UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE OPENAI COPYRIGHT INFRINGEMENT LITIGATION	Case No. 1:25-md-03143 Judge: Hon. Sidney H. Stein

BRIEF OF AMICUS CURIAE COPYRIGHT ALLIANCE IN SUPPORT OF PLAINTIFFS' OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

TABLE OF CONTENTS

I. Introduction 1

A. AI Training Does Not Serve a Transformative Purpose, but Even If It Did, a Finding of Transformativeness Does Not Control the Fair Use Analysis. 2

B. Robust Licensing Markets for AI Training Exist and Market Harm Is Cognizable. 5

C. The Fair Use Decisions Cited by Defendants Are Significantly Undermined by the Supreme Court's Decision in *Warhol* 7

D. The Non-AI Cases on Which Defendants Rely Are Distinguishable. 8

1. *Authors Guild v. Google, Inc.* 8

2. Image Search Engine Cases 10

3. *Google LLC v. Oracle America, Inc.* 11

4. Reverse Engineering Cases 12

E. Defendants' Reliance on the *Kadrey* and *Bartz* Cases Is Misplaced. 13

1. *Kadrey* 14

2. *Bartz* 15

F. Knowingly Sourcing Pirated Materials Should Weigh Heavily Against Fair Use 16

II. Conclusion 17

TABLE OF AUTHORITIES

Page(s)

Cases

American Geophysical Union v. Texaco, Inc.,
60 F.3d 913 (2d Cir. 1994) 3, 5

Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith,
598 U.S. 508 (2023) 1, 2, 3, 4, 7, 8, 11, 15

Authors Guild v. Google, Inc.,
804 F.3d 202 (2d Cir. 2015) 1, 8, 9, 11, 12, 13

Bartz v. Anthropic PBC,
787 F. Supp. 3d 1007 (N.D. Cal. 2025) 1, 13, 15, 17

Campbell v. Acuff-Rose Music, Inc.,
510 U.S. 569 (1994) 2, 3, 7, 8

Castle Rock Ent., Inc. v. Carol Publ'g Grp., Inc.,
150 F.3d 132 (2d Cir. 1998) 7

Fox News Network, LLC v. TVEyes, Inc.,
883 F.3d 169 (2d Cir. 2018) 8

Google LLC v. Oracle America, Inc.,
593 U.S. 1 (2021) 11, 12

Kadrey v. Meta Platforms, Inc.,
788 F. Supp. 3d 1026 (N.D. Cal. 2025) 1, 4, 7, 13, 14, 15, 16

Kelly v. Arriba Soft Corp.,
336 F.3d 811 (9th Cir. 2003) 10, 11

NXIVM Corp. v. Ross Inst.,
364 F.3d 471 (2d Cir. 2004) 17

Perfect 10, Inc. v. Amazon.com, Inc.,
508 F.3d 1146 (9th Cir. 2007) 10

Princeton Univ. Press v. Mich. Document Servs., Inc.,
99 F.3d 1381 (6th Cir. 1996) (en banc) 5

Sega Enterprises. Ltd. v. Accolade, Inc.,
977 F.2d 1510 (9th Cir. 1992) 13

Sony Computer Entertainment, Inc. v. Connectix Corp.,
203 F.3d 596 (9th Cir. 2000) 12, 13

Thomson Reuters Enter. Ctr. GmbH v. ROSS Intel. Inc.,
— F.4th —, 2026 WL 2932557 (3d Cir. Sept. 29, 2026) 4

VHT, Inc. v. Zillow Grp., Inc.,
918 F.3d 723 (9th Cir. 2019) 9

Other Authorities

AI Licensing by Copyright Owners, Copyright Alliance 6

Circuit to Warhol Estate: Google v. Oracle *Does Not Dictate a Different Result*,
Pryor Cashman (Aug. 26, 2021) 12

Ed Nawotka, *Princeton University Press Partners with Cashmere on AI Licensing*, Publishers Wkly. (Aug. 3, 2026) 10

Getty Images and Perplexity Strike Multi-year Image Partnership, Getty Images
(Oct. 31, 2025) 6

Jiarui Liu, *An Empirical Study of Transformative Use in Copyright Law*, 22 Stan.
Tech. L. Rev. 163, 185–86 (2019), ... 8

Keith Kupferschmid, *You Can't Steal Your Way to Fair Use: The Unclean Hands Doctrine and AI Training*, Copyright Alliance (Sep. 28, 2026) 16

Major Labels Sign Licensing Deals With AI Music Company Klay, Pitchfork
(Nov. 20, 2025) 6

Matt O'Brien, *ChatGPT-maker OpenAI Signs Deal with AP to License News Stories*, Associated Press (July 13, 2023) 10

Michael M. Grynbaum & Cade Metz, *The Times and Amazon Announce an A.I. Licensing Deal*, N.Y. Times (May 29, 2025) 6

Neil Weinstock Netanel, *Making Sense of Fair Use*, 15 Lewis & Clark L. Rev.
715, 754–55 (2011) 8

Partnership with Axel Springer to Deepen Beneficial Use of AI in Journalism,
OpenAI (Dec. 13, 2023) 6

