 8 (summarizing 60 F.3d at 922–23)). By contrast, says Anthropic, here it *did use* copies in the laboratory to train LLMs — a very transformative use. But this is a fast glide over thin ice. Like *Texaco*, Anthropic possessed copies it did not put into use in the laboratory and it kept those copies in a central library even after its transformative use had been completed. But, unlike *Texaco*, which bought those copies, Anthropic never paid for the central library copies stolen off the

⁵ Our court of appeals has not yet reappraised how bad faith (or good faith) figures in fair use after *Warhol*. Its prior appraisal applied the Supreme Court’s statement that “[f]air use presupposes good faith and fair dealing,” *Harper & Row*, 471 U.S. at 562 (cleaned up). See *Perfect 10*, 508 F.3d at 1164 n.8. Since then, the Supreme Court has renewed its “skepticism about whether bad faith has any role.” *Oracle*, 593 U.S. at 32–33 (reiterating doubts of *Campbell*, 510 U.S. at 585 n.18). And, recently, the Supreme Court has held squarely that it is not the “subjective intent” of a copyist that counts, but the “objective . . . use” of the copy. *Warhol*, 598 U.S. at 544–45. This order applies this most recent analysis. *Miller v. Gammie*, 335 F.3d 889, 900 (9th Cir. 2003) (en banc).

1 internet. *Texaco* also shows why Anthropic is wrong to suppose that so long as you create an
2 exciting end product, every “back-end step, invisible to the public,” is excused (Br. 10).

3 Notably, this is not a case where source copies were unavailable for separate purchase or
4 loan. *See, e.g., NXIVM Corp. v. Ross Inst.*, 364 F.3d 471, 475–76, 478–79 (2d Cir. 2004)
5 (using selections of training manual — otherwise available only to cult’s trainees subject to
6 NDAs — to expose cult in critical review); *Time Inc. v. Bernard Geis Assocs.*, 293 F. Supp.
7 130, 135–36, 138, 146 (S.D.N.Y. 1968) (Judge Inzer Bass Wyatt) (making charcoal drawings
8 of photographs taken of originals otherwise not on sale or loan out to illustrate a history
9 book).⁶ Nor were the copies made only incidentally and necessarily from pirated copies. *See,*
10 *e.g., Perfect 10*, 508 F.3d at 1164 n.8 (copies of images that had been pirated by third-party
11 websites were used to index those same websites while indexing the entire web). Here, piracy
12 was the point: To build a central library that one could have paid for, just as Anthropic later
13 did, but without paying for it.

14 Nor were the initial copies made immediately transformed into a significantly altered
15 form. In *Perfect 10*, images were copied by the search engine in thumbnail form only and
16 deployed immediately into the transformative use of identifying the full-sized images and the
17 pages from which they came. 508 F.3d at 1160, 1165, 1167. And, in *Kelly v. Arriba Software*
18 *Corp.*, images were copied at full size and then into thumbnails for immediate use in building a
19 search engine, after which the full-sized copies were immediately deleted. 336 F.3d 811, 815
20 (9th Cir. 2003). Not here. The *full-text* copies of books were downloaded and maintained
21 “forever.”

22 Nor does the initial copying here even resemble the full-text copying in the *Google Books*
23 cases. There, libraries of authorized copies *already* had been assembled, and *all* copies

24
25 ⁶ Anthropic repeats the misleading characterization of the copyright holder in *Oracle* that the
26 initial copies were there purloined (Reply 5). Not so. “All agree[d] that Google was and
27 remain[ed] free to use the Java language itself. All agree[d] that Google’s virtual machine [wa]s
28 free of any copyright issues. All agree[d] that the six-thousand-plus method implementations by
Google [we]re free of copyright issues. The copyright issue, rather,” was the use of Java for
purposes of creating competing software having the same familiar, functional schema. *Oracle*
Am., Inc. v. Google Inc., 872 F. Supp. 2d 974, 978 (N.D. Cal. 2012), *aff’d and rev’d in part*, 750
F.3d 1339 (Fed. Cir. 2014).

therefrom were made for direct employment in a one-to-one further fair use — whether the transformative use of pointing to the works themselves, the use of providing the works in formats for print-disabled patrons, or the use of insuring against going out of print, getting lost, and becoming otherwise unavailable. *HathiTrust*, 755 F.3d at 97, 101, 103; *Google*, 804 F.3d at 206, 216–18, 228 (further distinguishing search and snippet uses, which “test[ed] the boundaries of fair use”). Not so here concerning the pirated copies. No authorized copies existed from which Anthropic made its first copies. No full-text copy therefrom was put immediately into use training LLMs. Not every copy was even necessary nor used for training LLMs. No initial copy was ever deleted, even if never used or no longer used.⁷ The university libraries and Google went to exceedingly great lengths to ensure that all copies were secured against unauthorized uses — both through technical measures and through legal agreements among all participants. Not so here. The library copies lacked internal controls limiting access and use.

