UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

PAUL TREMBLAY, et al.,

Plaintiffs,

v.

OPENAI, INC., et al.,

Defendants.

Case Nos. 23-cv-03223-AMO

23-cv-03416-AMO

ORDER GRANTING IN PART AND DENYING IN PART THE MOTIONS TO DISMISS

This is a putative class action copyright case. OpenAI, Inc.'s motions to dismiss were heard before this Court on December 7, 2023. Having read the papers filed by the parties and carefully considered their arguments therein and those made at the hearing, as well as the relevant legal authority, the Court hereby **GRANTS** in part and **DENIES** in part the motions to dismiss for the following reasons.

I. BACKGROUND

Before the Court are two nearly identical putative class complaints in *Tremblay et al v. OpenAI, Inc. et al*, 23-cv-3223 and *Silverman et al v. OpenAI, Inc. et al*, 23-cv-3416. Plaintiffs are authors of books who allege that their books were used to train OpenAI language models that operate the artificial intelligence ("AI") software ChatGPT.[1] Silverman Compl. ¶¶ 1-4; Tremblay Compl. ¶¶ 1-4. Plaintiffs Paul Tremblay, Sarah Silverman, Christopher Golden, and Richard Kadrey (collectively, "Plaintiffs") hold registered copyrights in their books. Tremblay Compl. ¶¶

[1] For the purposes of the motion to dismiss at bar, the Court accepts all factual allegations in the Complaint as true and construes the pleadings in the light most favorable to the Plaintiffs. *Manzarek v. St. Paul Fire & Marine Ins. Co.*, 519 F.3d 1025, 1031 (9th Cir. 2008).

10, 12, Ex. A (*The Cabin at the End of the World* (Tremblay))[2]; Silverman Compl. ¶¶ 10-13, Ex. A (*The Bedwetter* (Silverman); *Ararat* (Golden), and *Sandman Slim* (Kadrey)).

Defendant OpenAI[3] creates and sells certain AI software known as large language models (or "LLM"). Tremblay Compl. ¶ 23. These language models are "trained" by inputting large amounts of texts known as the "training dataset." *Id.* The language models copy text from the training dataset and extract "expressive information." *Id.* ChatGPT is an OpenAI language model that allows paying users to enter text prompts to which ChatGPT will respond and "simulate human reasoning," including answering questions or summarizing books. *Id.* ¶¶ 22, 36-38. ChatGPT generates its output based on "patterns and connections" from the training data. *Id.* ¶ 39.

OpenAI copied Plaintiffs' copyrighted books and used them in its training dataset. *Id.* ¶ 24. When prompted to summarize books written by each of the Plaintiffs, ChatGPT generated accurate summaries of the books' content and themes. *Id.* ¶ 41 (citing Ex. B); Silverman Compl. ¶ 42 (citing Ex. B).

Plaintiffs seek to represent a class of all people in the U.S. who own a copyright in any work that was used as training data for OpenAI language models during the class period. Tremblay Compl. ¶ 42; Silverman Compl. ¶ 43. Plaintiffs assert six causes of action against various OpenAI entities: (1) direct copyright infringement (Count I); (2) vicarious infringement (Count II); (3) violation of Section 1202(b) of the Digital Millennium Copyright Act ("DMCA") (Count III); (4) unfair competition under Cal. Bus. & Prof. Code Section 17200 (Count IV); (5) negligence (Count V); and (6) unjust enrichment (Count VI).

OpenAI filed the instant motions to dismiss on August 28, 2023, seeking dismissal of Counts II through VI. ECF 33 ("Motion").[4]

---

[2] Plaintiff Mona Awad voluntarily dismissed her claims without prejudice. ECF 29.

[3] Defendants are seven entities that Plaintiffs collectively refer to as "OpenAI." ECF 33 ("Motion") at 14 (citing Tremblay Compl. ¶¶ 13-19; Silverman Compl. ¶¶ 14-20). The Court follows this naming convention.

[4] Defendants' motion to dismiss addresses both the Silverman and the Tremblay complaints and was filed concurrently on both dockets. Motion at 1 n1.

II. LEGAL STANDARD

Federal Rule of Civil Procedure 8(a) requires that a complaint contain "a short and plain statement of the claim showing that the pleader is entitled to relief." A defendant may move to dismiss a complaint for failing to state a claim upon which relief can be granted under Federal Rule of Civil Procedure 12(b)(6). "Dismissal under Rule 12(b)(6) is appropriate only where the complaint lacks a cognizable legal theory or sufficient facts to support a cognizable legal theory." *Mendiondo v. Centinela Hosp. Med. Ctr.*, 521 F.3d 1097, 1104 (9th Cir. 2008). To survive a Rule 12(b)(6) motion, a plaintiff must plead "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 a plaintiff pleads "factual content that allows 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).

In reviewing the plausibility of a complaint, courts "accept factual allegations in the complaint as true and construe the pleadings in the light most favorable to the nonmoving party." *Manzarek*, 519 F.3d at 1031. Nonetheless, courts do not "accept as true allegations that are merely conclusory, unwarranted deductions of fact, or unreasonable inferences." *In re Gilead Scis. Secs. Litig.*, 536 F.3d 1049, 1055 (9th Cir. 2008) (citation omitted).

