

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

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

*Attorneys for Defendant ANTHROPIC PBC*

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

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

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

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

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

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

Action Filed: August 19, 2024

**ANTHROPIC PBC'S REPLY IN SUPPORT  
 OF MOTION FOR SUMMARY  
 JUDGMENT**

Date: May 15, 2025

Time: 8:00 a.m.

Place: Courtroom 12, 19th Floor

Judge: Honorable William H. Alsup

**TABLE OF CONTENTS****Page**

INTRODUCTION AND SUMMARY OF ARGUMENT .....	1
ARGUMENT.....	1
I. THE ANSWERS TO THE COURT’S QUESTIONS SUPPORT FAIR USE .....	1
A. Case 1. An e-book is purchased at full price and read over five days by the purchaser. Each day he reloads the entire e-book. Case 2. Same as Case 1, but he buys the e-book for one cent from a notorious counterfeiter known as Books ‘R’ Cheap.....	1
B. Case 3. Same as Case 2, but his purpose throughout is to write a transformative parody of the work (and he does).....	1
C. Case 4. Same as Case 3, but his purpose further includes to write the parody with cowriters, so he buys one e-book and copies it for all co-writers (they jointly write the parody). .....	2
D. Case 5. Someone buys a copy of the e-book at full price and then gives it to a writer who uses it to write a transformative parody of the work. ....	3
E. Variant. For any case relevant (or for all at once), briefly explain to what extent it makes any difference for the copyright laws whether the work is purchased as an e-book or instead in print and converted by the purchaser immediately into a digital format before the next step. ....	3
II. THE FIRST FACTOR FAVORS FAIR USE BECAUSE USING COPYRIGHTED WORKS TO TRAIN AN LLM IS QUINTESSENTIALLY TRANSFORMATIVE.....	3
A. No case assesses transformative purpose by looking at isolated actions that are part of an overall transformative process. ....	3
B. There is no material dispute of fact that Anthropic’s use of books to train LLMs is highly transformative. ....	7
1. Plaintiffs do not seriously contest how LLM training works.....	7
2. Plaintiffs’ non-transformative use citations are inapposite. ....	8
III. THE SECOND AND THIRD FACTORS FAVOR FAIR USE .....	9
IV. THE FOURTH FACTOR FAVORS FAIR USE BECAUSE ANTHROPIC’S HIGHLY TRANSFORMATIVE USE DOES NOT HARM THE MARKET FOR PLAINTIFFS’ WORKS .....	9
A. As a matter of law, Anthropic’s highly transformative use forecloses Plaintiffs’ theory of market harm based on lost licensing fees. ....	10
B. There is no genuine dispute about whether a viable licensing market for LLM training data could develop. ....	11



D. Under Plaintiffs’ “initial acquisition” theory, there is no cognizable harm from lost book sales .....14

E. The public benefits of Claude support a fair use finding and do not present issues of fact. ....14

CONCLUSION.....15

**TABLE OF AUTHORITIES****Page(s)****Cases**

<i>A&amp;M Recs., Inc. v. Napster, Inc.</i> , 239 F.3d 1004 (9th Cir. 2001).....	6, 14
<i>A.V. ex rel. Vanderhye v. iParadigms, LLC</i> , 562 F.3d 630 (4th Cir. 2009).....	5, 6, 8
<i>A.V. v. iParadigms Liab. Co.</i> , 544 F. Supp. 2d 473 (E.D. Va. 2008).....	5
<i>Am. Geophysical Union v. Texaco Inc.</i> , 60 F.3d 913 (2d Cir. 1994).....	8, 12
<i>Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023) .....	2, 3
<i>Atari Games Corp. v. Nintendo of Am. Inc.</i> , 975 F.2d 832 (Fed. Cir. 1992).....	6
<i>Authors Guild, Inc. v. HathiTrust</i> , 755 F.3d 87 (2nd Cir. 2014).....	<i>passim</i>
<i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015).....	4, 5, 9
<i>Bill Graham Archives v. Dorling Kindersley Ltd.</i> , 448 F.3d 605 (2d Cir. 2006).....	10
<i>BMG Music v. Gonzalez</i> , 430 F.3d 888 (7th Cir. 2005).....	6
<i>Cambridge Univ. Press v. Patton</i> , 769 F.3d 1232 (11th Cir. 2014).....	8
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994) .....	2, 6, 10
<i>Dr. Seuss Enters. L.P. v. ComicMix LLC</i> , 983 F.3d 443 (9th Cir. 2020).....	11
<i>Google LLC v. Oracle Am., Inc.</i> , 593 U.S. 1 (2021) .....	<i>passim</i>
<i>Hachette Book Grp., Inc. v. Internet Archive</i> , 115 F.4th 163 (2d Cir. 2024).....	15

