United States District Court
Northern District of California

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

CONCORD MUSIC GROUP, INC., et al.,

Publishers,

v.

ANTHROPIC PBC,

Anthropic.

Case No. 24-cv-03811-EKL

ORDER DENYING MOTION TO DISMISS FIRST AMENDED COMPLAINT

Re: Dkt. No. 359

In this case, Plaintiffs allege that Defendant Anthropic PBC ("Anthropic") used Plaintiffs' copyrighted song lyrics to train a generative artificial intelligence ("AI") model, resulting in the unauthorized output of these copyrighted lyrics to third-party users. Plaintiffs, eight music publishing companies who own or control exclusive rights to millions of musical compositions,[1] assert claims for direct and secondary copyright infringement, and removal of copyright management information ("CMI"). The Court previously granted Anthropic's motion to dismiss the secondary infringement and removal of CMI claims, with leave to amend. Mot. to Dismiss Order, ECF No. 322 ("Order"). Publishers subsequently filed a first amended complaint, ECF No. 337 ("FAC"), and Anthropic moved to dismiss this complaint, Mot. to Dismiss, ECF No. 359 ("Mot."). The Court found this matter suitable for disposition without oral argument and vacated the hearing scheduled for September 3, 2025. *See* Civil L.R. 7-1(b). The Court has carefully reviewed the parties' briefs, the operative complaint, and the relevant authority. For the following reasons, the Court DENIES Anthropic's motion.

[1] Plaintiffs consist of Concord Music Group, Inc., Capitol CMG, Inc., Universal Music Corp., Songs of Universal, Inc., Universal Music - MGB NA LLC, Polygram Publishing, Inc., Universal Music - Z Tunes LLC, and ABKCO Music, Inc. (collectively, "Publishers").

I. BACKGROUND

The Court assumes the parties' familiarity with the facts of the case, the applicable legal standards, and the arguments made by the parties. Here, the Court discusses the facts and procedural background to the extent they are relevant to this motion.

A. Facts

Publishers control or own exclusive rights to millions of musical compositions, including 500 original works listed in Exhibit A to the operative complaint. FAC ¶ 44; *Id.*, Ex. A. Anthropic is a technology company that produces a series of AI large language models referred to as "Claude." *Id.* ¶ 53. Claude provides outputs in response to user prompts based on data that Anthropic has collected and used to train Claude. *Id.* ¶ 57.

According to Publishers, when curating the dataset used to develop and train Claude, Anthropic used Publishers' copyrighted lyrics. *Id.* ¶¶ 6, 58(a). It did so by supposedly leveraging third-party datasets, like "Common Crawl" and "The Pile," which contain a significant number of Publishers' copyrighted lyrics that are scraped from various websites. *Id.* ¶¶ 67, 69. The Publishers claim that, when training Claude, Anthropic decided to use tools, such as the Newspaper algorithm, to extract CMI from these datasets. *Id.* ¶¶ 68, 72-73. While Anthropic apparently had a choice to use algorithms like jusText that would retain more CMI in the dataset, it purposefully decided to use Newspaper, which removed CMI more effectively. *Id.* ¶ 73.

Additionally, Publishers assert that Anthropic can program "guardrails" that regulate Claude's output of copyrighted lyrics, including Publishers' lyrics. *Id.* ¶ 119. Publishers claim that, in developing and improving these guardrails, Anthropic "collected Claude user prompts and output data, including specific infringing output copying lyrics." And Anthropic supposedly became aware of the user prompts seeking such lyrics because, each time this occurred, the guardrails detected the users' actions and Anthropic received a corresponding notice. *Id.* ¶ 121. Publishers also contend that in conducting a broader study of Claude users' behavior, Anthropic analyzed "specific efforts by users to avoid these guardrails and identif[ied] instances in which the guardrails have failed and Claude has generated infringing output, including output copying Publishers' lyrics." *Id.*

Based on Anthropic's alleged use of Publishers' copyrighted lyrics, Publishers brought claims for (1) direct copyright infringement (Count I); (2) contributory copyright infringement (Count II); (3) vicarious copyright infringement (Count III); and removal or alteration of CMI under the Digital Millenium Copyright Act ("DMCA") (Count IV). *Id.* ¶¶ 159-204.

B. Procedural History

Anthropic moved to dismiss Publishers' secondary liability and DMCA claims, on August 15, 2024. ECF No. 205. After briefing and a hearing, the Court granted the motion to dismiss with leave to amend, on March 26, 2025. ECF No. 322. Publishers filed a first amended complaint, on April 25, 2025. ECF No. 337. Anthropic subsequently moved to dismiss Counts II, III, and IV of the operative complaint, on May 9, 2025. The Court took the motion under submission without oral argument, on August 29, 2025. ECF No. 424.