III. ANALYSIS

Defendants seek dismissal of all causes of action except for the claim for direct copyright infringement. Motion at 18. Defendants seek dismissal of Count II for vicarious copyright infringement; Count III for violation of Section 1202(b) of the Digital Millennium Copyright Act ("DMCA); Count IV for Unfair Competition under Cal. Bus. & Prof. Code § 17200; Count V for negligence; and Count VI for unjust enrichment. *Id.* The Court addresses each in turn.

A. Vicarious Copyright Infringement (Count II)

The Copyright Act grants the copyright holder exclusive rights to (1) "reproduce the copyrighted work in copies," (2) "prepare derivate works," and (3) "distribute copies . . . of the copyrighted work to the public." 17 U.S.C. § 106(1)-(3). Copyright protection does not extend to "every idea, theory, and fact" underlying a copyrighted work. *Eldred v. Ashcroft*, 537 U.S. 186,

mean that every element of the work may be protected." *Corbello v. Valli*, 974 F.3d 965, 973 (9th Cir. 2020) (citation omitted).

Copyright infringement requires that a plaintiff show (1) "he owns as valid copyright" and (2) the defendant "copied aspects of his work." *Corbello*, 974 F.3d at 973. The second prong "contains two separate components: 'copying' and 'unlawful appropriation.'" *Id.* at 974. "Copying can be demonstrated either through direct evidence or by showing that the defendant had access to the plaintiff's work and that the two works share similarities probative of copying, while the hallmark of 'unlawful appropriation' is that the works share substantial similarities." *Id.* (citations and internal quotations omitted); *see Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin*, 952 F.3d 1051, 1064 (9th Cir. 2020) ("the hallmark of 'unlawful appropriation' is that the works share *substantial* similarities") (emphasis in original).

A claim of vicarious infringement requires a threshold showing of direct infringement, *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1169 (9th Cir. 2007) ("*Amazon.com*"); *see A&M Recs., Inc. v. Napster, Inc.*, 239 F.3d 1004, 1013 n.2 (9th Cir. 2001) *aff'd*, 284 F.3d 1091 (2002) ("Secondary liability for copyright infringement does not exist in the absence of direct infringement by a third party"). A plaintiff must then show that "the defendant has (1) the right and ability to supervise the infringing conduct and (2) a direct financial interest in the infringing activity." *Perfect 10, Inc. v. Giganews, Inc.*, 847 F.3d 657, 673 (9th Cir. 2017) ("*Giganews*") (citation omitted).

Defendants argue that the vicarious infringement claim fails for three reasons: (1) Plaintiffs have not alleged that direct infringement occurred; (2) Plaintiffs have not alleged that Defendants had the "right and ability to supervise"; and (3) Plaintiffs have not alleged "direct financial interest." Motion at 19-21. The Court first considers whether Plaintiffs have adequately alleged direct infringement and, ultimately, does not reach Defendants' latter two arguments.

Plaintiffs suggest that they do not need to allege a "substantial similarity" because they have evidence of "direct copying." ECF 48 ("Response") at 15. They argue that because Defendants directly copied the copyrighted books to train the language models, Plaintiffs need not

F.3d 1148, 1154 (9th Cir. 2012) (explaining that "substantial similarity" helps determine whether copying occurred "when an allegedly infringing work appropriates elements of an original without reproducing it *in toto*."). Plaintiffs misunderstand *Range Rd.* There, the court did not need to find substantial similarity because the infringement was the public performance of copyrighted songs at a bar. *Range Rd.*, 668 F.3d at 1151-52, 1154. Since the plaintiffs provided unrebutted evidence that the performed songs were the protected songs, they did not need to show that they were substantially similar. *Id.* at 1154. Distinctly, Plaintiffs here have not alleged that the ChatGPT outputs contain direct copies of the copyrighted books. Because they fail to allege direct copying, they must show a substantial similarity between the outputs and the copyrighted materials. *See Skidmore*, 952 F.3d at 1064; *Corbello*, 974 F.3d at 973-74.

Plaintiffs' allegation that "every output of the OpenAI Language Models is an infringing derivative work" is insufficient. Tremblay Compl. ¶ 59; Silverman Compl. ¶ 60. Plaintiffs fail to explain what the outputs entail or allege that any particular output is substantially similar – or similar at all – to their books. Accordingly, the Court dismisses the vicarious copyright infringement claim with leave to amend.

B. Section 1202(b) of the DMCA (Count III)

In addition to protecting against vicarious and direct infringement, "[c]opyright law restricts the removal or alteration of copyright management information ('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). Section 1202(b) of the Digital Millennium Copyright Act (DMCA) provides that one cannot, without authority, (1) "intentionally remove or alter any" CMI, (2) "distribute . . . [CMI] knowing that the [CMI] has been removed or altered," or (3) "distribute . . . copies of works . . . knowing that [CMI] has been removed or altered." 17 U.S.C. § 1202(b). To state a Section 1202 claim for removal or alteration of CMI, plaintiffs must first identify "what the removed or altered CMI was." *Free Speech Sys., LLC v. Menzel*, 390 F. Supp. 3d 1162, 1175 (N.D. Cal. 2019). Plaintiffs must also show the

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.