1	<i>Harper &amp; Row v. Nation Enters.</i> ,	
2	471 U.S. 539 (1985) .....	5, 6
3	<i>In re Aimster Copyright Litig.</i> ,	
4	334 F.3d 643 (7th Cir. 2003) .....	6
5	<i>In re Barboza</i> ,	
6	545 F.3d 702 (9th Cir. 2008) .....	4
7	<i>John-Charles v. California</i> ,	
8	646 F.3d 1243 (9th Cir. 2011) .....	13
9	<i>Kelly v. Arriba Soft Corp.</i> ,	
10	336 F.3d 811 (9th Cir. 2003) .....	3, 4, 5
11	<i>Mattel, Inc. v. Walking Mountain Prods.</i> ,	
12	353 F.3d 792 (9th Cir. 2003) .....	8
13	<i>MCA, Inc. v. Wilson</i> ,	
14	677 F.2d 180 (2d Cir. 1981) .....	15
15	<i>NXIVM Corp. v. Ross Inst.</i> ,	
16	364 F.3d 471 (2d Cir. 2004) .....	6
17	<i>Oracle Am., Inc. v. Google Inc.</i> ,	
18	2016 WL 3181206 (N.D. Cal. June 8, 2016) .....	5
19	<i>Perfect 10, Inc. v. Amazon.com, Inc.</i> ,	
20	508 F.3d 1146 (9th Cir. 2007) .....	<i>passim</i>
21	<i>Sega Enters. Ltd. v. Accolade, Inc.</i> ,	
22	977 F.2d 1510 (9th Cir. 1992) .....	3, 9, 13
23	<i>Sony Computer Ent., Inc. v. Connectix Corp.</i> ,	
24	203 F.3d 596 (9th Cir. 2000) .....	3, 9
25	<i>Sony Corp. of Am. v. Universal City Studios, Inc.</i> ,	
26	464 U.S. 417 (1984) .....	8
27	<i>Thomson Reuters Enter. Ctr. GMBH v. Ross Intel. Inc.</i> ,	
28	2025 WL 458520 (D. Del. Feb. 11, 2025) .....	9, 15
	<i>Time Inc. v. Bernard Geis Assocs.</i> ,	
	293 F. Supp. 130 (S.D.N.Y. 1968) .....	5
	<i>Tresóna Multimedia, LLC v. Burbank High Sch. Vocal Music Ass'n</i> ,	
	953 F.3d 638 (9th Cir. 2020) .....	10
	<i>Triller Fight Club II LLC v. H3 Podcast</i> ,	
	2023 WL 11877604 (C.D. Cal. Sept. 15, 2023) .....	6

1	<i>Triton Energy Corp. v. Square D Co.</i> ,	
2	68 F.3d 1216 (9th Cir. 1995).....	12
3	<i>UMG Recordings, Inc. v. MP3.Com, Inc.</i> ,	
4	92 F. Supp. 2d 349 (S.D.N.Y. 2000).....	6
5	<i>United States v. Slater</i> ,	
6	348 F.3d 666 (7th Cir. 2003).....	6
7	<i>Viacom Int’l, Inc. v. YouTube, Inc.</i> ,	
8	676 F.3d 19 (2d Cir. 2012).....	13
9	<i>Wall Data Inc. v. Los Angeles Cnty. Sheriff’s Dep’t</i> ,	
10	447 F.3d 769 (9th Cir. 2006).....	9
11	<i>White v. W. Pub. Corp.</i> ,	
12	29 F. Supp. 3d 396 (S.D.N.Y. 2014).....	11

### **Statutes**

13	17 U.S.C. § 115.....	12
14	17 U.S.C. § 512(c) .....	13
15	17 U.S.C. § 512(h).....	6

### **Other Authorities**

16	U.S. Const., Art. I, Sec. 8.....	15
17	Pierre Leval, Toward a Fair Use Standard, 103 Harv. L. Rev. 1105, 1126 (1990) .....	2
18	Wendy J. Gordon, Fair Use As Market Failure, 82 Colum. L. Rev. 1600, 1614-15 (1982) .....	12

## INTRODUCTION AND SUMMARY OF ARGUMENT

Today's large language models achieve capabilities that no prior computer system could approach. Unlike traditional software with prescribed functions, Claude performs complex tasks—from software coding to document and data analysis—with a level of generality and adaptability previously confined to human intelligence. Plaintiffs do not seriously dispute that Claude is a radically transformative technology within the meaning of the first fair use factor, a technology that enables creativity and fosters scientific progress—exactly the type of advancement that copyright law was constitutionally designed to promote. And the parties' agreement on the facts goes further: Plaintiffs do not even contend that Claude generates outputs that are substantially similar to any of Plaintiffs' books, and they concede that Claude is not a substitute in the market for their books under the fourth fair use factor. With no genuine disputes of material fact, Anthropic is entitled to judgment as a matter of law on its fair use defense.

Because Plaintiffs do not raise genuine disputes of fact, they seek to defeat Anthropic's summary judgment motion by asking this Court to depart from established fair use law in at least two ways. *First*, they ask the Court to upend decades of fair use doctrine by divorcing the "initial acquisition" of a copyrighted work from the purpose for which the copy was acquired. *Second*, they assert that even extremely transformative uses must be subject to license, so long as the plaintiff can hypothesize a speculative future when a licensing market could develop for that use. The Court should reject these invitations to rewrite the law of fair use and grant Anthropic's motion.

## ARGUMENT

### I. THE ANSWERS TO THE COURT'S QUESTIONS SUPPORT FAIR USE

- A. **Case 1. An e-book is purchased at full price and read over five days by the purchaser. Each day he reloads the entire e-book. Case 2. Same as Case 1 but he buys the e-book for one cent from a notorious counterfeiter known as Books 'R' Cheap.**

Anthropic agrees with Plaintiffs that, on their assumptions, in Case 1, the purchaser does not infringe, and in Case 2, he does.

- B. **Case 3. Same as Case 2 but his purpose throughout is to write a transformative parody of the work (and he does).**

This is fair use. Parody is a quintessential fair use. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S.

569, 579 (1994). Courts have consistently held that unauthorized copies made as part of the process of creating new, non-infringing, public-facing content are fair use. *E.g., Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165-66 (9th Cir. 2007) (finding fair use despite Google’s making unauthorized copies of images from “infringing” sites). No precedent counsels a different result here, so the parodist’s unauthorized back-end copies are also fair use. “If the use is otherwise fair, then no permission need be sought or granted.” *Campbell*, 510 U.S. at 585 n.18. That is why the Supreme Court has rejected the idea that “bad faith has any role in a fair use analysis,” concluding that its previous skepticism about this was “justifiable” because “[c]opyright is not a privilege reserved for the well-behaved.” *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 32 (2021) (quoting Pierre Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105, 1126 (1990)). By contrast, Books ‘R’ Cheap commits copyright infringement by enabling and profiting from unauthorized distribution of e-books for its purchasers’ consumptive purposes.