Nor do the decisions on intermediate copying require anything less than the analysis applied here. Anthropic argues that our court of appeals in *Sega Enterprises Ltd. v. Accolade, Inc.* looked only at the “ultimate use” and “did not analyze a series of atomized acts of ‘infringement’ distinct from that overall purpose” (Reply 3). To the contrary, the appeals court examined the initial, intermediate, and ultimate copies used by the copyist. *The court explained that the copyist initially purchased commercially available copies of game cartridges and then made further copies necessarily and “solely in order to discover the functional requirements for compatibility.”* 977 F.2d 1510, 1522 (9th Cir. 1992). Thus, it reached only one result because on those facts there was only one “overall purpose” for the unauthorized copies. Indeed, the court reaffirmed prior caselaw holding that “intermediate

⁷ Training LLMs was not a use where perpetually maintaining a library copy was intrinsic to the proffered fair use (e.g., for a plagiarism-checker service). Nor is this an instance where retaining at least one copy was authorized by contract with the copyright owners (e.g., by agreement to express terms upon submission to a plagiarism-checker service, notwithstanding proposed terms scrawled on a paper prior to submission). *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 635–36 & n.5, 645 n.8 (4th Cir. 2009), *aff’d in relevant parts* 544 F. Supp. 2d 473, 480 (E.D. Va. 2008) (Judge Claude Hilton). Anthropic mischaracterizes this case.

1 copying of [a work] may infringe the exclusive rights granted to the copyright owner in
2 [S]ection 106 of the Copyright Act regardless of whether the end product of the copying also
3 infringes those rights.” *Id.* at 1518–19 (reaffirming *Walker v. Univ. Books*, 602 F.2d 859, 864
4 (9th Cir. 1979)).

5 Similarly, in *Sony Computer Entertainment, Inc. v. Connectix Corp.*, our appeals court
6 applied the same law to similarly focused conduct. Another copyist allegedly had purchased
7 an authorized copy and then made further copies solely and necessarily to reverse-engineer
8 compatibility requirements. 203 F.3d 596, 601, 602–03 (9th Cir. 2000).

9 Both *Sega* and *Sony* avoided imposing an “artificial hurdle” to fair use by generously
10 construing the intermediate copying necessary to the fair use. As one example, *Sega* stated
11 that an engineer should be permitted to reboot her computer while undertaking to reverse-
12 engineer software loaded onto it — even if doing so creates another digital copy of the
13 software and is not strictly necessary to reverse-engineering. *Id.* at 605. But neither *Sega* nor
14 *Sony* fathomed gifting an “artificial head start” to a fair user, either, by treating even the initial
15 copy as an intermediate one.

16 And, yes, some courts have “not inquire[d]” into intermediate or initial copying at all
17 (Reply 2 (citing *Campbell* as not inquiring into surplus copies in the studio)). But if a “close
18 reading of those cases [] reveals that in none of them was the legality of the [initial or]
19 intermediate copying at issue,” then it was not raised and not necessarily decided. *Sega*, 977
20 F.2d at 1519; see *Webster v. Fall*, 266 U.S. 507, 511 (1925). It was expressly decided
21 elsewhere: Our analysis must attend to different uses of different copies, and even to different
22 uses of *the same* copies. *Warhol*, 598 U.S. at 533.

23 Finally, Anthropic argues that even if the initial copies served a different use than the
24 intermediate and ultimate copies, it was not a use for which Anthropic necessarily would have
25 needed to pay Authors for a copy. In theory, argues Anthropic, it could have done as Google
26 did in *Google Books* — find an existing reference library willing to loan its copies for free as
27 source copies. Or, in theory, it could have done as Anthropic did later — go buy used copies
28 without having to pay Authors at all. See 17 U.S.C. § 109(a). But Anthropic did not do those

1 things — instead it stole the works for its central library by downloading them from pirated
2 libraries.