II. LEGAL STANDARD

Under Federal Rule of Civil Procedure 12(b)(6), a court must dismiss a complaint if it fails to state a claim upon which relief can be granted. To avoid dismissal, the plaintiff must allege "enough facts to state a claim to relief that is plausible on its face." *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007). A claim is facially plausible when the pleaded facts allow the court "to draw the reasonable inference that the defendant is liable for the misconduct alleged." *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009).[2] For purposes of a Rule 12(b)(6) motion, the court generally "accept[s] factual allegations in the complaint as true and construe[s] the pleadings in the light most favorable to the nonmoving party." *Manzarek v. St. Paul Fire & Marine Ins. Co.*, 519 F.3d 1025, 1031 (9th Cir. 2008). However, the court need not "assume the truth of legal conclusions merely because they are cast in the form of factual allegations." *Fayer v. Vaughn*, 649 F.3d 1061,

[2] Publishers claim that motions to dismiss are "viewed with disfavor" and "rarely granted." Opp. at 4, ECF No. 376 (quoting *Ernst & Haas Mgmt. Co. v. Hiscox, Inc.*, 23 F.4th 1195 (9th Cir. 2022)). But the precedent *Ernst & Haas* cited for this proposition is a 1991 Ninth Circuit opinion, which predates *Twombly* and *Iqbal*. *See* 23 F.4th at 1199 (quoting *McDougal v. County of Imperial*, 942 F.2d 668, 676 n.7 (9th Cir. 1991)). "*Ernst & Haas* did not endeavor to explain why the 'viewed with disfavor'/'rarely granted' concept was still viable post-*Twombly*/*Iqbal*." *First Call Int'l, Inc. v. S&B Glob., Inc.*, No. 23-cv-10016, 2024 WL 4404429, at *1 n.3 (C.D. Cal. July 5, 2024). And neither do Publishers. Ultimately, the parties' disagreement regarding the relevant standard is immaterial to the Court's analysis because the Court holds that Publishers have satisfied the *Twombly*/*Iqbal* standard articulated above.

1064 (9th Cir. 2011) (per curiam) (quoting *W. Mining Council v. Watt*, 643 F.2d 618, 624 (9th Cir. 1981)).

III. DISCUSSION

For the second time, Anthropic seeks to dismiss Publishers' claims for contributory infringement, vicarious infringement, and removal of CMI in violation of the DMCA (Counts II through IV), arguing that Publishers have failed to state plausible claims. The Court addresses these claims in turn.

A. Contributory Infringement (Count II)

Contributory infringement liability applies where the defendant "(1) has knowledge of another's infringement and (2) either (a) materially contributes to or (b) induces that infringement."[3] *Perfect 10, Inc. v. Visa Int'l Serv. Ass'n*, 494 F.3d 788, 795 (9th Cir. 2007). In the context of online platforms, "a computer system operator can be held contributorily liable if it 'has *actual* knowledge that *specific* infringing material is available using its system' and can 'take simple measures to prevent further damage' to copyrighted works, yet continues to provide access to infringing works." *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1172 (9th Cir. 2007) (first quoting *A & M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1022 (9th Cir. 2001); then quoting *Religious Tech. Ctr. v. Netcom On-Line Commc'ns Servs., Inc.*, 907 F. Supp. 1361, 1375 (N.D. Cal.1995)). "[T]he existence of direct infringement is a necessary element of a claim for contributory infringement." *Petroliam Nasional Berhad v. GoDaddy.com, Inc.*, 897 F. Supp. 2d 856, 867 (N.D. Cal. 2012) (collecting cases).

Here, Anthropic argues Count II should be dismissed solely based on the knowledge element. Mot. at 4. Specifically, Anthropic contends that the operative complaint "includes no plausible factual allegation that Anthropic knew of or was willfully blind to any specific infringing

[3] Plaintiffs argue that they only need to plausibly allege that Anthropic "knew or ***had reason to know*** of specific third-party infringement." Opp. at 5-6 (collecting cases). However, as courts in this Circuit have noted, "the law 'remains unsettled' with respect to whether 'reason to know' is sufficient." *E.g.*, *Waterman v. TikTok, Inc.*, No. 24-cv-04802, 2024 WL 5413655, at *3 (C.D. Cal. Oct. 30, 2024) (citation omitted) (citing *Atari Interactive, Inc. v. Redbubble, Inc.*, 515 F. Supp. 3d 1089, 1114 (N.D. Cal. 2021)); *see also Erickson Prods., Inc. v. Kast*, 921 F.3d 822, 832 (9th Cir. 2019). Because Publishers have met their burden under the more demanding standard, the Court does not address whether the less-demanding "reason to know" standard is viable.

