1

2

3

4   UNITED STATES DISTRICT COURT

5   NORTHERN DISTRICT OF CALIFORNIA

6   EUREKA DIVISION

7

8   PAUL TREMBLAY, et al.,   Case No. 23-cv-03223-AMO   (RMI)

9   Plaintiffs,

10   v.   **ORDER RE: FOURTH DISCOVERY
                                             DISPUTE**
11   OPENAI, INC., et al.,   Re: Dkt. No. 153

12   Defendants.

13

14   Now pending before the court is a jointly-filed letter brief setting forth a discovery dispute

15   through which Defendants seek to compel certain discovery over which Plaintiffs have asserted a

16   work product privilege. *See* Ltr. Br. (dkt. 153) at 1-5. Pursuant to Federal Rule of Civil Procedure

17   78(b) and Civil Local Rule 7-1(b), the court finds the matter suitable for disposition without oral

18   argument. For the reasons stated below, Defendants' request to compel the material in question is

19   granted.

20   By way of background, Plaintiffs (a group of authors) allege that Defendants' ChatGPT

21   software relies on a large language model by which it is trained through "copying massive

22   amounts of text and extracting expressive information from it," and that "[o]nce the large language

23   model has copied and ingested the text in its training dataset, it is able to emit convincingly

24   naturalistic text outputs in response to user prompts." *See* First Amend. Compl. ("FAC") (dkt.

25   120) at ¶ 2. Plaintiffs further allege that "when ChatGPT is prompted, ChatGPT generates

26   summaries of Plaintiffs' copyrighted works – something only possible if ChatGPT was trained on

27   Plaintiffs' copyrighted works." *Id*. at ¶ 5. More specifically, Plaintiffs' have alleged that "[w]hen

28   ChatGPT was prompted to summarize books written by each of the Plaintiffs, it generated very

accurate summaries. These summaries are attached [to the FAC] as Exhibit B. The summaries get some details wrong, which is expected, since a large language model mixes together expressive material derived from many sources. Still, the rest of the summaries are accurate, which means that ChatGPT retains knowledge of particular works in the training dataset and is able to output similar textual content." *Id*. at ¶ 51. Further, Exhibit-B to the FAC sets forth several prompts that ask Chat GPT to summarize in detail various parts of Plaintiffs' writings. *See* FAC *Exh. B* (dkt. 120-2) at 2-17, 20, 23, 26-27, 29, 35-38. The Exhibit also sets forth numerous queries to ChatGPT – along with their responses – about a number of Plaintiffs' works; the queries resemble questions and answers that one might encounter in a literature class (*e.g.*, What are the main themes of this work? What are examples of nature and beauty in this work? What are some examples of the immigrant experience in this work?). *See id*. at 18, 21, 24, 30, 33, 39. Additionally, the Exhibit also sets forth the queries and responses on a number of occasions where ChatGPT was asked to write a paragraph, or to compose a screenplay, either in the style of one of the Plaintiffs or "like" one of their works. *See id*. at 19, 22, 25, 28, 31-32, 34, 40.

The current dispute concerns Defendants' RFP 9, which seeks "[a]ll non-privileged Documents and Communications relating to [Plaintiffs'] investigation of the claims alleged in the Complaint." Ltr. Br. (dkt. 153) at 1. Defendants then narrowed this request to encompass: "(a) the OpenAI account information for individuals who used ChatGPT to investigate Plaintiffs' claims; and (b) the prompts and outputs for Plaintiffs' testing of ChatGPT in connection with their pre-suit ChatGPT testing, including prompts and outputs that *did not* reproduce or summarize Plaintiffs' works or otherwise support Plaintiffs' claims, along with documentation of Plaintiffs' testing process." *Id*. (emphasis added). Defendants maintain that Plaintiffs have "refuse[d] to respond in full based on a claim of work product protection, offering to produce only 'full threads of the prompts and outputs' that led to the examples in Exhibit-B to the Complaint." *Id*. In short, Defendants contend that Plaintiffs have "offered up only their preferred, cherry-picked results" by refusing to tender the prompts and results that *did not* improperly reproduce or summarize Plaintiffs' works. *Id*.

Defendants argue that Plaintiffs waived work product protection over the account and

2

prompt-and-output information for three reasons: (1) because Plaintiffs revealed or placed their work product at issue during the course of the litigation by including allegations regarding how ChatGPT responded to Plaintiffs' prompts made from those accounts in the FAC and in Exhibit-B thereto; (2) because Plaintiffs voluntarily disclosed the information in question to their adversary in litigation; and, (3) because the account information and the totality of the prompt-and-response data that Plaintiffs used to interrogate ChatGPT should, in fairness, be considered together with Plaintiffs' testing results (set forth in Exhibit-B) and that the totality of that data would be needed for Defendants to subject Plaintiffs' claims to meaningful adversarial testing. *See id*. at 1-3. Specifically, as to the account data component of the information that Defendants seek, Defendants submit that because "the 'custom instructions' feature allows users to 'add preferences or requirements' for 'ChatGPT to consider when generating its responses,'" Defendants "need[] this discovery to test Plaintiffs' allegations regarding ChatGPT's behavior in response to the 'interrogation' Plaintiffs chose to put at issue." *Id*. at 3.

