

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

IN RE: LYFT, INC. PASSENGER
SEXUAL ASSAULT LITIGATION

Case No. [26-md-03171-RFL](#) (LJC)

**ORDER RESOLVING ESI PROTOCOL
DISPUTE**

Re: Dkt. Nos. 412

This multidistrict litigation concerns passenger safety with respect to alleged sexual misconduct and sexual assault by drivers using the Lyft platform. The parties have agreed upon many provisions for the use of technology-assisted review (TAR), but they have filed a discovery dispute letter concerning the scope of documents subject to validation. Dkt. No. 412. Plaintiffs argue that validation of the methods used to identify responsive documents must include validation of the search terms deployed, along with TAR validation. Lyft posits that only validation of the TAR process is required. Having reviewed the parties' letter, and considered the requirements established in Rule 26 of the Federal Rules of Civil Procedure, Plaintiffs' proposed edits to Section 10 of the ESI Order are **GRANTED IN PART AND DENIED IN PART** as drafted in Exhibit A. Plaintiffs' proposed edits are approved insofar as they include a Keyword-Excluded Set in the Validation Sample. However, to the extent that Plaintiffs' proposal would, as a practical matter, require validation to occur at the end of the document production process, it is **DENIED**.

I. LEGAL STANDARD

Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party's claim or defense and proportional to the needs of the case, considering the importance of the issues at stake in the action, the amount in controversy, the parties' relative access to relevant

information, the parties' resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit. Fed. R. Civ. P. 26(b). Information within this scope of discovery need not be admissible in evidence to be discoverable. *Id.*

Rule 26(g)(1) of the Federal Rules of Civil Procedure requires that a party responding to discovery requests must conduct "a reasonable inquiry." Perfection is not required in the realm of electronic discovery involving significant volumes of data and high costs to identify every responsive document. *See Lawson v. Spirit AeroSystems, Inc.*, 2020 WL 1813395, at *7 (D. Kan. April 9, 2020). While discovery on discovery is often disfavored, transparency with respect to the methodologies used in the identification of responsive electronically stored information may also be a reasonable part of the discovery process. *See Taylor v. Google LLC*, 2024 WL 4947270, at *2 (N.D. Cal. Dec. 3, 2024); *In re Diisocyanates Antitrust Litigation*, 2021 WL 4295729, at *6–7 (W.D. Pa. Aug. 23, 2021), *report and recommendation adopted*, 2021 WL 4295719 (W.D. Pa. Sept. 21, 2021). Technology-assisted review of ESI requires considerable transparency and cooperation among counsel. Without transparency, specifically the disclosure of sufficient information, discussions between counsel to cooperate and negotiate reasonable discovery measures and court review where disputes arise will be difficult, if not impossible. *See In re Diisocyanates Antitrust Litigation*, 2021 WL 4295729, at *6-7.

II. PROCEDURAL HISTORY

In July 2026, the Court resolved the parties' competing proposals concerning the discovery of electronically stored information, Dkt. No. 286, and entered Pretrial Order No. 19 Governing Lyft's Production of Electronically Stored Information and Hard Copy Documents, Dkt No. 297 (ESI Protocol or PTO 19). The ESI Protocol was composed of the parties' negotiated and agreed-upon provisions, and certain provisions directed by the Court. Sections 9, 10 and 11 of the ESI Protocol address search terms, as well as TAR, which uses machine learning algorithms to aid in the review of documents for relevance to the case. When Pretrial Order No. 19 issued, Lyft had not decided whether to use TAR to assist in the collection, review, or production of records. However, the ESI Protocol addressed preliminary matters pertaining to Lyft's usage of TAR in

1 this litigation—if, in fact, Lyft decided to use TAR—such as details that Lyft was required to
 2 disclose to Plaintiffs, validation methodology, and meet and confer requirements. The ESI
 3 Protocol also established provisions for the use of search terms to help identify responsive
 4 documents. In some cases, the parties dispute whether search terms should be used to cull a set of
 5 documents that is then subject to TAR review. In this litigation, Plaintiffs have not disputed
 6 search term culling in advance of TAR review.