3 In sum, the first factor *points against* fair use for the central library copies made from
4 pirated sources — and no damages from pirating copies could be undone by later paying for
5 copies of the same works.

6 **2. THE NATURE OF THE COPYRIGHTED WORK.**

7 The second fair use factor is “the nature of the copyrighted work.” 17 U.S.C. § 107(2).
8 This factor “calls for recognition that some works are closer to the core of intended copyright
9 protection than others, with the consequence that fair use is more difficult to establish when the
10 former works are copied.” *Campbell*, 510 U.S. at 586. For one thing, less protection is due
11 published works than unpublished ones. For another, less protection is due “factual works than
12 works of fiction or fantasy.” *Harper & Row*, 471 U.S. at 563. But less protection is not no
13 protection. Even the arrangement of otherwise unprotectable facts surpasses the low bar for a
14 protectable original work of authorship. *Google*, 804 F.3d at 220.

15 Here, Anthropic accepts that all of Authors’ books — all published, whether non-fiction
16 or fiction — contained expressive elements (Reply 9). And, as set out above, this order
17 accepts Authors’ view of the evidence that their works were chosen for their expressive
18 qualities in building a central library and then in training specific LLMs (Opp. 11, 17 (citing,
19 e.g., Opp. Exh. 3 at -03433)).

20 The main function of the second factor is to help assess the other factors: to reveal
21 differences between the nature of the works at issue and the nature of their secondary use
22 (above), and to reveal any relation between the amount and substantiality of each work taken
23 and the secondary use (next). *E.g.*, *Campbell*, 510 U.S. at 586; *Kelly*, 336 F.3d at 820; *Google*,
24 804 F.3d at 220; *HathiTrust*, 755 F.3d at 98; *Bill Graham Archives v. Dorling Kindersley Ltd.*,
25 448 F.3d 605, 612–13 (2d Cir. 2006).

26 The second factor *points against* fair use for all copies alike.
27
28

1 **3. THE AMOUNT AND SUBSTANTIALITY OF THE PORTION USED.**

2 The third fair use factor is “the amount and substantiality of the portion” of the
3 copyrighted work used by the accused. 17 U.S.C. § 107(3). The crux of this factor is whether
4 the amount was “reasonable in relation to the purpose of the copying.” *Campbell*, 510 U.S. at
5 586. Thus, the amount of copying is considered first against the work itself, then more
6 importantly against the proposed transformative purpose. *See Warhol*, 598 U.S. at 543 & n.18.

7 **A. THE COPIES USED TO TRAIN SPECIFIC LLMs.**

8 Copies selected for inclusion in training sets were selected because they were complete
9 and because they contained rich protectible expression, or so this order accepts the record
10 shows for Authors. Was all this copying reasonably necessary to the transformative use?

11 Yes.

12 “What matters [] is not so much ‘the amount and substantiality of the portion used’ *in*
13 *making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a
14 public [in the purported secondary use] for which it may serve as a competing substitute [for
15 the primary use].” *Google*, 804 F.3d at 222. Here, once again, there is no allegation of any
16 traceable connection between the Claude service’s outputs and Authors’ works. The copying
17 used to train the LLMs underlying Claude was thus especially reasonable.

18 In response, Authors object primarily that the copying used in training was both
19 extremely extensive and not strictly necessary.

20 *As to extensive copying*, it is true that entire works were copied. And, “copying [] entire
21 work[s] ‘militate[s] against a finding of fair use.’” *Worldwide Church of God v. Philadelphia*
22 *Church of God, Inc.*, 227 F.3d 1110, 1118 (9th Cir. 2000) (quoting *Hustler Mag. Inc. v. Moral*
23 *Majority Inc.*, 796 F.2d 1148, 1155 (9th Cir. 1986)); *see Campbell*, 510 U.S. at 587. But we
24 just addressed why Authors’ argument is misdirected. The copies that count for this factor are
25 those that would merely serve the same use as the work’s ordinary one. Authors do not allege
26 such copying. The accused use here of the incremental copies is as orthogonal as can be
27 imagined to the ordinary use of a book.

1 *As to strict necessity*, Authors make a stronger point. When a productive use is made
2 possible only by borrowing from a specific work, fair use climbs towards its zenith. When a
3 productive use is possible without that borrowing, fair use falls to its nadir — and the
4 borrowing deserves a particularly compelling justification. *See Warhol*, 598 U.S. at 543 &
5 n.18, 547. Here, it is true that Anthropic could have used some other books or no books at all
6 for training its LLMs — or so this order accepts the record shows for Authors. But Anthropic
7 has presented a compelling explanation for why it was reasonably necessary to use them
8 anyway.

