

UNITED STATES DISTRICT COURT FOR THE NORTHERN DISTRICT OF
CALIFORNIA

In re Mosaic LLM Litigation

In re Mosaic LLM Litigation · No. 24-cv-01451-CRB (LJC) · Decided June 26, 2025

OPINION

In Re Mosaic LLM Litigation

PER CURIAM.

1 2 3

4 UNITED STATES DISTRICT COURT

5 NORTHERN DISTRICT OF CALIFORNIA

6 7 Case No. 24-cv-01451-CRB (LJC) 8 In Re Mosaic LLM Litigation

ORDER RESOLVING DISCOVERY

9 LETTER BRIEF AT ECF NO. 127

10 Re: Dkt. No. 127 11 12 13 14 Pending before the Court is the parties’ Joint Discovery Letter
regarding whether the CEO 15 of Defendant Databricks, Dr. Ali Ghodsi, may be designated as a
document custodian. ECF No. 16 127. Plaintiffs argue that Ghodsi is “uniquely situated to provide
relevant information[,]” given 17 his leadership role and involvement in the acquisition of
Defendant MosaicML. Id. at 1. 18 Defendants disagree, characterizing Plaintiffs’ arguments as
mere “speculation that [Ghodsi] must 19 have unique information simply by virtue of serving as
CEO.” Id. at 5. For the following reasons, 20 the Court GRANTS Plaintiffs’ request to designate
Ghodsi as a document custodian. 21 Defendants frame their objections to Ghodsi serving as a
custodian as being “on the basis 22 of Apex considerations.” Id. at 4. The apex doctrine “is
essentially the courts’ recognition that 23 deposing high-level officers of large corporate and
governmental entities imposes meaningful 24 burdens on those entities and witnesses, in part
because in the absence of any restrictions, sitting 25 for depositions could quickly become a full-
time job for leaders of entities that face extensive 26 litigation.” In re Uber Techs., Inc., Passenger
Sexual Assault Litig., No. 23-md-03084-CRB 27 (LJC), 2025 WL 896412, at *1 (N.D. Cal. Mar.
24, 2025). The Court notes that the apex doctrine 1 Co., Ltd, 282 F.R.D. 259, 263 (N.D. Cal.
2012) (analyzing the apex doctrine as applying to 2 depositions of high-level executives); In re
Facebook, Inc. Consumer Priv. User Profile Litig., No. 3 3:18-MD-02843, 2021 WL 10282213, at
*10 (N.D. Cal. Nov. 14, 2021) (finding that the apex 4 doctrine did not apply to a dispute
regarding whether executives “should be added as document 5 custodians”). However, the parties’
stipulated ESI Protocol expressly provides that, if the 6 requesting party (here, Plaintiffs) requests
to designate three additional document custodians, the 7 producing party may “object” to
Plaintiffs’ selections “only on the basis of Apex 8 considerations[.]” ECF No. 61 at 2. The Court