**C. Case 4. Same as Case 3 but his purpose further includes to write the parody with cowriters, so he buys one e-book and copies it for all co-writers (they jointly write the parody).**

This is also fair use. Again, the means by which each parodist obtains a copy of a copyrighted work does not affect their fair use—what matters is the purpose for which that copy was made. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 523 (2023). That purpose is the same whether the work is made by a single author or a group. For example, in *Campbell*, the Supreme Court did not inquire as to whether each member of 2 Live Crew bought their own copy of “Pretty Woman” or how many different mixes were made in the studio, and it would make no sense if the fair use inquiry turned on those questions. 510 U.S. at 569. At most, some courts have considered, under the third fair use factor, whether the “copies are excessive or unreasonable in relation to the purposes identified.” *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 99 (2nd Cir. 2014). And *HathiTrust* exemplifies why Case 4 is a fair use. There, the court found fair use when the defendant created a repository of digital copies obtained from an unauthorized source (Google), which was stored at four separate locations operated by different, unrelated entities. The copies were not excessive because they were “reasonably necessary” to facilitate “legitimate uses.” *Id.* at 98. Likewise, here, so long as the number of copies is reasonable in relation to the group’s transformative purpose, they are protected as fair use.



**D. Case 5. Someone buys a copy of the e-book at full price and then gives it to a writer who uses it to write a transformative parody of the work.**

This is likewise fair use. It is irrelevant that the parodist did not himself purchase a copy of the work. What matters is the purpose of the copy.

**E. Variant. For any case relevant (or for all at once), briefly explain to what extent it makes any difference for the copyright laws whether the work is purchased as an e-book or instead in print and converted by the purchaser immediately into a digital format before the next step.**

Conversion of a print work into a digital format does not affect the analyses above. Case 1 remains fair use. Case 2 is still infringement. For Cases 3, 4, and 5, print or digital copies are fair use for the same reason—the transformative nature of the use to which the copies are put. Nothing in fair use doctrine suggests the format of the original work should have any effect on the outcome of the fair use analysis.

## **II. THE FIRST FACTOR FAVORS FAIR USE BECAUSE USING COPYRIGHTED WORKS TO TRAIN AN LLM IS QUINTESENTIALLY TRANSFORMATIVE**

**A. No case assesses transformative purpose by looking at isolated actions that are part of an overall transformative process.**

Supreme Court and Ninth Circuit precedents foreclose Plaintiffs’ central argument that Anthropic’s purpose in downloading copies of books should be evaluated in isolation, rather than as an integral step in an overall transformative purpose. *Warhol*, 598 U.S. at 533 (“The same copying may be fair when used for one purpose but not another.”). Virtually every digital fair use case involves the creation of initial or intermediate copies before the final, transformative use. There are decades of cases that do not apply the analytical separation Plaintiffs propose, which would have come out differently if they had:

- In *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1522 (9th Cir. 1992), the Ninth Circuit found fair use of source code copies obtained by reverse-engineering after assessing their “ultimate use”—to create non-infringing but competing video games—and did not analyze a series of atomized acts of “infringement” distinct from that overall purpose.
- Similarly, in *Sony Computer Ent., Inc. v. Connectix Corp.*, 203 F.3d 596, 602 (9th Cir. 2000), the Ninth Circuit addressed another software reverse engineering case in which engineers created new copies “each time” that they “booted up their computer and the computer copied the program into the RAM” by looking to the overall fair use purpose.
- In *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 815 (9th Cir. 2003), the Ninth Circuit found fair use

when Arriba “obtained [a] database of pictures by copying images from other web sites,” which were then reduced to thumbnails and then displayed in a search engine by looking to Arriba’s overall transformative purpose.

- Similarly, in *Perfect 10*, 508 F.3d at 1165, 1168, the Ninth Circuit again found fair use after Google obtained unauthorized copies of pictures from infringing websites by looking at Google’s overall transformative use of building a search index.
- In *Authors Guild v. Google, Inc.*, 804 F.3d 202, 208 (2d Cir. 2015) (“*Google Books*”), the Second Circuit assessed the utility of the searchable database Google created, not the individual acts of making millions of unauthorized copies of books required to create that database.
- The same is true in *HathiTrust*, 755 F.3d at 103, in which HathiTrust received millions of unauthorized full-text, digitized copies of books from Google and put them to the transformative use of making them text-searchable and accessible to the visually impaired.

These seminal cases uniformly hold that what matters for fair use law is Anthropic’s ultimate purpose in copying, not any artificially siloed individual step in isolation.<sup>1</sup>

Relatedly, Plaintiffs argue separating the initial acquisition step from the broader purpose of Anthropic’s use is required because some of the copies made came from sources on the internet that were themselves unauthorized. Opp. (ECF No. 170-3) at 6-7. But whether the copy a defendant made came from an authorized or unauthorized copy does not affect the fair use analysis. Plaintiffs’ imagined rule is incompatible with decades of precedent, including the cases that considered the legality of search engines—all of which make millions of copies of texts and images from sources that are themselves infringing. See, e.g., *Perfect 10*, 508 F.3d at 1164 n.8; *Arriba*, 336 F.3d at 818-22.

More broadly, Plaintiffs contend that Anthropic has engaged in “piracy” or “theft” by

---

<sup>1</sup>Plaintiffs cite a footnote in *In re Barboza*, 545 F.3d 702, 704-05 & n.1 (9th Cir. 2008), for the proposition that each of Anthropic’s acts of copying can be separated as “distinct” acts for fair use purposes. Opp. at 6. But that court did not address that issue at all. In *Barboza*, the defendant had engaged in two distinct activities: creation of one set of “unlawfully duplicated” videos and distribution of a separate set of “legally purchased” videos. 545 F.3d at 706. The Ninth Circuit merely pointed out that only the duplicated videos were the subject of the trial because the district court found that plaintiffs had not rebutted application of the first sale doctrine to the purchased videos. Moreover, *Barboza* is irrelevant here because *it is not a fair use case*. Plaintiffs do not cite a single fair use case where closely related acts of copying were analyzed separately.