Shyamkrishna Balganesh, Jane C. Ginsburg & Peter S. Menell, *Comments on Preliminary Draft 9 of the American Law Institute's Restatement of Copyright*
at 15 (Sep. 26, 2023) 8

Supreme Court Affirms Andy Warhol's Prince Series Not Transformative Fair Use, Proskauer Rose LLP (June 14, 2023) 12

Technology & Marketing Law Blog (Apr. 8, 2021) 12

Text and Data Mining License, Elsevier (last visited Aug. 27, 2026) 10

Tyler Ochoa, *U.S. Supreme Court Upholds Fair Use in Google-Oracle Software Battle (Guest Blog Post)*,Technology & Marketing Law 12

Universal Music Group and Stability AI Announce Strategic Alliance to Co-Develop Professional AI Music Creation Tools, Universal Music Group (Oct. 30, 2025) 6

Universal Music Group to Transform Music Experience for Billions of Fans with NVIDIA AI, Universal Music Group (Jan. 6, 2026) 6

Warner Music Group and Stability AI Join Forces to Build the Next Generation of Responsible AI Tools for Music Creation, Warner Music Group (Nov. 19, 2025) 6

I. INTRODUCTION

The Copyright Alliance submits this amicus brief because our members have a strong interest in proper application of the Copyright Act, especially the application of copyright's fair use doctrine to generative artificial intelligence (GAI). The Supreme Court has emphasized that fair use must be applied on a case-by-case basis and is dependent on the specific facts of a case. Prior fair use cases may be instructive, but they do not dictate the outcome of cases, even when they involve similar facts, uses, and/or technologies. Yet Defendants seek to convince the Court otherwise and extract categorical holdings where there are none. Defendants argue that "fair use applies for all the same reasons" as in past fair use cases like *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015) [hereinafter *Google Books*], which "compels" a finding of fair use. OpenAI's Mem. of Points and Authorities in Supp. of Mot. for Summ. J. ("OpenAI MSJ") at 24 (Dkt. 1720). However, as other courts considering similar arguments in pending GAI infringement litigation have found, the cases upon which Defendants rely are distinguishable in a myriad of ways from the Defendants' use, and thus should be afforded very limited weight. Moreover, in arguing that the allegedly transformative nature of its models should tip the scales in favor of fair use, Defendants improperly equate "transformative" in the literal sense with transformative use as delineated by the Supreme Court. In doing so, they disregard the Supreme Court decision in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023) [hereinafter *Warhol*], which clarified that even when a court finds a use to be transformative, that finding should have a limited effect on the ultimate fair use determination. Finally, Defendants' reliance on the summary judgment orders in *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007 (N.D. Cal. 2025) and *Kadrey v. Meta Platforms, Inc.*, 788 F. Supp. 3d 1026 (N.D. Cal. 2025), is misplaced, as those cases demonstrate that the unauthorized use of copyrighted works for training will often not qualify as fair use. This is especially true when (1) a robust

licensing market exists for the use of copyrighted works for AI training, and (2) AI developers knowingly source pirated material for training, as OpenAI did. For these reasons, the Copyright Alliance respectfully submits that the Court should deny Defendants' motion for summary judgment.

A. AI Training Does Not Serve a Transformative Purpose, but Even If It Did, a Finding of Transformativeness Does Not Control the Fair Use Analysis.

Central to the Supreme Court's decision in *Warhol* was its confirmation that the defendant's purported "justification" must be considered as part of the first-factor analysis. This standard, which was first explained by the Supreme Court's discussion of parodic uses in *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994), was unmistakably reaffirmed in *Warhol*, which confirmed that when an original work and secondary use share the same or highly similar purposes, and the secondary use is commercial, "a particularly compelling justification is needed." *Warhol*, 598 U.S. at 547. Applying this standard to the facts of this case, there is insufficient justification to support Defendants' claims that copying massive amounts of pirated works and ignoring existing licenses offered by copyright owners were necessary to train its GAI models. *See* OpenAI MSJ at 20–21, 27; Microsoft Corp.'s Mem. of Law in Supp. of Mot. for Summ. J. in Books Plaintiffs' Consol. Cases ("Microsoft MSJ") at 18, 24.

Defendants wrongly seek to disaggregate the output of GAI models from the training process when describing the purpose of GAI. In describing the purpose of training, Defendants incorrectly claim the copying "was to derive broad, unprotectable statistical patterns" and "to teach the model generalizable rules of language and facts about the world." OpenAI MSJ at 1, 9. Defendants' motions ignore what comes after the initial training—most notably, the generation of output that *serves the same purpose* of the ingested works. The *Warhol* decision and many other pre-*Warhol* cases are careful to consider the purpose of a use holistically and not simply

end their analysis at a discrete or intermediate copying step—whether that be reverse engineering, shrinking a work into a thumbnail image, archiving a work for use in a plagiarism-detection service or a searchable digital repository for those with print disabilities, sampling, or training.