Plaintiffs assert that this material is shielded from discovery because it is attorney work product – Plaintiffs add that Defendants "seek[] wide-ranging discovery into Plaintiffs' counsels' investigatory files, regardless of whether the prompts and outputs were used in the complaint, and regardless of whether the prompts and outputs 'support Plaintiffs' claims.'" *Id*. at 3. Plaintiffs submit that they have agreed to tender prompts and outputs that support their claims and were set forth in the FAC and in Exhibit-B thereto. *Id*. at 5. However, the "prompts and outputs that did not reproduce or summarize Plaintiffs' works or otherwise support Plaintiffs' claims" are shielded from discovery because to produce those would be to divulge the "thoughts and analysis conducted by Plaintiffs' lawyers in preparation of the litigation." *Id*. at 3. Plaintiffs, therefore, consider the prompts and responses that did not improperly reproduce or summarize their works to be opinion work product. *Id*. at 4.

As to Defendants' waiver argument, Plaintiffs contend that the disclosure of some of their prompts and responses did not operate as a waiver regarding the undisclosed prompts and responses because those disclosures (in the FAC and Exhibit-B) were necessary in order to satisfy the plausibility requirement for their claims. *Id*. Plaintiffs also claim that the prompts and

United States District Court
Northern District of California

responses set forth in the FAC and in Exhibit-B "are not evidence which the jury will be asked to rely on . . . [as] Plaintiffs will rely on other, more concrete evidence obtained through discovery to identify the literary works OpenAI used to train their products." *Id*. The way Plaintiffs see it, "the fact that ChatGPT did not always produce a summary of Plaintiffs' work when prompted is immaterial because the summaries were simply used to plausibly allege that Defendants trained their products on Plaintiffs' Asserted Works[] [and] '[n]egative' test results are not relevant for the same reason the 'positive' test results are not particularly relevant at this stage." *Id*. At bottom, Plaintiffs state that Defendants seek to obtain "a free ride on the work of [their] attorney," which they urge the court to reject due to the suggestion that "this information is not relevant[] [a]nd OpenAI can obtain the material itself – [because] it can interrogate ChatGPT for itself." *Id*.

The Federal Rules of Civil Procedure define work product as "documents and tangible things that are prepared in anticipation of litigation or for trial by or for another party or its representative . . ." Fed. R. Civ. P. 26(b)(3)(A); *see also United States v. Richey*, 632 F.3d 559, 567-68 (9th Cir. 2011). Although work product is generally exempt from disclosure in discovery, a party can still obtain work product of an opposing party if the requesting party shows a "substantial need for the materials to prepare its case and cannot, without undue hardship, obtain their substantial equivalent by other means." Fed. R. Civ. P. 26(b)(3)(A).

There are two categories of work product. As to the first, fact work product is discoverable upon a showing of "substantial need" and "undue hardship." Fed. R. Civ. P. 26(b)(3). Substantial need consists of the relative importance of the information in the documents to the party's case and the ability to obtain that information from other means, and "[t]he test is not only relevancy but that there are no other available means by which to secure the same information." *AT&T Corp. v. Microsoft Corp.*, 2003 U.S. Dist. LEXIS 8710, 2003 WL 21212614, at *6 (N.D. Cal. Apr. 18, 2003). "Undue hardship is demonstrable if witnesses are unavailable or cannot recall the events in question." *Id*. On the other hand, opinion work product is defined as material prepared by an attorney that contains "mental impressions, conclusions, opinions, or legal theories of a party's attorney or other representative concerning the litigation." Fed. R. Civ. P. 26(b)(3)(B). In other words, opinion work product consists of the attorney's interpretation of legal theories and the

application of facts to those theories, rather than the bare facts or legal theories alone. While it is sometimes said that opinion work product is "virtually undiscoverable" (*see e.g.,* R*epublic of Ecuador v. Mackay*, 742 F.3d 860, 869 n.3 (9th Cir. 2014)), even opinion work product is not totally immune from discovery when an attorneys' mental impressions are at issue in a case and the need for the material is compelling. *Holmgren v. State Farm Mut. Auto. Ins. Co.*, 976 F.2d 573, 577 (9th Cir. 1992). Thus, "[t]he privilege derived from the work product doctrine is not absolute[] [and] [l]ike other qualified privileges, it may be waived." *United States v. Nobles*, 422 U.S. 225, 236 (1975). Similar to the waiver of the attorney-client privilege, a litigant can waive work product protection to the extent that the litigant reveals, <u>or places the work product at issue</u>, during the course of litigation. *United States v. Sanmina Corp. & Subsidiaries*, 968 F.3d 1107, 1119 (9th Cir. 2020).