1 downloading books as part of the process of creating its transformative technology. This is a familiar  
2 refrain. It is exactly what the plaintiff said in *Oracle*. See, e.g., 2020 WL 832871, at \*1 (Resp’t Oracle’s  
3 Supreme Ct. Br. in *Oracle*, 593 U.S. 1, accusing Google of an “egregious act of plagiarism” and referring  
4 to Google’s acts as “piracy” and “pirating”). As this Court noted in ruling on Oracle’s post-trial motion  
5 for judgment as a matter of law, Oracle had argued to the Federal Circuit that Google “exploited a  
6 **purloined** work.” *Oracle Am., Inc. v. Google Inc.*, 2016 WL 3181206, at \*2 n.1 (N.D. Cal. June 8, 2016)  
7 (emphasis added). These labels made no difference. See also *Perfect 10*, 508 F.3d at 1157, 1165-66 (fair  
8 use when Google provided thumbnail images from “website publishers that republish[ed] Perfect 10’s  
9 images on the Internet **without authorization**” and Perfect 10 notified Google they were unauthorized)  
10 (emphasis added); *Arriba*, 336 F.3d at 816-22 (fair use even though copyright holder “had **never given**  
11 **permission** to [defendant] to copy his images and objected when he found out that [defendant] was using  
12 them”) (emphasis added); *Google Books*, 804 F.3d at 208 (fair use for “**unauthorized digitizing** of  
13 copyright-protected works” used for searchable database with snippet display) (emphasis added); *Time*  
14 *Inc. v. Bernard Geis Assocs.*, 293 F. Supp. 130, 136 (S.D.N.Y. 1968) (fair use despite the defendant’s  
15 “recognition of the **impropriety**” of his making unauthorized copies of film frames kept in a private  
16 office, as cited by Supreme Court in *Harper & Row v. Nation Enters.*, 471 U.S. 539 (1985)).

17 Attempting to distinguish it from all of the cases rejecting their position, Plaintiffs  
18 mischaracterize *Google Books* by claiming that Google “lawfully obtained” the books via “bi-lateral  
19 agreements” with libraries. Opp. at 9. But that description is not true. Those libraries lacked rightsholder  
20 authorization to enter those deals with Google, no rightsholders were paid, and the rightsholders sued  
21 when they learned of these “agreements.” *Google Books*, 804 F.3d at 208.<sup>2</sup> Plaintiffs also ignore that, in  
22 *HathiTrust*, the Trust received unauthorized digital copies from Google, which was separately being sued  
23 for infringement for making those copies. 755 F.3d at 90.<sup>3</sup>

24  
25 <sup>2</sup> Contrary to Plaintiffs’ assertion (Opp. at 18), the fact that Google’s use of plaintiffs’ books might  
26 promote sales of plaintiffs’ books was not “critical” to the court’s ruling finding fair use. To the  
27 contrary, the court noted that the snippet function might cause a **loss** of sales. 804 F.3d at 224.

28 <sup>3</sup> Plaintiffs are wrong that *A.V. ex rel. Vanderhuy v. iParadigms, LLC*, 562 F.3d 630, 635 (4th Cir. 2009)  
 (“*iParadigms II*”), supports their position because “the defendant lawfully obtained its *initial* copy.”  
 Opp. at 9 (emphasis in original). In fact, the plaintiffs there expressly prohibited the defendant from  
 archiving their works at the time the works were initially handed in, and the defendant did so anyway.  
 *A.V. v. iParadigms Liab. Co.*, 544 F. Supp. 2d 473, 478 (E.D. Va. 2008). Those same plaintiffs later

All of the preceding cases are incompatible with the suggestion that “bad faith” somehow short-circuits the fair use analysis. Accordingly, the Court should reject Plaintiffs’ invitation to rely upon *Atari Games Corp. v. Nintendo of Am. Inc.*, 975 F.2d 832, 843 (Fed. Cir. 1992), which applies a *sui generis*, brightline rule barring fair use for unauthorized initial copies. *First*, this outlier decision is inconsistent with the Supreme Court’s “skepticism” that any “bad faith” is relevant to fair use at all. *Oracle*, 593 U.S. at 31-32; *Campbell*, 510 U.S. at 585 n.18. *Second*, in the decades since *Atari*, neither the Ninth Circuit nor any other circuit has adopted its rule that fair use protection is unavailable to defendants that obtain unauthorized copies. *See Triller Fight Club II LLC v. H3 Podcast*, 2023 WL 11877604 at \*8 (C.D. Cal. Sept. 15, 2023) (“[T]he Ninth Circuit has not required a defendant to have an authorized copy to invoke fair use.”) (citing *Perfect 10*, 508 F.3d at 1164 n.8). In fact, the Second Circuit has expressed its belief that *Atari* misread *Harper & Row*. *See NXIVM Corp. v. Ross Inst.*, 364 F.3d 471, 478-79 (2d Cir. 2004).

Even in the cases Plaintiffs characterize as “Internet piracy cases,” courts have conducted a robust fair use analysis of the defendants’ *ultimate purpose* regardless of their means of initial acquisition, and some courts have expressly acknowledged that even pirated copies could be put to fair use purposes. *See A&M Recs., Inc. v. Napster, Inc.*, 239 F.3d 1004, 1014 (9th Cir. 2001) (full fair use analysis despite allegation of acquisition by “piracy”); *United States v. Slater*, 348 F.3d 666, 669 (7th Cir. 2003) (unauthorized distribution of copyrighted software deemed not a fair use by analyzing relevant fair use factors); *UMG Recordings, Inc. v. MP3.Com, Inc.*, 92 F. Supp. 2d 349, 352 (S.D.N.Y. 2000) (distribution of unauthorized copies of songs found not a fair use only after analyzing all four fair use factors); *BMG Music v. Gonzalez*, 430 F.3d 888, 890 (7th Cir. 2005) (use of downloaded MP3s not a form of “time-shifting,” which would have been a fair use). *See also In re Aimster Copyright Litig.*, 334 F.3d 643, 647, 652-53 (7th Cir. 2003) (noting that using Aimster to download a copy of a CD one already owns “might be a fair use”).<sup>4</sup> The Court should likewise assess fair use here by reference to Anthropic’s ultimate purpose, regardless of how it acquired Plaintiffs’ works.

sued iParadigms, and the district court and the court of appeals held that the defendant’s use was fair. *iParadigms II*, 562 F.3d at 634, 645.