Demonstrating courts' considerations of the ultimate or actual purpose of a use, the Second Circuit in *American Geophysical Union v. Texaco, Inc.*, 60 F.3d 913 (2d Cir. 1994), rejected over-generalized fair use justifications for widespread infringement:

> The purposes illustrated by the categories listed in section 107 refer primarily to the work of authorship alleged to be a fair use, not to the activity in which the alleged infringer is engaged. Texaco cannot gain fair use insulation for [its employee]'s archival photocopying of articles (or books) simply because such copying is done by a company doing research. It would be equally extravagant for a newspaper to contend that because its business is "news reporting" it may line the shelves of its reporters with photocopies of books on journalism or that schools engaged in "teaching" may supply its faculty members with personal photocopies of books on educational techniques or substantive fields. Whatever benefit copying and reading such books might contribute to the process of "teaching" would not for that reason satisfy the test of a "teaching" purpose.

Id. at 924.

Warhol, *Texaco*, and many of the cases discussed below make clear that, when considering the purpose of the use in a factor-one fair use analysis, the analysis must not merely consider the intermediate step but rather must take into account the ultimate purpose of the use. Thus, the purpose of Defendants' models cannot be considered in a vacuum of "training." "Training," standing alone, is unlikely to constitute a sufficiently transformative use or purpose any more than "add[ing] new expression to [a] photograph," standing alone, was a transformative use or purpose in *Warhol*, 598 U.S. at 526, or "cop[ying]" the "musical phrase[] of the original," standing alone, was a transformative use or purpose in *Campbell*, 510 U.S. at 588. Defendants' copying of Plaintiffs' copyright-protected literary works for training must be

considered in view of its ultimate purpose: to generate output that serves the same purpose and acts as a substitute for the ingested works.

The Third Circuit recently applied this principle to AI training in the first appellate decision to address the question, rejecting the defendant's argument that copying the plaintiff's works to train an AI model was transformative. *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intel. Inc.*, — F.4th —, 2026 WL 2932557 (3d Cir. Sept. 29, 2026). The decision explains that while the defendant "took an intermediate step of using the [works] to train an AI program, which arguably presents a slight degree of difference in use," its use "shares the same ultimate purpose" as the plaintiff's use, making it "minimally transformative, at best." *Id.* at *6. The court also rejected the defendant's reliance on the same "intermediate copying" cases that Defendants invoke here, because in those cases "copying was necessary to access the unprotected functional aspects of computer code," whereas the defendant in *Thomson Reuters* had copied because doing so was "easy," and "[u]nlike necessity, ease is not a justification for copying." *Id.* at *7; *see infra* § D.4. While the court observed that Ross's AI model, unlike Defendants' models, could not "generate original expression," *id.* at *6 n.7, its method of analysis, which looks past the intermediate act of training to the ultimate purpose that the training serves, applies equally to generative AI.

Furthermore, even if the use is found to be for a transformative purpose, the Supreme Court made clear in *Warhol* that whether a use is transformative is not dispositive of the question of fair use; rather, it is merely one subfactor within the first statutory fair use factor. Indeed, as Judge Chhabria recognized in *Kadrey*, "[t]here is certainly no rule that when your use of a protected work is 'transformative,' this automatically inoculates you from a claim of copyright infringement." *Kadrey*, 788 F. Supp. 3d at 1035. Any factor-one analysis must involve a

weighing of other considerations, such as the commercial nature of and justification for the use, both of which will factor prominently. Defendants' models are clearly part of a commercial venture, designed to bolster its position in the market, which weighs against fair use under the first factor and offsets in whole or in part any finding of transformative purpose. Therefore, even if Defendants' use of copyrighted works for ingestion were found to be a transformative use in this case (though it is not), that would not necessarily mean factor one favors fair use.

B. Robust Licensing Markets for AI Training Exist and Market Harm Is Cognizable.

Defendants incorrectly argue that the fourth fair use factor weighs in its favor because "Plaintiffs are not entitled to a licensing fee" for use of their works for training because the use is transformative, and therefore any harm to the licensing market is "circular" and "not legally cognizable." OpenAI MSJ at 36, 37. However, circularity arguments have been invoked by defendants accused of infringement well before generative AI cases, and they have been rejected as dispositive under the fourth factor, especially when an actual licensing marker for the use already exists. *See Princeton Univ. Press v. Mich. Document Servs., Inc.*, 99 F.3d 1381, 1387 (6th Cir. 1996) (en banc) (rejecting a "circularity" argument against considering lost licensing revenue under factor four as "prov[ing] too much," and finding no circularity where an actual licensing market for the use already exists); *Texaco*, 60 F.3d at 931 (explaining that circularity is a concern only where "the availability of payment is conclusive against fair use"). There is nothing speculative or theoretical about the robust and ever-expanding licensing market for AI training. Since the public launch of GAI models a few years ago, there have been thousands of voluntary licenses directly entered into by copyright owners in connection with AI

development.[1] There are numerous examples of AI companies that are directly entering into licensing agreements with (and thereby compensating) copyright owners for the ingestion of their works, underscoring the viability and vibrancy of the voluntary AI licensing market supported by a robust copyright law framework.[2]