lyrics made available to Claude users, or of any third party who had prompted Claude to output infringing lyrics." *Id.* In response, Publishers raise five arguments regarding how their complaint plausibly alleges that Anthropic knew or had reason to know of the specific acts of infringement taking place via Claude. Opp. at 7. The Court only needs to address one of them in deciding this motion.

Publishers argue that the Court can infer that Anthropic had "actual or constructive knowledge" that "Claude users infringed specific lyrics" based on Anthropic's implementation and development of "guardrails." *Id.* at 10. First, Publishers allege that Anthropic employed pre-suit guardrails *because* it identified specific instances of Claude's output producing Publishers' lyrics. FAC ¶ 170 ("When Anthropic detected specific instances of infringement, it developed and adopted limited guardrails."). Second, Publishers claim that, in developing and improving its guardrails, Anthropic gathered "Claude user prompts and output data" that included such lyrics. FAC ¶ 121. Third, Publishers contend that, each time the guardrails were triggered by Claude user prompts containing these lyrics, "Anthropic became aware of those specific user prompts." *Id.* Fourth, Publishers assert that, when studying Claude user interactions, Anthropic evaluated both "specific efforts by users to avoid these guardrails" and situations where "the guardrails have failed and Claude . . . generated infringing output." *Id.* Based on these allegations, Anthropic had actual knowledge of specific acts of infringement by Claude users with respect to Publishers' lyrics. *See Amazon*, 508 F.3d at 1172. Accordingly, Publishers have pled a plausible claim with respect to Count II.

Anthropic maintains that these allegations are "not true." Reply at 7, ECF No. 381. At the motion to dismiss stage, the Court must accept the Publishers' non-conclusory allegations as true, although Anthropic's factual contentions may be revisited at a later stage. *Manzarek*, 519 F.3d at 1031. As described above, the allegations provide a plausible basis from which the Court can infer the requisite knowledge for purposes of Publishers' contributory infringement claim. Therefore, the Court denies the motion to dismiss as to Count II.

B. Vicarious Infringement (Count III)

"To state a claim for vicarious copyright infringement, a plaintiff must allege that the

defendant has (1) the right and ability to supervise the infringing conduct and (2) a direct financial interest in the infringing activity." *Visa*, 494 F.3d at 802 (citing *Ellison v. Robertson*, 357 F.3d 1072, 1078 (9th Cir. 2004)). In other words, a defendant "infringes vicariously by profiting from direct infringement while declining to exercise a right to stop or limit it." *Id.* (quoting *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 930 (2005)). "Like contributory liability, vicarious liability requires an underlying act of direct infringement." *Perfect 10, Inc. v. Yandex N.V.*, 962 F. Supp. 2d 1146, 1158 (N.D. Cal. 2013); *see also Luvdarts, LLC v. AT&T Mobility, LLC*, 710 F.3d 1068, 1071 (9th Cir. 2013).

Here, Anthropic's sole argument for dismissal is that Publishers fail to allege that Anthropic has a direct financial interest in Claude users' alleged infringement. The Court previously rejected this argument, Order at 8-9, but Anthropic reasserts it to "preserve the issue for further review," Mot. at 7. The Court adheres to its prior ruling. Contrary to Anthropic's framing, Publishers allege more than a financial interest in copyright infringement *generally*. *Id.* (citation omitted). Instead, Publishers allege that Anthropic directly benefits from the infringement of Publishers' works *specifically*, as required by the case law. *E.g., Perfect 10, Inc. v. Giganews, Inc.*, 847 F.3d 657, 673-74 (9th Cir. 2017).

Publishers allege that Anthropic "is paid every time . . . end users submit[] a request for Publishers' song lyrics, and it is paid again every time its Claude API generates output copying and relying on those lyrics." FAC ¶ 142. Moreover, it is plausible that the availability of Publishers' lyrics draws customers to use Claude because, as Publishers contend, Claude would not be as popular and valuable as it is but for "the substantial underlying text corpus that includes Publishers' copyrighted lyrics." *Id.* ¶ 146. Indeed, Publishers allege that "[c]ountless users . . . have prompted Claude for [their] lyrics." *Id.* ¶ 9, *see also id.* ¶ 90. These allegations are sufficient at this stage. Considering none of the other vicarious infringement elements are disputed, the Court denies Anthropic's motion to dismiss as to Count III.