<sup>4</sup> Plaintiffs grossly misquote *In re DMCA § 512(h) Subpoena to Twitter, Inc.*, 608 F. Supp. 3d 868 (N.D. Cal. 2022) (Opp. at 7), in which the court actually said, “[i]n some cases, no analysis is required; it is obvious, for example, that downloading *and distributing* copyrighted music via peer-to-peer systems does not constitute fair use.” 608 F. Supp. 3d at 879 (emphasis added).

**B. There is no material dispute of fact that Anthropic’s use of books to train LLMs is highly transformative.**

The training process transforms books from their original expressive form into abstract statistical patterns and mathematical weights that represent probabilistic relationships between words and concepts among millions of works, fundamentally altering both the form and function of the original works and adding new uses for them. Plaintiffs do not—and cannot—dispute the many transformative things that Claude can do. And Plaintiffs do not claim that Claude has ever reproduced a copy of their books.

**1. Plaintiffs do not seriously contest how LLM training works.**

Plaintiffs do not raise any genuine dispute of material fact regarding Anthropic’s account of the technical process for LLM training, as described in the declaration of its Chief Science Officer, Jared Kaplan. *See* Kaplan Decl. (ECF No. 119-5) ¶¶ 54-59. Plaintiff’s expert Dr. Zhao’s declaration makes vague criticisms about “nonspecific language” that is “not testable” (Zhao Decl. (ECF No. 170-4) ¶¶ 14(b), 38-53), but he does not identify any inaccuracy in Dr. Kaplan’s technical description. This strategic silence confirms there is no genuine dispute about the training process.

Dr. Zhao opines that he would generally characterize Anthropic’s pretraining process as similar to a “lossy compression” of its training data and criticizes Dr. Kaplan for “fail[ing] to grapple” with his “compression” framing (*id.* ¶ 55), which he contends causes model memorization. *See id.* ¶¶ 14(c), 54-77. Tellingly, Plaintiffs never took the opportunity to probe Anthropic on this framing. They asked Dr. Kaplan no questions about compression at his full-day deposition, even though it was taken *after* Dr. Kaplan’s declaration was submitted. Winthrop Decl. ¶ 3. And Dr. Zhao did not identify any technical inaccuracy in Dr. Kaplan’s declaration.

Nonetheless, even accepting Dr. Zhao’s compression analogy for the purposes of argument, it would not change the fair use analysis.<sup>5</sup> If one assumed “compression” resulted in an exact copy of training data—though Dr. Zhao does not even make that argument—courts frequently find uses are

---

<sup>5</sup> For the record, Anthropic does not agree that “compression” is an accurate or complete characterization of LLM training. Among other reasons, it fails to account for loss calculation optimization that occurs during pretraining and for the process that follows pretraining, including fine-tuning via supervised learning and reinforcement learning through human feedback, which optimize model behaviors. Kaplan Decl. ¶¶ 58, 60-64.



“transformative” even when the defendant’s ultimate use incorporates *exact copies* of the original work so long as the use is transformative. *See, e.g., Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 455 n.40 (1984) (holding verbatim unaltered copies of television shows to be fair use); *see also HathiTrust*, 755 F.3d at 97 (“digital copies of the entire books” were made to enable “creation of a full-text searchable database [that] is a quintessentially transformative use”); *Mattel, Inc. v. Walking Mountain Prods.*, 353 F.3d 792, 803 n.8 (9th Cir. 2003) (“We have, however, held that entire verbatim reproductions are justifiable where the purpose of the work differs from the original.”); *iParadigms II*, 562 F.3d at 639 (rejecting the plaintiffs’ argument that the defendant “merely stores the work unaltered and in its entirety” as “misguided,” because “[t]he use of a copyrighted work need not alter or augment the work to be transformative in nature . . . [r]ather, it can be transformative in function or purpose without altering or actually adding to the original work.”). Claude is not a substitute for a book and can do innumerable things that books cannot. That alone establishes that Anthropic’s use is transformative.

## **2. Plaintiffs’ non-transformative use citations are inapposite.**

Plaintiffs’ remaining cases on non-transformative uses actually highlight why Anthropic’s use is transformative, as they all involve defendants who merely repackaged copyrighted works and used them for the same purpose as the originals. *Opp.* at 14, 16-17. Not surprisingly, courts found such uses non-transformative. But these cases bear no resemblance to Anthropic’s use, which serves an entirely different purpose and creates new capabilities that the original works never had.