9 For one thing, all agree Anthropic needed billions of words to train any given LLM. If
10 using only books, Anthropic would have needed millions of books per model. If using a set
11 comprising only a small fraction of books and a larger fraction of other texts, Anthropic still
12 would have needed hundreds of thousands of books. Authors contend that because Anthropic
13 showed it could use such smaller sets of books, it surely could have used no books at all — or
14 at least not *their* books (Opp. 23). But Authors forget that “reasonably necessary” does not
15 mean “strictly necessary.” Authors do not contest that the volume of text required to train an
16 LLM is monumental. Because using so many works was reasonably necessary, using any one
17 work for *actually training* LLMs was about as reasonable as the next.

18 For another thing, no output to the public was even alleged to be infringing. So, yes,
19 Authors’ works were chosen as the strongest examples of writing. But the compelling benefits
20 of training the LLMs on strong examples were not offset by revelations to the public of any
21 portion of the works themselves. What was copied was therefore especially reasonable and
22 compelling.

23 The third factor thus *favors* fair use for the training copies.

24 ***B. THE COPIES USED TO BUILD A CENTRAL LIBRARY.***

25 But again, there was a separate use — a distinction that makes some difference as to
26 whether the amount and substantiality of the copying was “reasonable in relation to the
27 purpose of the copying” for the library copies. *Campbell*, 510 U.S. at 586.

(i) ***The Purchased Library Copies Converted from Print to Digital.***

For the print library copies that Anthropic purchased and then converted into digital library copies, Anthropic already enjoyed entitlement to keep the copies in its library. The purpose of the copying was to keep them in its library but with more favorable storage and searchability properties. Copying the entire work was exactly what this purpose required. There was no surplus copying. The source copy was destroyed.

The third fair use factor *favours* fair use for the purchased library copies converted from print to digital.

(ii) ***The Pirated Library Copies.***

For the pirated library copies, however, Anthropic lacked any entitlement to hold copies of the books at all. Its purpose, it says, was to train LLMs. But its objective conduct was to seek “all the books in the world” and then retain them even after deciding it would not make further copies from them for training — indicating there were other further uses. Against the purpose of acquiring all the books one could on the chance some might prove useful for training LLMs and maybe other stuff too, almost any unauthorized copying would have been too much. Anthropic copied millions of books *in toto*, Authors’ works among them.

The third factor *points against* fair use for the pirated library copies.

4. THE EFFECT OF THE USE UPON THE MARKET FOR OR VALUE OF THE COPYRIGHTED WORK.

The final factor is “the effect of the use upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107(4). This factor points against fair use when a copyist makes copies available that displace demand for copies the copyright owner already makes available or readily could. *Texaco*, 60 F.3d at 926–28 (reproduced copies); *Dr. Seuss Enters., L.P. v. ComicMix LLC*, 983 F.3d 443, 461 (9th Cir. 2020) (derivative copies). “While the first factor considers whether and to what extent an original work and secondary use [*in principle* could] have substitutable purposes, the fourth factor focuses on *actual or potential* market substitution.” *Warhol*, 598 U.S. at 536 n.12 (emphasis added).

1 **A. THE COPIES USED TO TRAIN SPECIFIC LLMs.**

2 The copies used to train specific LLMs did not and will not displace demand for copies
3 of Authors' works, or not in the way that counts under the Copyright Act.

4 Again, Authors concede that training LLMs did not result in any exact copies nor even
5 infringing knockoffs of their works being provided to the public. If that were not so, this
6 would be a different case. Authors remain free to bring that case in the future should such
7 facts develop.

8 Instead, Authors contend generically that training LLMs will result in an explosion of
9 works competing with their works — such as by creating alternative summaries of factual
10 events, alternative examples of compelling writing about fictional events, and so on. This
11 order assumes that is so (Opp. 22–23 (citing, e.g., Opp. Exh. 38)). But Authors' complaint is
12 no different than it would be if they complained that training schoolchildren to write well
13 would result in an explosion of competing works. This is not the kind of competitive or
14 creative displacement that concerns the Copyright Act. The Act seeks to advance original
15 works of authorship, not to protect authors against competition. *Sega*, 977 F.2d at 1523–24.

