

UNITED STATES DISTRICT COURT FOR THE NORTHERN DISTRICT OF CALIFORNIA

In re Mosaic LLM Litigation

Mosaic LLM Litigation · No. 24-cv-01451-CRB (LJC) · Decided December 8, 2025

SUMMARY

The United States District Court for the Northern District of California denied without prejudice Plaintiffs’ request to take five additional third-party depositions in the Mosaic LLM litigation. The Court ordered that a Rule 30(b)(6) deposition of Hanlin Tang be reopened for two additional hours concerning Topics 37 and 40, finding that Defendants had not adequately prepared the designee on matters reasonably available from documents and former employees.

CASE DETAILS

COURT	United States District Court for the Northern District of California
WRITING FOR THE COURT	Lisa J. Cisneros
JURISDICTION	United States District Court for the Northern District of California
DECIDED	December 8, 2025
DOCKET	24-cv-01451-CRB (LJC)
DISPOSITION	other
PROCEDURAL POSTURE	Plaintiffs sought leave to take five additional third-party depositions after the close of fact discovery in an ongoing copyright-infringement action.
STANDARD OF REVIEW	The court evaluated the proposed discovery under Federal Rules of Civil Procedure 26(b)(1) and 26(b)(2), considering relevance, proportionality, timeliness, cumulative or duplicative discovery, availability from less burdensome sources, and the existing case schedule.
PRECEDENTIAL VALUE	Unknown; district court discovery order with no reporter citation.

TOPICS

- discovery dispute
- copyright infringement
- civil procedure
- commercial litigation

PRACTICE AREAS

QUESTIONS PRESENTED

1. Whether Plaintiffs should be permitted to take five additional third-party depositions after the close of fact discovery.
2. Whether the alleged failure to adequately prepare a Rule 30(b)(6) witness justified reopening that deposition for limited additional questioning.

HOLDINGS

1. The proposed five third-party depositions were relevant but not proportional to the needs of the case because the request was untimely, potentially duplicative of discovery already obtained, and information was available from less burdensome sources.
2. A corporate party must make a conscientious, good-faith effort to designate and prepare a knowledgeable Rule 30(b)(6) witness, using information reasonably available from documents, former employees, or other sources; Defendants had not adequately prepared Hanlin Tang on Topics 37 and 40.

KEY QUOTATIONS

“Although a ‘Rule 30(b)(6) designee is not required to have personal knowledge on the designated subject matter,’ a corporate party ‘has a duty to make a conscientious, good-faith effort to designate knowledgeable persons for Rule 30(b)(6) depositions and to prepare them to fully and unequivocally answer questions about the designated subject matter.’” (at 5)

“The fact that an organization no longer has a person with knowledge on the designated topics does not relieve the organization of the duty to prepare a Rule 30(b)(6) designee.” (at 6)

“Defendants must comply with their obligations under Rule 30(b)(6) and prepare their ‘Rule 30(b)(6) designee to the extent matters are reasonably available, whether from documents, past employees, or other sources’ to testify on Topics 37 and 40.” (at 6)

FACTUAL BACKGROUND

Plaintiffs assert direct copyright infringement claims against MosaicML and vicarious copyright infringement claims against Databricks, while also seeking leave to add additional infringement theories. Plaintiffs sought third-party depositions concerning the alleged hosting, sharing, and distribution of datasets, including depositions of AWS, a former MosaicML employee, and three customers. The request was made shortly before the close of fact discovery, after Plaintiffs had already obtained substantial party and documentary discovery, and AWS access logs were being pursued through a subpoena.

PROCEDURAL HISTORY

Plaintiffs had already received permission to take five additional party depositions and sought to depose Amazon Web Services, a former MosaicML employee, and three customers. The magistrate judge denied the request

without prejudice but reopened the Rule 30(b)(6) deposition of Hanlin Tang for two additional hours concerning Topics 37 and 40.

DISPOSITION

other

REMAND INSTRUCTIONS

The request to depose AWS, Nicholas Sapp, and three customers was denied without prejudice. Defendants must prepare Hanlin Tang in accordance with Rule 30(b)(6) and the order and must permit an additional two hours of questioning concerning Topics 37 and 40. The joint administrative motion to file excess pages as exhibits was granted.