For example, Plaintiffs analogize Anthropic’s [REDACTED] to the facts in *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913 (2d Cir. 1994). *Opp.* at 14. But that defendant photocopied scientific journals obtained through an existing licensing program and then distributed those copies within the company for consumptive use by its employees rather than purchasing additional licenses. Because the secondary purpose and use of the copies was identical to the original purpose and use, this was non-transformative. 60 F.3d at 922-23. The remaining cases Plaintiffs cite also involved ultimate uses that were highly similar to the original use. *See Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1262-63 (11th Cir. 2014) (university’s distribution of photocopies of textbooks to students was not transformative because defendants did not use the digital copies “for anything other than the same intrinsic purpose . . . served by Plaintiffs’ works: reading material for students in university courses”);

*Wall Data Inc. v. Los Angeles Cnty. Sheriff's Dep't*, 447 F.3d 769, 778 (9th Cir. 2006) (sheriff's department installing several thousand unlicensed copies of software was not transformative because it "put those copies to the identical purpose as the original software," and "did not provide the marketplace with new creative works," or "advance[] . . . public knowledge") (internal quotations omitted); *Thomson Reuters Enter. Ctr. GMBH v. Ross Intel. Inc.*, 2025 WL 458520, at \*8 (D. Del. Feb. 11, 2025) (copying Westlaw headnotes to "develop a competing legal research tool" was not transformative).<sup>6</sup>

### III. THE SECOND AND THIRD FACTORS FAVOR FAIR USE

Both the second and third fair use factors favor Anthropic. With respect to the second factor, Plaintiffs argue that books are "creative works" deserving the highest protection under copyright law, as compared to "informational and functional works" like computer code. Opp. at 11 (internal quotations omitted). But Plaintiffs, who are authors of both fiction *and* non-fiction works, ignore that all books contain both expressive and functional elements—just like computer code. *Google Books*, 804 F.3d at 224 (plaintiff's "copyright does not extend to the facts communicated by his book" but instead "protects only the author's manner of expression"). Plaintiffs attempt to distinguish their books from the computer code in *Oracle*, *Sega*, and *Connectix* (Opp. at 11), but, similarly to those cases, Anthropic has used Plaintiffs' works to extract factual and functional information necessary for its LLMs to understand human language. Moreover, Plaintiffs ignore that Anthropic uses only published works. Accordingly, the second factor favors fair use. *E.g.*, *Sega*, 977 F.2d at 1527.

As to factor three, Plaintiffs do not contest that the transformative use in question requires Anthropic to use the full text of books to function effectively. Mot. (ECF No. 119-7) at 17-18; Kaplan Decl. ¶¶ 39, 42-43. The third factor favors fair use. *HathiTrust*, 755 F.3d at 98 (fair use for copying full books where it was "necessary to copy the entire copyrighted work.").

### IV. THE FOURTH FACTOR FAVORS FAIR USE BECAUSE ANTHROPIC'S HIGHLY TRANSFORMATIVE USE DOES NOT HARM THE MARKET FOR PLAINTIFFS' WORKS

The fourth factor strongly favors Anthropic because: (i) lost licensing fees are not cognizable market harms for the highly transformative use of LLM training, (ii) alternatively, there is no genuine dispute of fact about whether a licensing market to enable Anthropic's transformative use could develop;

---

<sup>6</sup> As explained in the Motion, that case was not about generative AI. Mot. at 14.

and (iii) Plaintiffs’ suggestions of other theories of harm based on lost book sales due to the LLM training process or AI-generated books have been waived, are not cognizable as a matter of law, and/or are unsupported.

**A. As a matter of law, Anthropic’s highly transformative use forecloses Plaintiffs’ theory of market harm based on lost licensing fees.**

When the defendant makes a highly transformative use that does not create a market substitute for the original works—such as Anthropic’s use here—lost licensing fees are not a cognizable harm as a matter of law. Mot. at 18. Here, Plaintiffs concede that “Claude itself is not [a] substitute for books.” James Malackowski Decl. (“JM Decl.”) (ECF No. 170-5) ¶ 38; Peterson Decl. (ECF No. 119-6) ¶¶ 5, 21-23, 31-33. This leaves “lost licensing revenues” as the only theory of harm based on Anthropic’s use of books as training data that is articulated in the Opposition. Opp. at 21.

The Court can quickly dispense with Plaintiffs’ suggestion that the law limiting recovery of licensing fees for transformative uses is cabined to cases involving “criticism, parody, or commentary.” *Id.* at 22. Plaintiffs cite no authority supporting such a narrow construction. And the Supreme Court’s decision in *Oracle*, which most definitely was not a case concerning “criticism, parody, or commentary,” invoked this principle in rejecting Oracle’s claim to certain lost profits from Google’s allegedly infringing use. 593 U.S. at 35-36 (claimed lost licensing fees not “cognizable under the Copyright Act”) (quoting *Campbell*, 510 U.S. at 591-92). And, there is a litany of cases holding that lost licensing fees for highly transformative uses do not constitute cognizable harms.

For example, *Tresóna Multimedia, LLC v. Burbank High Sch. Vocal Music Ass’n*, 953 F.3d 638, 651-52 (9th Cir. 2020) is directly on point. In *Tresóna*, the court held that plaintiff was “not harmed by the loss of any [licensing] fees” for the use of sheet music incorporating the Olivia Newton-John song “Magic” as part of an original eighteen-minute competitive choir arrangement because the arrangement was “not a substitute” for the full song. *Id.*; see also *HathiTrust*, 755 F.3d at 99 (lost licensing revenue did not “count” as a cognizable harm because the transformative full-text searches of books did not substitute for books themselves); *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614 (2d Cir. 2006) (finding no market harm from lost licensing fees when transformative use—including a



1 thumbnail concert poster in a book—served different purpose from original).<sup>7</sup>

2 As a matter of law, Plaintiffs cannot prevail on the circular argument that the fourth-factor market  
3 harm they have suffered consists solely of the failure to obtain royalties for a highly transformative use  
4 that does not function as a market substitute for their books.

5 **B. There is no genuine dispute about whether a viable licensing market for LLM training**  
6 **data could develop.**

7 Because the only market harm from LLM training Plaintiffs assert is not cognizable under the  
8 fourth fair use factor, the Court need not address the question whether LLM training data licensing  
9 markets exist or plausibly could. But if it does so, it should conclude as a matter of law that no such  
10 market exists or is likely to develop (absent congressional action).