Defendants claim that there is no harm under the fourth factor because the training market for Plaintiffs' works is not "traditional, reasonable, or likely to be developed." OpenAI MSJ at 37 (citation omitted). However, there is no question that, given the strong existing market for books and news articles for AI training, a market for plaintiffs' works is reasonable and likely to be developed. Moreover, the Third Circuit's decision in *Thomson Reuters* confirms that harm to the emerging market for AI training licenses is cognizable under the fourth factor. The defendant there argued, as Defendants do here, that no such market existed. *Thomson Reuters*, 2026 WL 2932557, at *9. But the court explained that "a potential derivative market is not illusory just because an author 'has evidenced little if any interest in exploiting this market for

[1] *See AI Licensing by Copyright Owners*, Copyright Alliance, https://copyrightalliance.org/artificial-intelligence-copyright/licensing/copyright-owners (last visited Oct. 29, 2026) (listing over 200 such licenses—a list that does not include numerous confidential agreements or those that are not publicly reported).

[2] Recent examples of free-market licensing and partnership agreements include the ones between Universal Music Group and NVIDIA, *Universal Music Group to Transform Music Experience for Billions of Fans with NVIDIA AI*, Universal Music Group, (Jan. 6, 2026), https://www.universalmusic.com/universal-music-group-to-transform-music-experience-for-billions-of-fans-with-nvidia-ai; between the *New York Times* and Amazon, Michael M. Grynbaum & Cade Metz, *The Times and Amazon Announce an A.I. Licensing Deal*, N.Y. Times (May 29, 2025), https://www.nytimes.com/2025/05/29/business/media/new-york-times-amazon-ai-licensing.html; between various news publishers and OpenAI, *e.g.*, *Partnership with Axel Springer to Deepen Beneficial Use of AI in Journalism*, OpenAI (Dec. 13, 2023), https://openai.com/index/axel-springer-partnership; between Universal Music Group and Stability AI, *Universal Music Group and Stability AI Announce Strategic Alliance to Co-Develop Professional AI Music Creation Tools*, Universal Music Group (Oct. 30, 2025), https://www.universalmusic.com/universal-music-group-and-stability-ai-announce-strategic-alliance-to-co-develop-professional-ai-music-creation-tools; between Getty Images and Perplexity, *Getty Images and Perplexity Strike Multi-year Image Partnership*, Getty Images (Oct. 31, 2025), https://newsroom.gettyimages.com/en/getty-images/getty-images-and-perplexity-strike-multi-year-image-partnership; between Klay Vision, Inc. and Universal Music Group, Sony Music Group, and Warner Music Group, Nina Corcoran, *Major Labels Sign Licensing Deals With AI Music Company Klay*, Pitchfork (Nov. 20, 2025), https://pitchfork.com/news/warner-music-group-signs-licensing-deal-with-ai-music-company-klay; and between Warner Music Group and Stability AI, *Warner Music Group and Stability AI Join Forces to Build the Next Generation of Responsible AI Tools for Music Creation*, Warner Music Group (Nov. 19, 2025), https://www.wmg.com/news/warner-music-group-and-stability-ai-join-forces-to-build-the-next-generation-of-responsible-ai-tools-for-music-creation.

derivative works,'" *id.* (quoting *Castle Rock Ent., Inc. v. Carol Publ'g Grp., Inc.*, 150 F.3d 132, 145–46 (2d Cir. 1998)), and that the fact that the copyright owner "did not license its [works] to others does not disprove that a market exists to do so." *Id.* Finding that "the market for licensing [the works] as text to train AI is rapidly developing," the court held that by copying the works "for use as training data without . . . authorization," the defendant "usurped [the copyright owner's] opportunity to enter that derivative market and license its [works] for that purpose," and that this "market harm . . . weighs against fair use." *Id.* The same is true here. OpenAI's suggestion that Plaintiffs suffered no harm because they have not themselves licensed their books for AI training, *see* OpenAI MSJ at 37, is no more an answer in this case than it was in *Thomson Reuters*.

Defendants also argue that licensing markets "make[] little economic sense" because "the transaction costs for licensing any individual works (or even groups of works) would vastly exceed the marginal benefit of those works." OpenAI MSJ at 37–38. However, it is not the job of courts to evaluate market efficiencies—that is for Congress. Moreover, Defendants' licensing claims are nothing more than the self-serving arguments that were rejected by Judge Chhabria in *Kadrey*: "These products are expected to generate billions, even trillions, of dollars for the companies that are developing them. If using copyrighted works to train the models is as necessary as the companies say, they will figure out a way to compensate copyright holders for it." *Kadrey*, 788 F. Supp. 3d at 1036.