accordingly considers the appropriate apex 9 principles—namely, the executive’s access to relevant information, such information’s 10 uniqueness, and burden of production—in this decision. See *Apple v. Samsung*, 282 F.R.D. at 263 11 (rejecting the “two-prong test” of permitting apex depositions only if the deponent has “unique 12 first-hand, non-repetitive knowledge” and other “less intrusive discovery methods” have been 13 exhausted in favor of a “more nuanced equation” considering relevance, burden, unique 14 knowledge, and availability of other discovery methods). 15 Plaintiffs argue that Ghodsi “near undeniably has unique documents relevant to Plaintiffs’ 16 vicarious infringement claim,” as he was the “architect, the engine, and the closer” of Databricks’ 17 acquisition of MosaicML. ECF No. 127 at 2. In support of this contention, they cite to an 18 interview given by Ghodsi where he discussed meeting MosaicML’s CEO at a conference and 19 deciding to try to purchase MosaicML based on his believe that MosaicML’s generative AI 20 models were extremely valuable. *Id.* The Court finds it credible that, given Ghodsi’s role in the 21 acquisition, he will have documents regarding Databricks’ acquisition of Mosaic, which are 22 arguably relevant to Databricks’ knowledge of Mosaic’s alleged infringing activity and the 23 financial benefit Databricks believed it could obtain from Mosaic’s infringing activity. See *In re 24 Glumetza Antitrust Litig.*, No. 19-cv-05822, 2020 WL 3498067, at *7 (N.D. Cal. June 29, 2020). 25 Defendants try to distinguish this case from *Kadrey v. Meta*, where the plaintiffs sought to “add 26 Meta’s CEO based on produced documents indicating that he was a principal decision-maker on 27 acquiring the datasets at issue.” ECF No. 127 at 5. Plaintiffs cite to interviews where Ghodsi 1 datasets including Plaintiffs’ copyrighted works. As in *Kadrey v. Meta*, Ghodsi’s role as a 2 “principal decision-maker” supports Plaintiffs’ position that he has relevant information. While it 3 is somewhat more speculative that Ghodsi has documents regarding Databricks and MosaicML’s 4 relationship post-acquisition, given Ghodsi’s role in acquiring MosaicML and integrating it into 5 Databricks, it seems highly plausible that Ghodsi has information regarding their ongoing 6 relationship.¹ Judge Breyer’s recent decision to allow Plaintiffs to add a direct infringement claim 7 against Databricks strengthens Plaintiffs’ argument that Ghodsi will have relevant documents and 8 is an appropriate custodian. See ECF No. 129. 9 Defendants rely heavily on the argument that Plaintiffs have not shown Ghodsi will have 10 unique documents, contending that he does not have any documents relevant to Plaintiffs’ claims 11 that “would not be in corporate documents, [former Mosaic CEO] Dr. Rao’s documents, or 12 documents from the other eleven custodians.” ECF No. 127 at 4. But that there will be overlap 13 between Ghodsi’s custodial documents and the other custodians’ is not dispositive. “The idea that 14 responsive documents will necessarily be found in other custodians’ records is not sufficient to 15 defeat a search of [a custodian’s] files.” *Shenwick v. Twitter*, No. 16-cv-05314, 2018 WL 833085, 16 at *1 (N.D. Cal. Feb. 7, 2018); see *In re Uber Techs.*, 2025 WL 896412, at *2 (explaining, in the 17 context of depositions, that “an overemphasis on unique knowledge is inconsistent with otherwise 18 1 Defendants argue that Plaintiffs claim that Ghodsi has unique and relevant information is based 19 on speculation. See ECF No. 127 at 5. Certainly, a requesting

party with limited access into a responding party's decision-making process and corporate structure often must engage in some 20 degree of speculation when choosing custodians. Given this information imbalance, the responding party is generally best positioned to select custodians, as they know and can "identify 21 those individuals within its organization likely to have information relevant to the case." In re EpiPen (Epinephrine Injection, USP) Mktg., Sales Pracs. & Antitrust Litig., No. 17-md-2785, 22 2018 WL 1440923, at *2 (D. Kan. Mar. 15, 2018); see The Sedona Principles, Third Edition, 19 Sedona Conf. J. 1, 118 (2018) (explaining, in the introduction to Principle 6, that "a responding 23 party is best situated to preserve, search, and produce its own ESI"). The ESI Protocol reflects this general principle, as it allows the responding party to select their first ten document 24 custodians. ECF No. 61 at 2. Once the responding party has identified ten custodians, the requesting party can identify three more, subject to limited objections. Id. Plaintiffs have 25 followed this protocol, and, based on the limited information available to them, made a selection they believe will produce relevant information. While their decision to select Ghodsi inevitably 26 involved some degree of speculation, they "will have to bear the consequences" of making an unwise selection. Tremblay v. OpenAI, Inc., 23-cv-03223 (N.D. Cal. Ot. 18, 2024), ECF No. 191 27 at 3. 1 common approaches to discovery and trial advocacy, where corroborating evidence and testing 2 || credibility through the testimony of more than one witness is often appropriate"). Moreover, 3 Ghodsi has a unique leadership role relative to other custodians. The majority of agreed-upon 4 custodians, identified at ECF No. 127-1, are members and directors of MosaicML's research and 5 engineering teams. As Plaintiffs argue, Ghodsi, as CEO of the parent company Databricks, may 6 || well have unique information regarding "strategic-level LLM development commercialization, 7 including communications with Databricks' board and investors." ECF No. 127 at 2. 8 Defendants do not argue that designating Ghodsi as a document custodian would be unduly 9 || burdensome. As the parties agreed to initially identify up to thirteen document custodians, and 10 || Defendants do not contend that it would be uniquely burdensome to produce Ghodsi's custodial 11 file, the Court has no reason to find that designating Ghodsi as one of the thirteen custodians 12 || would pose an undue burden on Defendants. 5 13 Plaintiff's request to designated Ghodsi as a document custodian is accordingly || GRANTED. 3 15 IT IS SO ORDERED. 16 Dated: June 26, 2025 1g 7

LIBA 4. CISNEROS

19 Liga CIs Magistrate Judge 20 21 22 23 24 25 26 27 28