11 Contrary to Plaintiffs’ contention (Opp. at 20), there is no battle of the experts here—because  
12 Plaintiffs’ expert Malackowski answers the wrong question. His analysis focuses narrowly on “*copyright*  
13 *registered books*” rather than addressing the existence or prospect of a licensing market for the aggregate  
14 data required to train a frontier LLM like Claude, which is what is actually needed to enable Anthropic’s  
15 purpose. JM Decl. ¶ 59. Licensing copyright registered books alone would not come close to yielding  
16 the data necessary for Anthropic to build a general purpose LLM. Kaplan Decl. ¶ 39; Peterson Decl. ¶¶  
17 15, 56. In fact, that volume of data would represent only a tiny fraction of the copyrighted data that is  
18 required for Anthropic’s transformative purpose and use. Kaplan Decl. ¶ 46; Peterson Decl. ¶ 58. It makes  
19 no sense to assess the viability of a relevant licensing market by reference to such a crimped fraction of  
20 the whole (*id.* ¶¶ 47-48), and no court frames the operative legal question in that way. To the contrary,  
21 the question courts ask is whether a market exists or plausibly could exist to license the material *needed*  
22 *to achieve the transformative purpose*. See, e.g., *White v. W. Pub. Corp.*, 29 F. Supp. 3d 396, 400  
23 (S.D.N.Y. 2014) (explaining that the fourth factor favors fair use because the transaction costs associated  
24 with any conceivable licensing regime that would be required for the defendant’s product “would be  
25 prohibitively high”). Plaintiffs have not even attempted to argue that a licensing market exists or could  
26 exist for LLM training data at the scale required to build a tool like Claude, and as a result there is no

27  
28 <sup>7</sup> For this reason, Plaintiffs’ citation to *Dr. Seuss Enters. L.P. v. ComicMix LLC*, 983 F.3d 443, 461 (9th Cir. 2020) is misplaced, as the court there was concerned with the risk of harm from non-transformative derivatives of Dr. Seuss books, for which there was an established market.

material dispute of fact that it does not and could not. *See* Mot. at 20-21.

Further, while the market for books that Malackowski addresses is wrong as a matter of law, his analysis of that narrow market **confirms** the far broader problem with licensing here. *Triton Energy Corp. v. Square D Co.*, 68 F.3d 1216 (9th Cir. 1995) (affirming summary judgment when “expert opinion and the inferences [the plaintiff] seeks to draw from it are not of sufficient quantum or quality to create genuine issues of material fact”). Plaintiffs highlight one example of an agreement for the use of copyrighted books as LLM training data—the HarperCollins deal. ECF No. 147-27. But that contract shows the **non-viability** of a market for licensing the registered copyrighted books necessary to train an LLM. The agreement [REDACTED] [REDACTED] JM Decl. ¶¶ 59, 61. Even assuming (with no evidence) that all [REDACTED] that would provide only a tiny fraction of **just the books** needed to train Claude. Replicating the same deal for an entire training corpus would be impossible; it is undisputed that identifying all the necessary rightsholders would be impossible. Supplemental Declaration of Steven R. Peterson ¶¶ 8-10. Even assuming that were not true, scaling the HarperCollins agreement [REDACTED] *Id.* ¶ 11. This is the very definition of market failure, which has long been recognized to be one of the core justifications for the fourth factor favoring fair use. *Texaco*, 60 F.3d at 929; *cf.* Wendy J. Gordon, *Fair Use As Market Failure*, 82 Colum. L. Rev. 1600, 1614-15 (1982).

There is also no precedent for the development of a viable licensing regime under circumstances remotely comparable to these. Malackowski’s contrary speculation, invoking digital music services and YouTube, turns only on facially inapt comparisons, not any dispute of fact. *Cf.* JM Decl. ¶¶ 26-27. With respect to music copyrights, the obvious impossibility of licensing the requisite rights at scale has triggered multiple rounds of *legislative intervention*, with Congress adopting and then revising a complex compulsory licensing regime to allow streaming services to pay statutorily-set royalties to a statutorily-identified centralized clearinghouse, as a unique solution to the problem that many of the operative rightsholders will never be identified at all. *See* 17 U.S.C. § 115 (establishing the compulsory license, last modified in 2018); *id.* § 115(d)(3) (authorizing the designation of a centralized royalty collector); *id.* § 115(d)(3)(I) (establishing a process to try to identify absentee rights owners); *id.* § 115(d)(3)(J)

(prescribing how unclaimed royalties are to be allocated). With respect to YouTube, the critical distinction, which Plaintiffs elide altogether, is again that bespoke copyright laws—not organic market development—enable the licensing regime they point to. Specifically, under Section 512(c) of the Digital Millennium Copyright Act, the default legal position for YouTube is *immunity* from copyright infringement liability for content posted by users. *Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 42-43 (2d Cir. 2012). Against that backdrop, rights owners can opt in to deals or send highly particularized takedown notices—but if they don’t, the content remains on the site and the site has no liability for it. No such legislative safe harbor exists to enable a market for AI training data.

In sum, developing a licensing market to address the undisputed scale of Anthropic’s needs in training Claude would require a new statutory framework, of the sort that Congress has enacted with respect to licensing music rights or providing safe harbor from monetary liability for vast amounts of copyrighted data on the internet. Plaintiffs’ expert raises no dispute of fact with respect to these conclusions; he underscores them.

**C. Other theories of harm from LLM training are waived, not cognizable, or unsupported.**

In a brief detour, Malackowski also speculates about two other market harms: (i) a “risk of future market substitution” because Claude helps people “write prose” (JM Decl. ¶ 35); and (ii) AI-generated “knockoff” books (JM Decl. ¶¶ 36-37). Plaintiffs do not argue either of these theories of harm in their Opposition itself, and they therefore have waived them. *John-Charles v. California*, 646 F.3d 1243, 1247 n.4 (9th Cir. 2011) (a party who has “failed to develop any argument . . . has waived it”). In any event, the first theory is wrong as a matter of law and the second theory is completely bereft of record support.