16 Authors next contend that training LLMs displaced (or will) an emerging market for
17 licensing their works for the narrow purpose of training LLMs (Opp. 21–22). Anthropic
18 argues that transactional costs would exceed Anthropic's expected benefit from any such
19 bargain, prompting it to cease dealing with any rightsholders or else to cease developing such
20 technology altogether (Br. 22–23). Our record could support either account — so this order
21 must assume Authors are correct. A market could develop (Opp. 19–21 (citing record)). Even
22 so, such a market for that use is not one the Copyright Act entitles Authors to exploit.

23 None of the cases cited by Authors requires a different result. All contemplated losses of
24 something the Copyright Act properly protected — not the kinds of fair uses for which a
25 copyright owner cannot rightly expect to control. *See TVEyes, Inc.*, 883 F.3d at 181 (use of a
26 right legally reserved to and factually already being licensed by copyright owner); *Texaco*, 60
27 F.3d 931 (same); *Ringgold v. BET, Inc.*, 126 F.3d 70, 80–81 (2d Cir. 1997) (use of a right
28 legally reserved to and factually likely to be marketable by copyright owner — displaying

images of her artistic work in television shows); *cf. Seltzer v. Green Day, Inc.*, 725 F.3d 1170, 1179 (9th Cir. 2013) (no evidence use could be or “was likely to” be marketable).

The fourth factor thus *favors* fair use for the training copies.

B. THE COPIES USED TO BUILD A CENTRAL LIBRARY.

(i) The Purchased Library Copies Converted from Print to Digital.

For these copies, this order assumes Anthropic’s format change from print to digital displaced purchases of new digital copies that Anthropic would have made directly from Authors (had it not been able to purchase print copies in used condition). But for reasons stated under the first factor, such losses did not relate to something the Copyright Act reserves for Authors to exploit. It was a format change.

Authors’ next argument, it seems, is that the format change nonetheless exposed it to usurpation of the opportunity to sell rightful copies because Anthropic *might* transmit additional unauthorized digital copies more readily than it could have transmitted additional unauthorized print copies — and that the same would be true for all format converters (*cf. Opp.* 25 n.14; *Opp. Expert Malackowski* ¶ 52). But after much discovery, there is no inkling in our record of intent to redistribute library copies once acquired nor of inability to secure that valuable library against outside actors. And, if the internal, central library copies did or do in fact lead to further reproduction or distribution, those further copies remain redressable separately by Authors. The format change did not itself usurp the Authors’ rightful entitlements.

This factor is thus *neutral* for the purchased library copies converted from print to digital.

(ii) The Pirated Library Copies.

The copies used to build a central library *and* that were obtained from pirated sources plainly displaced demand for Authors’ books — copy for copy. Not every person who merely intends to make a fair use of a work is thereby entitled to a full copy in the meantime, nor even to steal a copy so that achieving this fair use is especially simple or cost-effective. Here, the copies employed in training LLMs were one thing, but the copies acquired to assemble a

convenient, general-purpose library of works for various uses for which the company might have of them, if any, was a different use altogether.

Anthropic has almost no rebuttal on these points. *First*, Anthropic argues that “Claude’s services do not reduce [or usurp] the value of Plaintiffs’ works through substitution in their traditional markets” (*see* Br. Expert Peterson ¶ 33). But stealing pirated copies of Authors’ works plainly did. *Second*, Anthropic argues that it may have been able to purchase some books on the open market (and some other texts), but not other texts it copied (*cf. id.* ¶ 48 (re licensing)). But this case does not concern those other texts it could not have purchased. It could have purchased Authors’ books (and many others). In fact it later did. *Finally*, Anthropic argues that the effect on these texts from one book foregone was too small to be considered (*see id.* ¶ 77). But the test requires that we contemplate the likely result were the conduct to be condoned as a fair use — namely to steal a work you could otherwise buy (a book, millions of books) so long as you at least loosely intend to make further copies for a purportedly transformative use (writing a book review with excerpts, training LLMs, etc.), without any accountability. As Anthropic itself suggested, “That would destroy the [entire] publishing market if that were the case” (*see* Tr. 53; *see also* Tr. 32, 41; Opp. Expert Malackowski ¶¶ 31–34, 38).

The fourth factor *points against* fair use for the pirated library copies.