CASES CITED

- Kadrey v. Meta Platforms, Inc., No. 23-cv-03417-VC (TSH), 2024 WL 4502099, at *1 (N.D. Cal. Oct. 15, 2024)
- Great Am. Ins. Co. of New York v. Vegas Const. Co., 251 F.R.D. 534, 539-40 (D. Nev. 2008)
- Flodin v. Cent. Garden & Pet Co., No. 21-cv-01631-JST (DMR), 2024 WL 3387620, at *2 (N.D. Cal. July 11, 2024)

OPINION

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 DISPUTE REGARDING THIRD- PARTY DEPOSITIONS 10 Re: ECF No. 201
11 12 13 14 Before the Court is the parties’ discovery letter brief, filed on November 25, 2025, 15
disputing whether Plaintiffs may take five third-party depositions.

ECF Nos. 199-2, 201.1 16 Plaintiffs wish to depose Amazon Web Services, former MosaicML
employee Nicholas Sapp, and 17 three of Defendants’ customers. The Court held a hearing on the
matter on December 4, 2025. 18 Having considered the relevant legal authorities and arguments
presented, Plaintiffs’ request to 19 depose five third parties is DENIED without prejudice.

However, the Court orders that the 20 30(b)(6) deposition of Hanlin Tang2 be reopened for an
additional two hours for further 21 questioning regarding Topics 37 and 40. The parties’ joint
administrative motion for leave to file 22 excess pages as exhibits to their letter brief is granted.
ECF No. 200. 23 I. BACKGROUND 24 The Court assumes the parties’ familiarity with the
overall procedural and factual 25 background of the case and provides a limited summary of the
procedural history only to 26 1 A redacted copy of the letter brief was filed at ECF No. 201.

An unredacted copy was filed 27 under seal at ECF No. 199-2. Throughout this Order, the
undersigned cites to the version of the 1 contextualize its ruling. Plaintiffs’ First Amended

Consolidated Complaint asserts a claim against 2 Defendant MosaicML for direct copyright infringement and a claim against Defendant Databricks 3 for vicarious copyright infringement.

See ECF Nos. 131, 162. Plaintiffs have filed a motion for 4 leave to file a second amended consolidated complaint to, among other things, revive their direct 5 infringement claim against Databricks and add in claims for contributory infringement and 6 inducement of infringement against both Defendants. See ECF No. 196-3 at ¶¶ 66-100. 7 Defendants oppose. ECF No. 213.

The matter is set for a hearing in front of Judge Breyer in 8 January 2026. See ECF No. 196. 9 The close of fact discovery was on November 21, 2025. The Court previously granted 10 Plaintiffs' request to take five extra party depositions (in addition to the ten allowed under Rule 11 30(a)(2) (A)(i) of the Federal Rules of Civil Procedure) and provided the parties "a brief extension" 12 of the discovery deadline—until December 5, 2025—to complete these depositions.

ECF No. 187. 13 Plaintiffs now seek to depose five additional nonparties. ECF No. 201. Plaintiffs filed an 14 administrative motion to amend the case schedule and continue the discovery cutoff to January 5, 15 2026, which Defendants opposed. ECF Nos. 191, 198. Judge Breyer has yet to rule on that 16 administrative motion. 17 II. LEGAL STANDARD 18 A party may generally "depose any person" without leave of the court.

Fed. R. Civ. P. 19 30(a)(1). A nonparty "deponent's attendance may be compelled by subpoena under Rule 45." Id. 20 However, absent stipulation, a party must obtain leave of the court to take more than ten 21 depositions. Fed. R. Civ. P. 30(a)(2). "[T]he court must grant leave" to take more than ten 22 depositions "to the extent consistent with Rule 26(b)(1) and (2)." Id. Rule 26(b)(1) provides that 23 "[p]arties may obtain discovery regarding any nonprivileged matter that is relevant to any party's 24 claim or defense and proportional to the needs of the case, considering," among other things, "the 25 importance of the discovery in resolving the issues, and whether the burden or expense of the 26 proposed discovery outweighs its likely benefit." Rule 26(b)(2)(A) authorizes courts to "alter the 27 limits in these rules on the number of depositions ... or on the length of depositions under Rule 1 On motion or on its own, the court must limit the frequency or extent of discovery otherwise allowed by these rules or by local rule if it 2 determines that: 3 (i) the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is 4 more convenient, less burdensome, or less expensive; 5 (ii) the party seeking discovery has had ample opportunity to obtain the information by discovery in the action; or 6 (iii) the proposed discovery is outside the scope permitted by 7 Rule 26(b)(1).

8 Accordingly, in evaluating whether to grant Plaintiffs leave to take the additional five depositions, 9 the Court considers the factors enumerated in Rule 26(b). "[T]he burden inquiry includes 10 considering whether the proposed discovery can feasibly be completed" within the existing case 11 schedule. *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC (TSH), 2024 WL 4502099, at *1 12 (N.D.