C. Removal or Alteration of CMI (Count IV)

The DMCA "restricts the removal or alteration of [CMI] – information such as the title, the author, the copyright owner, the terms and conditions for use of the work, and other identifying

information set forth in a copyright notice or conveyed in connection with the work." *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 671 (9th Cir. 2018) (citing 17 U.S.C. § 1202(b)-(c)). To establish a violation of the DMCA under 17 U.S.C § 1202(b), the plaintiff must plausibly allege that defendant had the requisite scienter.

Here, Publishers assert they have plausibly alleged scienter under both Sections 1202(b)(1) and 1202(b)(3). Section 1202(b)(1) provides that, absent authorization, a person may not "intentionally remove or alter any" CMI, while Section 1202(b)(3) prohibits the distribution of "copies of works . . . knowing that [CMI] has been removed or altered." *Doe v. Github, Inc.*, 672 F. Supp. 3d 837, 857 (N.D. Cal. 2023) (citing 17 U.S.C. § 1202(b)). "Both provisions [also] require the defendant to possess the mental state of knowing, or having a reasonable basis to know, that his actions 'will induce, enable, facilitate, or conceal' infringement." *Stevens*, 899 F.3d at 673. Some courts have referred to this as the "double-scienter" requirement. *See, e.g.*, *Andersen v. Stability AI Ltd.*, 744 F. Supp. 3d 956, 970 (N.D. Cal. 2024); *Logan v. Meta Platforms, Inc.*, No. 22-cv-01847, 2023 WL 3668520, at *1, *6 (N.D. Cal. May 24, 2023).

Anthropic does not presently challenge the first scienter requirement under either Section 1202(b)(1) or 1202(b)(3) – *i.e.*, that Anthropic "intentionally remove[d] or alter[ed] CMI or distributed[] lyrics knowing CMI had been removed." Reply at 9. Rather, it disputes the second requirement by arguing that Publishers have failed to sufficiently allege that Anthropic knew or had a reasonable basis to know that its actions would conceal its infringement.[4] *See id.* The Court finds otherwise.

1. Section 1202(b)(1)

Accepting Publishers' allegations as true, the Court infers that Anthropic had a reasonable basis to know that its intentional removal of CMI from the datasets used to train Claude would conceal its alleged infringement.

Publishers allege that Anthropic knew Claude's output of Publishers lyrics would contain

[4] Publishers also argue that they have plausibly alleged that Anthropic knew, or had a reasonable basis to know, that its intentional removal of CMI, and distribution of Publishers' lyrics without CMI, would conceal Claude users' infringement. Opp. at 19. But the Court finds it unnecessary to reach this argument in deciding the DMCA claim.

CMI unless it engaged in a process to remove the CMI from its training datasets. FAC ¶¶ 67-72, 74. Anthropic's cofounders elected to use a tool, Newspaper, that effectively removed CMI from these datasets – and they chose this tool specifically because it was more effective at removing CMI compared to another tool, jusText. *Id.* ¶¶ 72-73 (asserting that, when Anthropic applied jusText to "a scraped webpage, containing footnotes, a copyright owner name, and '© 2019' copyright notice," this algorithm did not remove any of this information but, "[i]n contrast, Newspaper, which removed the footnotes, copyright owner name, and copyright notice entirely, was considered 'a significant improvement.'"). At least some of these cofounders were aware that, during training, large language models tend to memorize "A LOT" of data and then "regurgitate" it. *Id.* ¶¶ 108-09 & n.29 (quoting Tom B. Brown *et al.*, *Language Models are Few-Shot Learners* 8-9, ARXIV (July 22, 2020), https://arxiv.org/pdf/2005.14165 [https://perma.cc/25RZ-B8VB]). Taking these allegations together (as true), the Court can plausibly infer that Anthropic had a reasonable basis to know that the dataset curation and training processes it engaged in, whereby it removed CMI (*i.e.*, its intentional removal of CMI), concealed its own infringement.

