

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
EUREKA DIVISION

In re OpenAI ChatGPT Litigation

Case No. 23-cv-03223-AMO (RMI)

ORDER RE: DISCOVERY DISPUTE

Re: Dkt. No. 136

United States District Court
Northern District of California

Now pending before the court is a jointly-filed letter brief setting forth a dispute regarding Plaintiffs’ Interrogatory (“ROG”) Nos. 12 and 14. (*see* dkt. 136). The court concludes that pursuant to Civil Local R. 7-1(b), these matters are suitable for disposition without oral argument.

Plaintiffs, a group of authors who hold registered copyrights in their works, bring this putative class action against a group of defendants that they collectively refer to OpenAI. *See* First Amend. Consol. Compl. (“FAC”) (dkt. 120) at 2. ChatGPT is a software product created, maintained, and sold by OpenAI, and is powered by two AI software programs called GPT-3.5 and GPT-4, which are also known as “large language models.” *Id.* Plaintiffs allege that “[r]ather than being programmed in the traditional way, a large language model is ‘trained’ by copying massive amounts of text and extracting expressive information from it. This body of text is called the training dataset. Once a large language model has copied and ingested the text in its training dataset, it is able to emit convincingly naturalistic text outputs in response to user prompts.” *Id.* In essence, Plaintiffs contend that because a large language model’s output is reliant on the material in its training dataset, “[e]very time it assembles a text output [in response to user queries], the model relies on the information it extracted from its training dataset,” which results in ChatGPT

1 sometimes generating summaries of Plaintiffs' copyrighted works and benefiting commercially
2 from the use of Plaintiffs' and Class members' copyrighted works. *Id.* at 2-3. In light of which,
3 Plaintiffs have brought claims for direct copyright infringement and unfair competition. *Id.* at 12-
4 14.

5 Through ROG No. 12, Plaintiffs contend that they seek "basic information with respect to
6 the identity of social-media usernames of certain current and former employees of Defendants
7 who used personal social-media accounts to communicate on the subjects of the litigation." *See*
8 Ltr. Br. (dkt. 136) at 1. Plaintiffs have narrowed ROG No. 12 and request an order directing
9 Defendants as follows: (1) as to *current* employees or board members listed in response to any
10 interrogatory, Defendants will inquire as to whether any such person has used any of their personal
11 social media accounts to discuss anything relevant to this litigation, and, if so, Defendants will
12 produce those individuals' social media usernames; and, (2) as to *past* employees or board
13 members, Defendants will produce the social media usernames of those individuals to the extent
14 they are known to Defendants. *Id.* Plaintiffs explain that Defendants' directors and personnel may
15 operate some of their personal social media accounts pseudonymously, making it difficult or
16 impossible for Plaintiffs to uncover discussions relevant to this action having taken place through
17 such accounts by other means. *Id.* at 2. Plaintiff note, for example, that Elon Musk, a co-founder
18 and early investor in OpenAI, has admitted in a recent deposition in a separate action that he
19 operates two pseudonymous Twitter accounts. *Id.*

20 Defendants contend that this "[i]sn't a narrow request," and that "OpenAI does not collect
21 from its employees and Board Members information about personal social media accounts, or
22 monitor those accounts in the ordinary course of business." *Id.* at 2-3. Thus, Defendants contend
23 that "Plaintiffs' request goes far beyond what OpenAI has, and none of the requested accounts are
24 relevant in any event." *Id.* at 3. In essence, Defendants state that they do not have the requested
25 information in their "possession or custody" because "the company does not systematically collect
26 from its employees and Board members information about personal social media accounts, or
27 monitor those accounts in the ordinary course of business." *Id.* Defendants add that neither is this
28 information within their "control," because "Plaintiffs have identified no legal basis for OpenAI to

1 demand its employees and Board members turn over [username] information about their personal
2 [social media] accounts.” *Id.* Defendants then state that “Plaintiffs’ cited cases are of no help
3 [because] [i]n *Roblox Corp. v. WowWowee Grp. Ltd.*, [2023 WL 8374741, at *2 (N.D. Cal. Dec. 4,
4 2023)], the court ordered the identification of personal accounts only after the defendant admitted
5 at least one such account was relevant[,] [and] [h]ere, OpenAI disputes relevance as to all the
6 personal accounts.” *Id.*

7 Having considered the Parties’ arguments, and having reviewed the authorities cited,
8 Plaintiffs’ request is **GRANTED in part**, and the court fashions the following ruling to balance
9 each Party’s rights. Defendants are **ORDERED** to promptly investigate and inquire from *current*
10 directors and employees whether they have engaged in any discussions, the contents of which
11 might be relevant to the claims or defenses involved in this case, using their personal social media
12 accounts. The court finds that the burden associated with undertaking this inquiry is minimal and
13 that it is proportionate and responsive to the needs of the case. If all current directors and
14 employees report that they have engaged in no such discussions on their social media accounts,
15 Defendants are **ORDERED** to certify that fact to Plaintiffs, which will put the matter to rest. If, on
16 the other hand, any current director or employee answers that inquiry in the affirmative,
17 Defendants are **ORDERED** to gather and disclose that person’s relevant social media username(s)
18 forthwith. As to *past* employees, Defendants are **ORDERED** to produce the social media
19 usernames of any such persons if Defendants know, or learn, that any that such persons have
20 engaged in discussions on social media that might be relevant to claims or defenses in this case,
21 and the social media username(s) of such persons are known to Defendants.

22 Through ROG No. 14, Plaintiffs seeks information about individuals and entities who
23 possess or have possessed stock or ownership interests in OpenAI greater than five percent. *See*
24 Ltr. Br. (dkt. 136) at 4. In essence, Plaintiffs seek the identity of these persons or entities because
25 such “individuals or entities may have discoverable documents or information relevant to the
26 claims in this action,” and because “such individuals may have sought to exert influence and/or
27 voice concerns vis-à-vis decisions by Defendants.” *Id.* at 4-5 (emphasis added). Plaintiffs also
28 contend that information regarding “Defendants’ ownership structure will allow Plaintiffs to

United States District Court
Northern District of California

1 understand the Defendants’ corporate relationship and structure of relationship . . . [and] will
2 address the financial condition of the Defendants, which is relevant to the ability of Defendants to
3 respond to a judgment.” *Id.* 5.

4 Plaintiffs’ portion of the letter brief does not indicate any concrete basis on which the court
5 could conclude that entities or persons owning more than 5% of the shares of OpenAI actually
6 *would* have (rather than *could* have) any relevant documents or information, or that they *actually*
7 (rather than *might have*) sought to exert influence or voice concerns about OpenAI’s relevant
8 business decisions. The asserted basis for compelling this discovery appears to be purely
9 speculative – as does the contention that the identities of these shareholders would help Plaintiffs
10 “understand the Defendants’ corporate relationship and structure of relationship.” As to the
11 contention that the *identity* of those owning shares in excess of the 5% threshold would “address
12 the financial condition of the Defendants, which is relevant to the ability of Defendants to respond
13 to a judgment,” the court finds this suggestion to be unpersuasive. The *identity* of a company’s
14 shareholders does not appear to be something that would shed any light on the company’s
15 financial condition, or the company’s ability to respond to a judgment. Thus, the request to compel
16 information pursuant to ROG No. 14 must be **DENIED** because this discovery appears to be more
17 in the nature of a fishing expedition based on suspicion and assumption rather than a request for
18 discoverable material premised on concrete facts and assertions.

19 **IT IS SO ORDERED.**

20 Dated: May 7, 2024



ROBERT M. ILLMAN
United States Magistrate Judge

21
22
23
24
25
26
27
28