Cal. Oct. 15, 2024). 13 III. ANALYSIS 14 Defendants argue that the discovery Plaintiffs seek through their third-party depositions is 15 not relevant to the current claims and defenses, and that Plaintiffs are improperly fishing for 16 evidence to support their unpled contributory infringement claim. ECF No. 199-2 at 4-5.

The 17 Court sees where Defendants are coming from—the discovery Plaintiffs seek is more obviously 18 relevant to Plaintiffs’ unpled claims—but agrees with Plaintiffs that the third-party depositions 19 seek discovery relevant to their existing claims and Defendants’ fair use defense. ECF No. 199-2 20 at 4. Plaintiffs’ existing direct infringement claim asserts that MosaicML distributed Plaintiffs’ 21 works without permission, and how Defendants’ used Plaintiffs’ works—including, arguably, if 22 they distributed the works for commercial purposes—will likely bear on the fair use defense.

See 23 ECF No. 131 at ¶¶ 38, 50; 17 U.S.C. § 107(1). Plaintiffs seek to depose one of the platforms 24 Defendants used to allegedly host and share the datasets, three of the customers to whom 25 Defendants allegedly distributed the datasets, and a former employee familiar with the alleged 26 distribution and technological procedures for making datasets available to customers.

It is likely 27 that such depositions would yield testimony relevant to Defendants’ alleged distribution of the 1 datasets. Although the Court does not view the alleged distribution as a central part of Plaintiffs’ 2 claims or the fair use defense, discovery into this topic is relevant. 3 But the analysis does not end with relevance. See Fed. R. Civ. P. 26(b).

Although some of 4 the Rule 26(b)(1) factors weigh in favor of permitting the extra depositions—namely, the amount 5 in controversy and the importance of the issues at stake in this action—the Court finds that 6 allowing the five depositions is not proportional to the needs of the case.

First, Plaintiffs’ request 7 is untimely. Plaintiffs explain that they “notified Defendants of their intent to seek third-party 8 depositions on November 14.” ECF No. 199-2 at 4. That was seven days before the close of fact 9 discovery.

Although Plaintiffs argue that the belatedness of their request was “a problem of 10 Defendants’ own making” because Defendants’ 30(b)(6) witnesses were unprepared to testify 11 about the distribution of datasets, this does not excuse Plaintiffs’ delay in raising this issue.³ Id. 12 As Plaintiffs acknowledged during the hearing, Defendants substantially completed document 13 production in August.

Plaintiffs were aware of the dataset distribution issue well before the 14 discovery cut-off and should have diligently sought third-party depositions within the discovery 15 timeframe. Permitting the third-party depositions at this late stage would derail the case schedule. 16 Second, the depositions appear somewhat duplicative of discovery Plaintiffs have obtained 17 from Defendants.

See Fed. R. Civ. P. 26(b)(2)(C)(i) (permitting the court to limit discovery if “the discovery sought is unreasonably cumulative or duplicative”). As Defendants explain, “Plaintiffs have already deposed multiple witnesses at length” about “Defendants’ alleged sharing of datasets with a limited number of enterprise customers” and Plaintiffs have obtained “significant documentary discovery into alleged dataset sharing.” *Id.* at 5-6; see, e.g., ECF Nos. 199-10 at 10-22 11, 199-11 at 5-12.

Information from third parties would likely supplement what Plaintiffs have already obtained from Defendants or allow Plaintiffs to corroborate Defendants’ witnesses’ testimony about the alleged distribution, but this is not a situation where Plaintiffs’ only source for relevant material is from third parties.

As discussed below, to the extent Plaintiffs were unable to obtain discovery from Defendants because Defendants’ 30(b)(6) witness was not prepared to testify about the noticed topics, such deficiencies may be cured by reopening the 30(b)(6) deposition. See *Great Am. Ins. Co. of New York v. Vegas Const. Co.*, 251 F.R.D. 534, 540 (D. 3 Nev. 2008) (“[I]f an organization designates a witness it believes in good faith would be able to provide knowledgeable responsive testimony and it becomes apparent during the deposition that the designee produced is unable to respond to relevant areas of inquiry, the responding party has a duty to designate an additional knowledgeable deponent.”).