5. OVERALL ANALYSIS.

After the four factors and any others deemed relevant are “explored, [] the results [are] weighed together, in light of the purposes of copyright.” *Campbell*, 510 U.S. at 578.

The copies used to train specific LLMs were justified as a fair use. Every factor but the nature of the copyrighted work favors this result. The technology at issue was among the most transformative many of us will see in our lifetimes.

The copies used to convert purchased print library copies into digital library copies were justified, too, though for a different fair use. The first factor strongly favors this result, and the third favors it, too. The fourth is neutral. Only the second slightly disfavors it. On balance, as

the purchased print copy was destroyed and its digital replacement not redistributed, this was a fair use.

The downloaded pirated copies used to build a central library were not justified by a fair use. Every factor points against fair use. Anthropic employees said copies of works (pirated ones, too) would be retained “forever” for “general purpose” even after Anthropic determined they would never be used for training LLMs. A separate justification was required for each use. None is even offered here except for Anthropic’s pocketbook and convenience.

And, as for any copies made from central library copies but not used for training, this order does not grant summary judgment for Anthropic. On this record in this posture, the central library copies were retained even when no longer serving as sources for training copies, “hundreds of engineers” could access them to make copies for other uses, and engineers did make other copies. Anthropic has dodged discovery on these points (*e.g.*, Opp. Exh. 17 at 93–94 (retained); Opp. Exh. 22 at 196 (no limits); Opp. Exh. 30 at 3, 4 (no accounting); *see also* Opp. 15). We cannot determine the right answer concerning such copies because the record is too poorly developed as to them. Anthropic is not entitled to an order blessing all copying “that Anthropic has ever made after obtaining the data,” to use its words (Opp. Exh. 30 at 3, 4).

CONCLUSION

With respect to the training copies and the print-to-digital converted copies, this order has drawn all ambiguities and inferences in favor of the opposing side, namely Authors. With respect to the pirated copies, this order has also accepted the Authors’ version of the facts. Authors did not move for summary judgment but if they had, then we would have been obligated to accept all reasonable views given the evidence in defendant’s favor instead.

This order grants summary judgment for Anthropic that the training use was a fair use. And, it grants that the print-to-digital format change was a fair use for a different reason. But it denies summary judgment for Anthropic that the pirated library copies must be treated as training copies.

We will have a trial on the pirated copies used to create Anthropic’s central library and the resulting damages, actual or statutory (including for willfulness). That Anthropic later

bought a copy of a book it earlier stole off the internet will not absolve it of liability for the theft but it may affect the extent of statutory damages. Nothing is foreclosed as to any other copies flowing from library copies for uses other than for training LLMs.

IT IS SO ORDERED.

Dated: June 23, 2025.

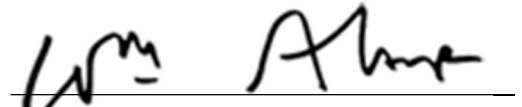

WILLIAM ALSUP
UNITED STATES DISTRICT JUDGE

EXHIBIT C

No. _____

**IN THE UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT**

**ANDREA BARTZ, CHARLES GRAEBER,
and KIRK WALLACE JOHNSON, Plaintiffs,**

v.

ANTHROPIC PBC, Defendant.

On Petition for Permission to Appeal from the United States District Court
for the Northern District of California,
Case No. 3:24-cv-05417-WHA
The Honorable William H. Alsup

**DECLARATION OF KRISHNA RAO IN SUPPORT
OF DEFENDANT'S PETITION FOR PERMISSION
TO APPEAL UNDER RULE 23(F)**

Kathleen R. Hartnett
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
Telephone: (415) 693-2000
(khartnett@cooley.com)

Ephraim A. McDowell
Alexander J. Kasner
Elias S. Kim
COOLEY LLP
1299 Pennsylvania Avenue, N.W.
Washington, DC 20004-2400
Telephone: (202) 842-7800

Daralyn J. Durie
MORRISON & FOERSTER LLP
425 Market Street
San Francisco, CA 94105-2482
Telephone: (415) 268-7000

Douglas A. Winthrop
ARNOLD & PORTER KAYE
SCHOLER LLP
Three Embarcadero Center, 10th Floor
San Francisco, CA 94111-4024
Telephone: (415) 471-3100

Allison Wettstein O'Neill
COOLEY LLP
10265 Science Center Drive
San Diego, CA 92121-1117
Telephone: (858) 550-6000

Attorneys for Petitioner Anthropic PBC

I, Krishna Rao, declare as follows:

1. I am Chief Financial Officer for Defendant Anthropic PBC (“Anthropic”). I submit this Declaration in support of Defendant’s Petition for Permission to Appeal Under Rule 23(f). I have personal knowledge of the matters set forth herein, and if called upon as a witness, I could competently testify thereto.