C. The Fair Use Decisions Cited by Defendants Are Significantly Undermined by the Supreme Court's Decision in *Warhol*.

Many of the cases on which Defendants rely for their argument that the copying of Plaintiffs' works is transformative were decided after *Campbell*, but before *Warhol*. During that three-decade period, courts took a very broad view of transformativeness, and in fact, a finding

of transformativeness was very often dispositive of the ultimate fair use question.[3] But *Warhol* dramatically changed the landscape, both narrowing what "transformative" means and clarifying that a finding of transformativeness does not lead ineluctably to a conclusion that the use was fair. Thus, all of the fair use cases on which Defendants rely, which were decided between *Campbell* and *Warhol*, and all of which relied heavily on a transformative use finding for their ultimate holding of fair use, simply do not reflect the current, post-*Warhol*, state of the law.[4]

D. The Non-AI Cases on Which Defendants Rely Are Distinguishable.

1. *Authors Guild v. Google, Inc.*

While Defendants claim that *Google Books* "compels" a finding of fair use in this case, it is distinguishable because it involved a completely different purpose for the use of the copyrighted materials—to provide information location services to drive readers to the relevant source material. OpenAI MSJ at 24. Concluding that Google's use was transformative, the appellate court held that "Google's making of a digital copy to provide a search function . . . augments public knowledge by making available information *about* [p]laintiffs' books without

[3] Indeed, between 1994 and 2023, a determination that a use was transformative almost always meant that a court concluded it was fair use. A 2011 study found that of all the fair use cases decided in 2006 through 2010, the defendant won nearly 100% of the time when the court found the subject use to be transformative. Neil Weinstock Netanel, *Making Sense of Fair Use*, 15 Lewis & Clark L. Rev. 715, 754–55 (2011). Between 2010 and 2023, the Copyright Alliance is aware of only one circuit court decision holding that a use was transformative but not fair use. Even then, judges within this Circuit took issue with expansive notions of transformativeness. *See Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 183 (2d Cir. 2018) (Kaplan, D.J., concurring) (explaining that "[a]s noted commentators have observed, courts 'appear to label a use "not transformative" as a shorthand for "not fair," and correlatively "transformative" for "fair." Such a strategy empties the term of meaning.'" (citation omitted)); *see also* Jiarui Liu, *An Empirical Study of Transformative Use in Copyright Law*, 22 Stan. Tech. L. Rev. 163, 185–86 (2019), *available at* https://ssrn.com/abstract=3330236 (finding that transformative use was the only statistically significant subfactor driving the first factor determination).

[4] Shyamkrishna Balganesh, Jane C. Ginsburg & Peter S. Menell, *Comments on Preliminary Draft 9 of the American Law Institute's Restatement of Copyright* at 15 (Sep. 26, 2023), https://scholarship.law.columbia.edu/cgi/viewcontent.cgi?article=5177&context=faculty_scholarship ("The Supreme Court's Warhol decision shifted fair use analysis back to the course that Congress intended. It is important to recognize many cases decided during the interim period [between *Campbell* and *Warhol*] that boil fair use analysis down to a mere transformativeness inquiry are no longer good law.").

providing the public with a substantial substitute for matter protected by the [p]laintiffs' copyright interests in the original works or derivatives of them." *Google Books*, 804 F.3d at 207.

Significantly, the decision made clear that the case "tests the boundaries of fair use," *id.* at 206—a position that a sister circuit has agreed with—and may have come out differently had any one of several factors varied. *See VHT, Inc. v. Zillow Grp., Inc.*, 918 F.3d 723, 743 (9th Cir. 2019) ("We agree with the Second Circuit's observation that the copyright dispute over the Google Books search engine 'tests the boundaries of fair use.'" (citation omitted). First, the Second Circuit explained that the fair use analysis would have been different if the purpose of Google's scanning of literary works was to create substitutes for the original works. *See Google Books*, 804 F.3d at 214, 218–19. In other words, it was critical that Google used the books to provide information *about* the works and by serving as a "pointer" for readers to consume the original works. By contrast, Defendants cannibalize Plaintiffs' copyrighted works for the purpose of allowing its users to generate the same type of works and thus is much more likely to usurp the market for (and obviate any need to consume) Plaintiffs' works.

Second, in *Google Books*, Google used the works solely *for the information* in the works, but in the instant case, Defendants are using the information and the *expressive elements* of the works. What Defendants consider to be unprotectable information is actually protected copyrightable expression.

Third, in *Google Books*, Google copied legitimate copies of books. But in the instant case, OpenAI copied copyrighted works from a source known to contain *pirated* works. *See* OpenAI MSJ at 3; Microsoft MSJ at 17 (describing Library Genesis as "an online repository known to contain unauthorized copies of books").

Fourth, and perhaps most important, the court in *Google Books* concluded (with regard to the fourth fair use factor) that there was no actual or potential market for the licensing of copyrighted works to search engines. *Id.* at 226–27. This is significantly different than with GAI, where there is very clearly an actual market for the licensing of copyrighted works for ingestion.[5]

Any of these differences standing alone would likely be sufficient to push this case beyond "the boundaries of fair use" set forth in *Google Books*. When taken together, there can be no doubt that they do.