Third, it is not apparent that the discovery Plaintiffs seek could not “be obtained from some other source that is more convenient, less burdensome, or less expensive.” Fed. R. Civ. P. 26(b)(2)(C)(i). Plaintiffs argue that they must be permitted to depose Amazon Web Services (AWS) to learn who downloaded the datasets hosted on AWS. ECF No. 199-2 at 2-3.

But, as Plaintiffs acknowledged, they have subpoenaed the access logs from AWS and are in talks with AWS regarding their production. Without prejudice to any arguments the parties or AWS may raise regarding production of the access logs, that Plaintiffs are in the process of obtaining information they need through less burdensome means undercuts their need to depose AWS.

Plaintiffs’ request to depose AWS, Nicholas Sapp, and three of Defendants’ customers identified in the letter brief is accordingly DENIED without prejudice to Plaintiffs’ renewing their request if Judge Breyer modifies the case schedule or permits Plaintiffs to amend their complaint, or if ongoing discovery reveals information that would significantly alter the relevance and proportionality analysis above.

As noted, Plaintiffs argue that one of the reasons the third-party depositions are justified is that Defendants’ 30(b)(6) designee, Hanlin Tang, was not prepared to testify fully about deposition Topics 37 (“The instances in which You provide or have provided access to your

Training data to third parties, including customers and the public, and any policies regarding the 24 same.”) and 40 (“The technological support You provide to third parties, including customers and 25 the public, to download and use Pirated books datasets to train LLMs.”).

ECF No. 199-2 at 2. 26 They cite to testimony where Tang was unable to answer specific questions regarding customers’ 27 access of the datasets and where Tang explained that former MosaicML employee Nicholas Sapp 1 contracted in preparation for the 30(b)(6) deposition. Id. at 2-3. Defendants dispute that Hanlin 2 Tang was unprepared and argued at the hearing that they had no obligation to contact a former 3 employee as part Tang’s 30(b)(6) preparation.

4 Although a “Rule 30(b)(6) designee is not required to have personal knowledge on the 5 designated subject matter,” a corporate party “has a duty to make a conscientious, good-faith effort 6 to designate knowledgeable persons for Rule 30(b)(6) depositions and to prepare them to fully and 7 unevasively answer questions about the designated subject matter.” Great American, 251 F.R.D.

8 at 539 (quotation marks omitted). “The fact that an organization no longer has a person with 9 knowledge on the designated topics does not relieve the organization of the duty to prepare a Rule 10 30(b)(6) designee.” Id. Although “it is not uncommon to find that a corporation no longer 11 employs individuals” knowledgeable about past events, this does “not relieve a corporation from 12 preparing its Rule 30(b)(6) designee to the extent matters are reasonably available, whether from 13 documents, past employees, or other sources.” Id. (quotation marks omitted); see *Flodin v. Cent.*

14 Garden & Pet Co., No. 21-cv-01631-JST (DMR), 2024 WL 3387620, at *2 (N.D. Cal. July 11, 15 2024) (explaining that a responding organization has a duty to educate its 30(b)(6) designee with 16 information “reasonably available to the corporation, which includes ... information from past 17 employees”). Tang testified that former MosaicML employee Nicholas Sapp was knowledgeable 18 about the specifics of providing customers access to the datasets.

See, e.g., ECF No. 199-5 at 12 19 (“We’d probably have to ask the IT team ... Probably ask Nicholas Sapp.”). Mr. Tang testified, 20 and Defendants’ counsel confirmed, that Mr. Sapp was not contacted in preparation for the 21 30(b)(6) deposition. This supports Plaintiffs’ argument that Defendants did not prepare Mr. Tang 22 “to fully and unevasively answer questions about” Topics 37 and 40, and warrants the reopening 23 of Mr. Tang’s 30(b)(6) deposition for the limited purpose of questioning him about these topics.

24 Great American, 251 F.R.D. at 539 (quotation marks omitted). Defendants must comply with 25 their obligations under Rule 30(b)(6) and prepare their “Rule 30(b)(6) designee to the extent 26 matters are reasonably available, whether from documents, past employees, or other sources” to 27

testify on Topics 37 and 40. Id. The Court is cognizant of the burden on Defendants, particularly 1
on the record.

2 IT IS SO ORDERED. 3 Dated: December 8, 2025 4 5 * L J. CIXNEROS 6 ed States Magistrate
Judge 7 8 9 10 11 a 12 14 15 16 17 6 18 19 20 21 22 23 24 25 26 27 28

Retrieved from lawtools.ai · <https://lawtools.ai/cases/federal/united-states-district-court-for-the-northern-district-of-ca/2025/in-re-mosaic-llm-litigation-1a68dtl4>