7 **III. ANALYSIS**

8 Plaintiffs contend that the ESI Protocol requires “end-to-end” validation, which involves
 9 validation not only of the documents subject to TAR review, but also the documents that are never
 10 subject to TAR review because they are culled using search terms. Though Lyft argues that this
 11 type of validation is a nonstandard practice, Section 9 set forth that the “[u]se of search terms will
 12 be validated post-review as discussed in Section 11. Dkt. No. 297 at 8. The “end-to-end”
 13 validation efforts outlined in PTO 19, Section 11, clearly contemplated validation of “non-relevant
 14 sets and of the entire collection against which the search terms were run.” *Id.* at 10. Section 11
 15 states in full:

16 The Parties shall participate in an iterative and cooperative approach
 17 in which the Parties will meet and confer regarding reasonable and
 18 appropriate validation procedures and random sampling of
 19 Defendant’s Documents (both of relevant and non-relevant sets and
 20 of the entire collection against which search terms were run or TAR
 21 or AI or other identification or classification methodology was used),
 22 in order to establish that an appropriate level of end-to-end recall (the
 percentage of responsive Documents in the initial collection before
 any search terms or TAR or AI or manual review was applied which
 were classified as responsive after Defendant’s search, TAR, AI and
 review processes) has been achieved and ensure that the Defendant’s
 search, classification and review methodology was effective and that
 a reasonable percentage of responsive ESI was identified as
 responsive.

23 *Id.* Lyft contends that this language was a “placeholder.” Dkt. No. 412 at 7. However, while
 24 PTO 19 left room for the parties to meet and confer on methodology, and to reach amenable
 25 resolutions amongst themselves, the target validation populations and purpose of validation is
 26 clear. Validation is not intended only to determine whether TAR was reasonably effective, but to
 27 “ensure that the Defendant’s search, classification and review methodology was effective and that
 28 a reasonable percentage of responsive ESI was identified as responsive.” Dkt No. 297 at 10.

Lyft’s reliance on *Schulte v. LinkedIn Corp.* is misplaced. Dkt. No. 412 (citing *Schulte v. LinkedIn Corp.*, 2026 U.S. Dist. LEXIS 146482, at *8–10; 2026 WL 1905851 (N.D. Cal. July 1, 2026). In *Schulte*, Judge Beeler denied the plaintiffs’ request to prohibit LinkedIn altogether from using search strings to pre-cull the custodial documents reviewed by Relativity AiR. *Schulte*, 2026 WL 1905851, at *2–3. Here, Plaintiffs do not oppose Lyft from pre-culling the documents reviewed by TAR—they simply seek validation of the culled set to ensure the error rate of false negatives is reasonable.

Importantly, Judge Beeler rejected the plaintiffs’ request to compel LinkedIn to produce certain aiR Metrics because the ESI protocol did not appear to support requiring such disclosures. The ESI protocol only required LinkedIn to “disclose to the receiving party if they intend to use Technology Assisted Review (‘TAR’) to filter out non-responsive documents.” *Id.* at *3. LinkedIn satisfied that requirement and also provided additional information to the plaintiffs upon request. *Id.* Against that backdrop, Judge Beeler was unwilling to require LinkedIn to disclose more information absent a showing of a deficiency, and the plaintiffs simply pointed to the number of documents in the target population. *Id.* Here, for reasons explained above, the ESI Protocol supports a broad approach to validation that includes a sample of documents excluded by search terms. The record in this MDL also justifies including a sample of Keyword-Excluded documents because the percentage of custodial documents culled by search terms will well exceed the number of documents reviewed by TAR.