2. Image Search Engine Cases

Defendants cite the Ninth's Circuit's decisions in *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003) and *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007) as evidence that courts have found that it is fair use "where transformative uses display portions of the original copyrighted content, so long as the new use does not serve a 'competing substitute for the original.'" OpenAI MSJ at 18 (citation omitted). In these cases, however, the court held that the defendants' reproduction of the plaintiff's photos as thumbnail images qualified as fair use because the thumbnail images served an entirely different purpose than the original images. In *Arriba*, the court held that the plaintiff's photographs were "artistic works intended to inform and to engage the viewer in an aesthetic experience," in contrast to Arriba's use, which the court found was "unrelated to any aesthetic purpose" and that the search engine instead "functions as a

[5] There is already high demand for corpora of copyrighted works for ingestion by AI systems, and copyright owners are offering and entering into various licensing agreements. Publishers and copyright owners of scientific and research works have either offered or entered into licensing agreements that allow for text and data mining (TDM) or other GAI-related uses. *See, e.g.*, *Text and Data Mining License*, Elsevier, https://www.elsevier.com/about/policies-and-standards/text-and-data-mining/license (last visited Aug. 27, 2026); Ed Nawotka, *Princeton University Press Partners with Cashmere on AI Licensing*, Publishers Wkly. (Aug. 3, 2026), https://www.publishersweekly.com/pw/by-topic/industry-news/publisher-news/article/100971-princeton-university-press-partners-with-cashmere-on-ai-licensing.html. Multiple news organizations, including NewsCorp, the Associated Press, the *Atlantic*, and the *Financial Times*, have reached deals with OpenAI for use of their works to train ChatGPT. *See, e.g.*, Matt O'Brien, *ChatGPT-maker OpenAI Signs Deal with AP to License News Stories*, Associated Press (July 13, 2023), https://www.ap.org/media-center/ap-in-the-news/2023/chatgpt-maker-openai-signs-deal-with-ap-to-license-news-stories/.

tool to help index and improve access to images on the internet and their related web sites." *Arriba*, 336 F.3d at 818.

Here, by contrast, both the purpose of Plaintiffs' works that are ingested and the purpose of the material generated by Defendants' models are exactly the same: "to be read and enjoyed by a human audience." OpenAI MSJ at 20. Because both the ingested work and AI-generated output serve the same purpose, that "seriously weakens" the fair use claim. *Arriba*, 336 F.3d at 819 (citation omitted).

The court also ruled in favor of fair use in *Arriba* because "[t]he thumbnails do not stifle artistic creativity because they are not used for illustrative or artistic purposes and therefore *do not supplant the need for the originals*." *Id.* at 820 (emphasis added). This is very different from the generation of literary material by Defendants' AI models, where in many cases it is likely that such AI-generated output would "supplant the need for the original [i.e., the ingested work]." *Id.*

3. *Google LLC v. Oracle America, Inc.*

Defendants assert that the Supreme Court's decision in *Google LLC v. Oracle America, Inc.*, 593 U.S. 1 (2021), also supports a finding of fair use because of the "great weight" that should be assigned to a use that enables "new applications." *See* OpenAI MSJ at 18 (citation omitted); *see also* Microsoft MSJ at 14 ("[I]f the copying in *Oracle* is transformative, LLM training is, too."). But it is widely recognized that *Google*, which involved copying of functional computer code, has a very limited application—and is simply inapplicable when the work being used is a traditional copyrighted work (e.g., books, news articles, songs, movies, photographs, etc.).[6] The decision in *Google* is expressly limited to the specific type of computer declaring

[6] *See, e.g.*, Brief for the United States as Amicus Curiae Supporting Respondents at 25, *Warhol*, 598 U.S. 508 (No. 21-869), 2022 WL 3574326 ("[Andy Warhol Foundation's] reliance on *Google* is likewise misplaced.

code at issue in that case and cannot be applied to past or future fair use analyses outside of that specific context. *See Google*, 593 U.S. at 40 (clarifying that its decision "do[es] not overturn or modify [its] earlier cases involving fair use").

Significantly, the Court was careful to distinguish computer code from highly expressive works (like those at issue in this case) that have no functional elements that may impact a factor-two analysis, stating that "computer programs differ from books, films, and many other 'literary works' in that [software] programs almost always serve functional purposes." *Id.* at 21. When considering the facts and circumstances surrounding Defendants' use of Plaintiffs' creative, expressive works of authorship, the instant case is easily distinguishable from the software-specific facts of *Google*.

4. Reverse Engineering Cases

Defendants also wrongly rely on *Sony Computer Entertainment, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), in which the Ninth Circuit concluded that Connectix's intermediate copying for reverse engineering qualified as fair use, necessary to permit Connectix to make its non-infringing game station function with PlayStation games. *See* Microsoft MSJ at 15