*First*, Malackowski’s theory that Claude’s writing ability harms the Plaintiffs is not cognizable under the Copyright Act. For purposes of the fourth factor, it does not matter whether the copier’s work enables creation of new works that compete *generally* for attention with the plaintiff’s work. Rather, the question is limited to what impact defendant’s use has on the value of plaintiff’s work in particular—*i.e.*, whether the resulting work directly usurps the market for plaintiff’s work. *See, e.g., Sega*, 977 F.2d at 1523-24 (finding no market harm based on copying game system code notwithstanding that it enabled Accolade to create Mike Ditka Power Football, which competed with Sega’s Joe Montana Football).

*Second*, any purported risk of “knockoff” books written by AI technology *generally* is

1 unsupported by any record facts or claims. There is no claim that Claude has ever generated a single  
2 “knockoff” book, or any claim that any consumer has ever purchased a Claude-generated book in favor  
3 of any of the Plaintiffs’ books. Plaintiffs themselves testified that they were not aware of any decline in  
4 their book sales related to AI. Mot. at 7. *See also* Peterson Decl. ¶¶ 41, 74-77 (AI-generated works  
5 unlikely to meaningfully alter the competitive dynamics of book sales). Speculation about hypothetical  
6 harms from AI-generated “knockoffs” does not create a dispute of material fact.

7 **D. Under Plaintiffs’ “initial acquisition” theory, there is no cognizable harm from lost book**  
8 **sales.**

9 As established above, the Court should reject as a matter of law Plaintiffs’ “initial acquisition”  
10 theory, *i.e.*, that the act of downloading should be viewed in isolation for fair use purposes. *See* Section  
11 II.A, *infra*. However, if the Court does a separate fair use analysis under that theory, it should find that  
12 Plaintiffs have not asserted a cognizable harm merely because they allege that they “did not receive the  
13 financial benefits they would have received from the purchase of lawful copies.” Opp. at 12. Copyright  
14 law does not automatically require that rightsholders receive a financial benefit when a work is  
15 “acquired.” There are many ways to lawfully acquire books without payment to the rightsholder, such as  
16 if a parodist checks a book out from a library or buys a used copy. Plaintiffs’ theory only makes sense in  
17 cases about non-transformative uses, such as a distribution of copies. For example, in *Napster* (as well  
18 as the other similar cases cited by Plaintiffs), the defendant made quintessentially *non-transformative*  
19 uses—*i.e.*, allowing millions of people to download and listen to exact copies of plaintiffs’ music without  
20 paying. *Napster*, 239 F.3d at 1017. Indeed, Plaintiffs expose this entire theory as a smokescreen by also  
21 taking the position that [REDACTED]

22 [REDACTED] is also copyright infringement. Because Anthropic’s sole purpose in  
23 acquiring books was to put them to a highly transformative use, its downloading was not a substitute for  
24 readers buying Plaintiffs’ books, and Plaintiffs suffered no cognizable harm as a matter of law.

25 **E. The public benefits of Claude support a fair use finding and do not present issues of fact.**

26 Finally, Plaintiffs do not contest the overwhelming evidence that Claude provides substantial  
27 public benefits—namely, the fact that millions of scholars, scientists, businesses, and creators are using  
28 Claude to advance creative endeavors such as scientific research, software coding, and writing. *See, e.g.*,

*Oracle*, 593 U.S. at 39-40 (absent fair use doctrine, Java API would have limited “the future creativity of new [computer] programs,” posing a “risk of creativity-related harms to the public”); *Perfect 10*, 508 F.3d at 1165 (“a search engine provides social benefit . . . namely, an electronic reference tool”).

Instead, Plaintiffs set up two strawman arguments. *First*, they assert that the “public will benefit more” from a regime in which authors are compensated for LLM training (Opp. at 22-23), but this argument is not supported by the record or any of the cases Plaintiffs cite, which all involved non-transformative uses that provided direct market substitutes for the original works. *See Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 195 (2d Cir. 2024) (internet library that allowed “original works [to] be copied and disseminated for free” would undermine incentives to create new books); *Ross*, 2025 WL 458520, at \*10 (legal research tool was a directly competitive market substitute); *MCA, Inc. v. Wilson*, 677 F.2d 180, 185 (2d Cir. 1981) (not fair use to “plagiarize a competitor’s copyrighted substitute dirty lyrics of his own, [and] perform it for commercial gain”). *Second*, contrary to Plaintiffs’ suggestion, the Court does not need to make any sweeping factual finding in order to find fair use, such as that AI technology will not “make human labor obsolete” or “whether LLMs hold more promise than peril.” Opp. at 23-24. The inquiry here in a copyright infringement action is solely into the “risk[s] of creativity-related harms to the public” *within the framework of the Copyright Act*. *Oracle*, 593 U.S. at 40. The record is undisputed that a creative tool like Claude advances the purpose of copyright law.

### CONCLUSION

Copyright law strikes a balance between the rights of copyright holders, who are granted a limited monopoly on copyrighted expression, and the rights of others to make “fair use” of those copyrighted works for the greater good. But there is no balance in the arguments that Plaintiffs make here. Claude is a radically transformative technology that serves an entirely different purpose from Plaintiffs’ books. Under settled principles of copyright fair use, Claude should be permitted to exist alongside Plaintiffs’ books, a result that promotes the “Progress of Science and useful Arts” for society as a whole. U.S. Const., Art. I, Sec. 8. Anthropic’s Motion for Summary Judgment should be granted.

Dated: May 8, 2025

Respectfully submitted,

**ARNOLD & PORTER KAYE SCHOLER LLP**

By: /s/ Douglas A. Winthrop  
DOUGLAS A. WINTHROP

*Attorneys for Defendant*  
ANTHROPIC PBC

**CERTIFICATE OF SERVICE**

I, Douglas A. Winthrop, am the ECF user whose identification and password are being used to file the foregoing **ANTHROPIC PBC'S REPLY IN SUPPORT OF MOTION FOR SUMMARY JUDGMENT.**

Dated: May 8, 2025

/s/ Douglas A. Winthrop