Anthropic contends that Publishers' allegations regarding the extractor algorithms "speak at most to Anthropic's desire *to* remove CMI but not at all as to *why* Anthropic allegedly did this." Mot. at 8. It urges the Court to follow *Tremblay v. OpenAI, Inc.*, 716 F. Supp. 3d 772 (N.D. Cal. 2024). Reply at 11-12. *Tremblay* is distinguishable because the allegations in that case were conclusory, and the court analyzed whether the defendant knowingly *enabled* infringement by end users of the generative AI model. *Id.* at 779. But here, Publishers allege that Anthropic *concealed* infringement. And the detailed allegations described above raise an inference that Anthropic had reasonable grounds to know its training datasets would conceal its infringement – based on its cofounders understanding of how large language models are trained and their selection of the Newspaper algorithm over the jusText algorithm. *Kadrey v. Meta Platforms, Inc.* is instructive. No. 23-cv-03417, 2025 WL 744032 (N.D. Cal. Mar. 7, 2025). In that case, the court found that plaintiffs sufficiently alleged that defendant "intentionally removed CMI to conceal copyright infringement." *Id.* at *2. It reached this conclusion based on plaintiffs' allegations that defendant (1) was aware its AI system was particularly prone to "memorizing and generating outputs of CMI

unless CMI was removed from its training data"; and (2) took steps to reduce the chances that its program would produce CMI. *Id.* Publishers have made similar allegations here. Accordingly, Publishers have met their burden under § 1202(b)(1).

2. Section 1202(b)(3)

In its previous motion to dismiss Order, the Court concluded that "Publishers plausibly allege that the [Claude] output omits the CMI regarding the Works." Order at 10. And Anthropic does not assert otherwise this time around. Thus, it is undisputed, for purposes of this motion, that Anthropic distributed Publishers' lyrics without including CMI. The critical question is whether Anthropic did so while knowing, or having a reasonable basis to know, that it would conceal its own infringement. The Court finds Publishers have plausibly alleged that Anthropic had the requisite scienter.

As discussed above, accepting Publishers' allegations as true, Anthropic decided to use the Newspaper algorithm to extract CMI from the datasets used to train Claude, including "Common Crawl" and "The Pile," even though it had a choice to use, jusText, which would have retained more, if not all, of the CMI. FAC ¶¶ 72-73. And some of Anthropic's cofounders understood the tendency of large language models to memorize large amounts of data and regurgitate it. *Id.* ¶¶ 108-09 (citation omitted). Because Anthropic allegedly copied the datasets discussed above, to curate a dataset to train Claude, and removed the CMI therefrom, the Court can plausibly infer that Anthropic did so "knowing that removing the CMI would help conceal the alleged infringement." *Batra v. PopSugar, Inc.*, No. 18-cv-03752, 2019 WL 482492, at *2 (N.D. Cal. Feb. 7, 2019).

Notwithstanding Anthropic's insistence that the Court follow *Tremblay*, this case is once again distinguishable. Reply at 11-12. In *Tremblay*, the court was analyzing whether plaintiffs' 1202(b)(3) claim was plausibly alleged in connection with third-party user infringing activity. *See* 716 F. Supp. 3d at 779-780. Whereas, in this case, Publishers allege that Anthropic had a reasonable basis to know that it was concealing its own infringement. *See Mango v. BuzzFeed, Inc.*, 970 F.3d 167, 173 (2d Cir. 2020) (distinguishing situation where "defendant's own infringement satisfies Section 1202(b)(3)'s second scienter requirement," from scenario where plaintiff alleges "constructive knowledge of likely future, third-party infringement."). Likewise,

in *Tremblay*, plaintiffs did "not allege[] that Defendants distributed their books or copies of their books," nor "that ChatGPT reproduces Plaintiffs copyrighted works without CMI." 716 F. Supp. 3d at 780. That is not the case here, given Publishers allege that Anthropic "distributes Publishers' lyrics and copies of those lyrics knowing that [CMI] has been removed." *See, e.g., id.* ¶¶ 128, 131. They also allege that "Anthropic wanted to train its Claude AI models specifically on the content of Publishers' lyrics, so that the models' output would reproduce that expressive content," without CMI. *Id.* ¶ 74. Thus, Anthropic's reliance on *Tremblay* is misplaced and Publishers have met their burden under § 1202(b)(3).

Based on the above, Publishers have plausibly alleged the scienter required under § 1202(b), for purposes of its DMCA claim. As a result, the Court denies Anthropic's motion to dismiss as to Count IV.

IV. CONCLUSION

For the foregoing reasons, the Court DENIES Defendant's motion to dismiss the operative complaint. Defendant is ORDERED to file its answer within 14 days after the Court rules on Plaintiffs' motion for leave to file a second amended complaint. ECF No. 411.

IT IS SO ORDERED.

Dated: October 6, 2025

Eumi K. Lee
United States District Judge

United States District Court
Northern District of California