Emphasizing the difficulty of 'apply[ing] traditional copyright concepts in th[e] technological world' of 'functional' computer programs, the *Google* Court principally focused on the second statutory factor, emphasizing the copied code's distance 'from the core of copyright[.]'" (citation omitted))"; *Circuit to Warhol Estate:* Google v. Oracle *Does Not Dictate a Different Result,* Pryor Cashman (Aug. 26, 2021), https://www.pryorcashman.com/publications/circuit-to-warholestate-google-v-oracle-does-not-dictate-a-different-result ("The Circuit explained how the Supreme Court 'took pains to emphasize' that the unique context of the [*Google*] case (software code) may make its holding less applicable in other contexts, especially in the context where the copyrighted material 'serves an artistic rather than a utilitarian function.'"); *Supreme Court Affirms Andy Warhol's Prince Series Not Transformative Fair Use*, Proskauer Rose LLP (June 14, 2023), https://www.proskauer.com/blog/supreme-court-affirms-andy-warholsprince-series-not-transformative-fair-use ("Recently, the court in *Google v. Oracle* interpreted fair use broadly but limited its decision to the context of software codes."); Tyler Ochoa, *U.S. Supreme Court Upholds Fair Use in Google-Oracle Software Battle (Guest Blog Post)*, Technology & Marketing Law Blog (Apr. 8, 2021), https://blog.ericgoldman.org/archives/2021/04/u-s-supreme-court-upholds-fair-use-in-google-oracle-software-battleguest-blog-post.htm ("Any Supreme Court opinion is important, and this one no doubt will be quoted often in future briefs and opinions. But other than clarifying the standard of review, I doubt the decision will have much impact on fair use cases that do not involve software.").

(contending that OpenAI used Plaintiffs' works "to reverse engineer the functional components that enable the act of expression itself"); OpenAI MSJ at 24 n.14. Specifically, the court held that Sony's copyrighted software code included functional elements that resulted in a lower degree of protection. Similar to the decision in *Google,* the decision in *Sony* largely depended on the analysis of factor two and the work at issue being a computer program. The factor-two analysis was central to the ultimate fair use finding, and it would be inapplicable to the highly expressive works copied by Defendants here.

The Ninth Circuit in *Sega Enterprises. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), a case involving facts very similar to *Sony*'s, further confirmed the inapplicability of reverse engineering computer code cases when the works at issue are highly expressive. The court explained that "[b]ecause Sega's video game programs contain unprotected aspects that cannot be examined without copying, we afford them a lower degree of protection than more traditional literary works." *Id.* at 1526. The works that Defendants copied are exactly the type of traditional literary works that the Ninth Circuit confirmed deserve a higher degree of protection and thereby significantly alter any fair use analysis. Defendants concede as much. *See* Microsoft MSJ at 24 ("Training does use full copies of works that are at the core of copyright's subject matter . . . .").

E. Defendants' Reliance on the *Kadrey* and *Bartz* Cases Is Misplaced.

Defendants misleadingly present *Kadrey* and *Bartz* as a settled judicial consensus that training generative AI models on copyrighted books and news articles is fair use. However, while reaching different conclusions, *Kadrey* and *Bartz* demonstrate that in most cases the unauthorized use of copyrighted works for training will not qualify as fair use.

1. *Kadrey*

Defendants rely on *Kadrey* to argue that their use of copyrighted works is "highly transformative" and should thus qualify as fair use. OpenAI MSJ at 1, 21. However, as discussed above, even if a use is found to be transformative under the first fair use factor, that is not dispositive of the analysis.

Defendants cite *Kadrey* to argue that the purpose of their models is distinct from that of the ingested works. But *Kadrey*'s conclusion turned on different facts and the specific record before the court, and Judge Chhabria warned that training an LLM designed "to create works that competed with the originals" "could be less transformative," because "that use would have the purpose and character of enabling an LLM to develop substitute works." *Kadrey*, 788 F. Supp. 3d at 1045 n.6. Here, Defendants describe the purpose of the copyrighted works as "to be read and enjoyed by a human audience," while they claim the purpose of the use for training was merely "to distill broad statistical patterns," and "teach the model." *Id*. at 9, 20. At a very high level that may be one purpose of copying the works, but there is no doubt that another purpose is to generate output to be read and enjoyed by a human audience. When the output of a model serves the same purposes as the works used to build it, the training that makes that output possible cannot be characterized as serving a wholly distinct purpose.

Kadrey also rejected arguments that Defendants repeat here, claiming that they copied Plaintiffs' books and news articles only to extract "unprotectable" statistical patterns. OpenAI MSJ at 26; *see* Microsoft MSJ at 15, 24. Meta made the same argument, and the court was unpersuaded: "Word order, word choice, grammar, and syntax are how people express their ideas," so "even though LLMs may only learn about 'statistical relationships,' those relationships are the product of creative expression." *Kadrey*, 788 F. Supp. 3d at 1049.

2. *Bartz*

Defendants cite the *Bartz* case to argue that its technology is "quintessentially" and "spectacularly" transformative, which it claims favors fair use. OpenAI MSJ at 1 (citation omitted). But in *Bartz*, Judge Alsup erroneously applied a literal or common meaning of "transformative" instead of applying the Supreme Court's definition of "transformative use." Under well-established case law and re-emphasized in the Supreme Court's recent decision in *Warhol*, a use is not transformative merely because it produces something new and innovative. To be transformative, the Supreme Court confirmed that a use must (1) have a further purpose or different character, (2) be "justified because copying is reasonably necessary to achieve the user's new purpose" and (3) "further[] the goal of copyright, namely, to promote the progress of science and the arts, without diminishing the incentive to create." *Warhol*, 598 U.S. at 531–32. In the *Bartz* case, Judge Alsup failed to apply this critical analysis, and to take the same approach in this case would only serve to propagate the error.