For instance, in *Nobles*, 422 U.S. at 237-38, a criminal defendant's decision to present his defense investigator as a witness waived work-product privilege over the investigator's report "with respect to matters covered in his testimony." *Id*. at 236, see also *id*. at 239 ("[R]espondent sought to adduce the testimony of the investigator and contrast his recollection of the contested statements with that of the prosecution's witnesses. Respondent, by electing to present the investigator as a witness, waived the privilege with respect to matters covered in his testimony. Respondent can no more advance the work-product doctrine to sustain a unilateral testimonial use of work-product materials than he could elect to testify in his own behalf and thereafter assert his Fifth Amendment privilege to resist cross-examination on matters reasonably related to those brought out in direct examination.").

The court will first note that the account settings and the negative test results are more in the nature of fact work product than opinion work product. While Plaintiffs suggest that revealing these facts would necessarily reveal "the thoughts and analysis conducted by Plaintiffs' lawyers in preparation of the litigation" (*see* Ltr. Br. (dkt. 153) at 3), the court finds this suggestion to be exaggerated and unpersuasive. As stated above, opinion work product consists of the attorney's interpretation of legal theories and the application of facts to those theories, rather than the bare facts or legal theories alone. The account settings and negative test result sought here are more in

the nature of bare facts. However, even assuming *arguendo* that to reveal these bare facts would provide a revelatory glimpse into Plaintiffs' counsels' thought process such as to reveal counsels' interpretation of legal theories and the application of some facts to those theories, Plaintiffs cannot avoid the notion that by placing a large subset of these facts in the FAC and in Exhibit-B, Plaintiffs have waived the ability to assert work product protection. As was the case with the investigator's report in *Nobles*, 422 U.S. at 237-38 – here too, Plaintiffs' negative results and the account settings in question are "reasonably related" to the positive results that Plaintiffs have placed at issue and included in the FAC and in Exhibit-B. Thus, an evaluation of those account settings and negative test results is a necessary component of Defendants' ability to understand Plaintiffs' positive results and to fairly subject those results to meaningful scrutiny.

It is therefore no answer for Plaintiffs to simply declare that the negative results are "not relevant" simply because Plaintiffs did not include them in the FAC. As stated, Defendants' review of the negative test results, as well as the account setting used to interrogate ChatGPT, appears likely to aid in Defendants' understanding of Plaintiffs' positive results as well as Defendants' ability to subject those positive test results to scrutiny. Or, to put it another way, the court disagrees with Plaintiffs' contention that the account settings and the prompts and outputs that do not "support Plaintiffs' claims" (*see* Ltr. Br. (dkt. 153) at 3) are not "particularly relevant" (*see id*. at 2) because while they may not be relevant to Plaintiffs' pursuit of their claims, they are plainly relevant to Defendants' defenses. Of course, the reason is that Plaintiffs have placed a subset of their test results at issue in this case; and, having done so, Defendants should be entitled to examine the entirety of Plaintiffs' test results (including the negative results), as well as the account settings used in interrogating ChatGPT.

Lastly, the court is unpersuaded by Plaintiffs' argument that Defendants can simply interrogate ChatGPT themselves. *See id*. at 5. Without knowing the account settings used by Plaintiffs to generate their positive and negative results, and without knowing the exact formulation of the prompts used to generate Plaintiffs' negative results, Defendants would be unable to replicate the same results. *See e.g., Mushroom Assoc. v. Monterey Mushrooms, Inc.*, 1992 U.S. Dist. LEXIS 19664, *13-14 (N.D. Cal. May 19, 1992) ("Also, the discovering party

cannot obtain substantially equivalent materials because the specific work product is directly at issue. For instance, although the discovering party could obtain its own test results concerning a patent, the tests results known to counsel at the time advice is given are what is relevant when a party asserts the advice of counsel defense."). Here, Plaintiffs' FAC and Exhibit-B contain a series of test results, and while Defendants can certainly produce some prompts and responses themselves, doing so will not be useful in understanding or scrutinizing Plaintiffs' test results. On the other hand, discovery into the account settings used by Plaintiffs, and discovery into Plaintiffs' negative test results will achieve that end.

For the reasons stated herein, as well as those argued by Defendants (*see* Ltr. Br. (dkt. 153) at 1-3), the request to compel the OpenAI account information for the individuals who used ChatGPT to investigate Plaintiffs' claims, and the prompts and outputs for Plaintiffs' testing of ChatGPT in connection with their pre-suit ChatGPT testing, including prompts and outputs that *did not* reproduce or summarize Plaintiffs' works or otherwise support Plaintiffs' claims, along with documentation of Plaintiffs' testing process, is **GRANTED**.

IT IS SO ORDERED.

Dated: June 24, 2024

ROBERT M. ILLMAN
United States Magistrate Judge