2. Since its founding in 2021, Anthropic has raised to date a total of approximately \$15 billion in equity and convertible securities.

3. Anthropic currently expects to generate no more than \$5 billion in revenue in 2025, however, we expect to operate at a loss of billions of dollars for 2025.

4. Due to its investments in research, development, and infrastructure, and the extremely capital-intensive nature of developing frontier large language models (LLMs), Anthropic is not yet profitable as a business (and in fact is significantly loss making) and does not expect to be profitable in the next few years.

5. I understand, given the district court’s class certification order and the range of statutory damages, the potential total damages award at issue. Based on Anthropic’s initial assessment and inquiries, it appears very unlikely that Anthropic could obtain an appeal bond for the total possible amount of statutory damages should Anthropic be found liable.

6. As of 2025, Anthropic continues to raise capital to fund its significant

investments in research and development. Anthropic remains a private company.

I declare under penalty of perjury that the foregoing is true and correct.

Executed on July 31, 2025, at San Francisco, California.



KRISHNA RAO

**UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT**

Form 15. Certificate of Service for Electronic Filing

Instructions for this form: <http://www.ca9.uscourts.gov/forms/form15instructions.pdf>

9th Cir. Case Number(s)

I hereby certify that I electronically filed the foregoing/attached document(s) on this date with the Clerk of the Court for the United States Court of Appeals for the Ninth Circuit using the Appellate Electronic Filing system.

Service on Case Participants Who Are Registered for Electronic Filing:

I certify that I served the foregoing/attached document(s) via email to all registered case participants on this date because it is a sealed filing or is
☐ submitted as an original petition or other original proceeding and therefore cannot be served via the Appellate Electronic Filing system.

Service on Case Participants Who Are NOT Registered for Electronic Filing:

I certify that I served the foregoing/attached document(s) on this date by hand delivery, mail, third party commercial carrier for delivery within 3 calendar days, or, having obtained prior consent, by email to the following unregistered case participants (*list each name and mailing/email address*):
☒

See attached Service List.

Description of Document(s) (*required for all documents*):

PETITION FOR PERMISSION TO APPEAL UNDER RULE 23(F)

Signature

/s/ Alexander J. Kasner

Date

July 31, 2025

Feedback or questions about this form? Email us at forms@ca9.uscourts.gov

SERVICE LIST

Justin A. Nelson
Alejandra C. Salinas
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002-5096
Telephone: (713) 651-9366
jnelson@susmangodfrey.com
asalinas@susmangodfrey.com

Rohit D. Nath (SBN 316062)
Michael Gervais (SBN 330731)
SUSMAN GODFREY L.L.P.
1900 Avenue of the Stars, Suite 1400
Los Angeles, CA 90067-2906
Telephone: (310) 789-3100
rnath@susmangodfrey.com
mgervais@susmangodfrey.com

Jordan W. Connors
SUSMAN GODFREY L.L.P.
401 Union Street, Suite 3000
Seattle, WA 98101
Telephone: (206) 516-3880
jconnors@susmangodfrey.com

J. Craig Smyser
Samir Doshi
SUSMAN GODFREY L.L.P.
One Manhattan West, 51st Floor,
New York, NY 10019
Telephone: (212) 336-8330
csmyser@susmangodfrey.com
sdoshi@susmangodfrey.com

Co-Lead Class Counsel

Rachel Geman
Jacob S. Miller
Danna Z. Elmasry
**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**
250 Hudson Street, 8th Floor
New York, NY 10013-1413
Telephone: (212) 355-9500
rgeman@lchb.com
jmiller@lchb.com
delmasry@lchb.com

Daniel M. Hutchinson (SBN 239458)
Jallé Dafa (SBN 290637)
Reilly T. Stoler (SBN 310761)
**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339
Telephone: (415) 956-1000
rstoler@lchb.com
jdafa@lchb.com
dhutchinson@lchb.com

Betsy A. Sugar
**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**
222 Second Ave., #1640
Nashville, TN 37201
Telephone: (615) 313-9000
bsugar@lchb.com

Co-Lead Class Counsel