(Ret.) Judge Francis’s reasoning in *In re Diisocyanates Antitrust Litigation* makes clear the limitations of a validation process that is focused exclusively on the TAR portion of review. 2021 WL 4295729, at *9. “This requirement for validation applies to search terms just as it does to TAR.” *Id.* at *6 (citing *In re Seroquel Products Liability Litigation*, 244 F.R.D. 650, 662 (M.D. Fla. 2007) (“[W]hile key word searching is a recognized method to winnow relevant documents from large repositories, use of this technique must be a cooperative and informed process . . . Common sense dictates that sampling and other quality assurance techniques must be employed to meet requirements of completeness.”). The concerns detailed by Dr. Grossman in *In re Diisocyanates Antitrust Litigation* would exist even if a requesting party had some influence over

the selection of search terms, as in this MDL. *See id.* at *9. Though the parties in *In re Diisocyanates Antitrust Litigation* presented competing search term proposals, it is not obvious from that decision whether the defendants’ proposed terms were unilaterally selected or reflected any iterative input from the plaintiffs. *See id.* at 9–12. Nonetheless, in this case, the parties agree that a satisfactory hit rate for search terms falls within the 20–36% range, though they dispute the final number within that range. Dkt. No. 422 at 3, 5; *compare to In re Diisocyanates Antitrust Litigation*, WL 4295729, at *10 (“plaintiffs’ suggested search strings returned between 67% and 87% of the documents collected, depending upon defendant, while the hit rate for the defendants’ suggested search strings ranged from 29% to 55%”). This means that, in this MDL, at least 64% of the custodial records will never be subject to TAR. Validating the search term review is a measure that is proportionate to the needs of this litigation because a substantial portion of custodial records collected will never be subject to TAR review. Given that fact and the ESI Protocol, Plaintiffs need not rely solely on their preliminary investigations as they have negotiated search terms with Lyft to guard against the inadvertent culling of a significant number of responsive documents.

Furthermore, Lyft has not persuaded the Court that validating a set of Keyword-Excluded records will be unduly burdensome because the scope of the representative sample sought by Plaintiffs is modest. Plaintiffs represent—and Lyft does not dispute—that a “sample calculated at a 95 percent confidence level with a margin of error of plus or minus 2.5 percent would include about 1,537 additional documents,” in addition to the “approximately 3,500 documents already subject to review under the TAR validation process.” Dkt. No. 412 at 5. These figures are not unduly burdensome and, indeed, should be relatively quick to review because the Keyword-Excluded documents should be largely non-responsive.¹

Plaintiffs request for an order compelling Lyft to include a “Keyword-Excluded Set Sample” is **GRANTED**.

Turning to the issue of timing, Plaintiffs’ proposal that the “Keyword-Excluded Set will be

¹ Plaintiffs are also willing to consider alternative approaches to reduce the sample size while still providing meaningful information regarding the effectiveness of the search-term filter. *Id.*

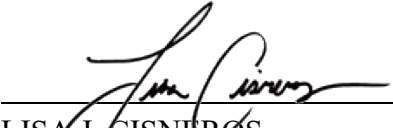
1 drawn from the entire population of documents subject to keyword filtering and TAR,” raises the
2 possibility that this set would not be drawn until the close of the document discovery period so as
3 to capture any and all documents subject to review. Plaintiffs appear to understand the need for
4 early validation. Dkt. No. 412 at 3 (“The question is, therefore, whether the parties should define
5 an objective process for validating the search terms now, when any problem can still be corrected
6 efficiently, or defer the issue until the cost of correction is far higher.”). The Court agrees that the
7 Keyword-Excluded validation set should be drawn sooner rather than later.

8 The parties and the Court will soon determine what search terms will be deployed. *See*
9 Dkt. No. 426. Reasonably soon after Lyft has completed the collection and production of
10 documents using those search terms, and well before the substantial completion deadline on
11 January 4, 2027, Lyft shall produce the “Keyword-Excluded Set Sample.” The parties shall meet
12 and confer no later than September 30, 2026 regarding a schedule and deadline for their validation
13 efforts to be completed. By October 5, 2026, the parties shall file either a stipulation setting forth
14 a schedule and deadline for their validation efforts to be completed, or a PTO 19 letter with their
15 respective positions.

16 Lyft’s alternative request for Plaintiffs to pay associated costs is **DENIED**.

17 **IT IS SO ORDERED.**

18 Dated: September 20, 2026

19
20
21 
22 LISA J. CISNEROS
23 United States Magistrate Judge
24
25
26
27
28