Defendants also rely on *Bartz* to analogize the training of models to a "student" preparing for an exam and to "humans who have read Plaintiffs' books and been inspired by them." OpenAI MSJ at 7–8, 32–33. But that is not how computers work and is merely an attempt to humanize the functions of a machine and avoid liability. Judge Chhabria explains the critical distinction in the *Kadrey* decision:

> [A]n LLM's consumption of a book is different than a person's. An LLM ingests text to learn "statistical patterns" of how words are used together in different contexts. It does so by taking a piece of text from its training data, removing a word from that text, predicting what that word will be, and updating its general understanding of language based on whether it was right or wrong—and then repeating this exercise billions or trillions of times with different text. This is not how a human reads a book.

Kadrey, 788 F. Supp. 3d at 1044.

The reality is that GAI models ingest and reproduce perfect copies of copyrighted works (and generate material) at a speed and scale that makes any comparison to human learning utterly inapt, and Defendants' attempt to liken the two should be roundly rejected.

F. Knowingly Sourcing Pirated Materials Should Weigh Heavily Against Fair Use

OpenAI acknowledges that it obtained copyrighted works for training by downloading "compilations of books from Library Genesis." OpenAI MSJ at 3. Neither Defendant suggests that OpenAI was unaware that Library Genesis is a shadow library "known to contain unauthorized copies of books." Microsoft MSJ at 17. Defendants instead argue that how they "came upon" copies of works is "incidental," Microsoft MSJ at 2, and that downloading from a pirate site is merely one step in "a single transformative use," OpenAI MSJ at 23. However, downloading illegal copies of works from known piracy repositories is not immunized by a use that may come after. Even *Kadrey* recognized that "downloading is a different use from any copying done in the course of LLM training." 788 F. Supp. 3d at 1047.

While Defendants claim that "bad faith" or "the propriety of the defendant's conduct" is not relevant to the fair use analysis, Microsoft MSJ at 19 (citation omitted), when AI companies knowingly use massive amounts of pirated copies of works without authorization, such use is not a mere act of bad faith—it is an act that borders on criminal infringement that should weigh against fair use.[7] The key difference between past cases that have considered bad faith and the current AI infringement cases is that those past cases involved the use of one or many *legitimate* copies, and that use was found to be fair use because the copy was not otherwise available. *See*

[7] *See* Keith Kupferschmid, *You Can't Steal Your Way to Fair Use: The Unclean Hands Doctrine and AI Training*, Copyright Alliance (Sep. 28, 2026), https://copyrightalliance.org/ai-piracy-problem-part-four/ (arguing that, rather than treating an AI developer's knowing use of pirated works only as evidence of bad faith under the first fair use factor, courts may invoke the equitable doctrine of unclean hands to bar the fair-use defense altogether, given the Supreme Court's characterization of fair use as an "equitable rule of reason" (citation omitted)).

NXIVM Corp. v. Ross Inst., 364 F.3d 471, 478 (2d Cir. 2004) (finding fair use where defendants' only access to the unpublished, non-public course manual came through an alleged breach of a non-disclosure agreement). In contrast, many of the pending AI training infringement cases involve the use of *millions of illegal copies*. Moreover, copyrighted works that AI companies wish to use for training are largely available through legitimate sources, which is in stark contrast to past cases where works were not available. The fact that AI companies' intentional use of these illicit sites and services does not presumptively disqualify them from claiming fair use is antithetical to the foundations and goals of our copyright system and the rights it guarantees. Judge Alsup recognized this in *Bartz*, finding that Anthropic's sourcing of pirated works did not qualify as fair use and that "bad faith is not the basis for this decision." *Bartz*, 787 F. Supp. 3d at 1026.

II. CONCLUSION

The cases relied on by Defendants for their fair use arguments are of limited relevance here given the narrowed interpretation of transformativeness post-*Warhol*, and the existence of a robust licensing market for training and use of pirated works should weigh against fair use. This Court should deny Defendants' motions for summary judgment.

Respectfully Submitted,

DATED: October 29, 2026.

KELLER ROHRBACK L.L.P.

By */s/ Ron Kilgard*

Ron Kilgard (SBN 4976916)
rkilgard@kellerrohrback.com

1140 6TH AVE, 9TH FLOOR
NEW YORK, NY 10036
TELEPHONE: (206) 623-1900
FACSIMILE: (206) 623-3384

Benjamin Gould
bgould@kellerrohrback.com
1201 Third Ave, Suite 3400
Seattle, WA 98101
TELEPHONE: (206) 623-1900
FACSIMILE: (206) 623-3384

Attorneys for Amicus Copyright Alliance

CERTIFICATE OF COMPLIANCE

In accordance with Local Civil Rule 7.1(c), I hereby certify that the foregoing Brief of Amicus Curiae is 5,881 words, exclusive of the caption page, table of contents, table of authorities, and signature block as determined by the word processing program used to prepare it.

Dated: October 29, 2026.

By: */s/ Ron Kilgard*
Ron Kilgard