

LEE LINDERMAN

llinderman@hueston.com

523 West 6th Street

D: 213-788-4053

Suite 400

T: 213 788 4340

Los Angeles, CA 90014

F: 888 866 4825

HUESTON HENNIGAN LLP

September 11, 2026 – Via ECF

The Honorable Thomas S. Hixson, U.S. Dist. Court (N.D. Cal.)

Re: *Entrepreneur Media, LLC v. Meta Platforms, Inc.*, Case No. 3:25-cv-09579-VC
Joint Letter Brief re: Meta's Responses to Interrogatory Nos. 12, 19, and 20

Dear Magistrate Judge Hixson:

Plaintiff Entrepreneur Media, LLC (“EM”) and Defendant Meta Platforms, Inc. (“Meta”) submit this letter brief regarding Meta’s responses to EM’s Interrogatory Nos. 12, 19, and 20 (the “Contested Interrogatories”). Relevant excerpts of the Contested Interrogatories and Meta’s responses are attached as Exhibits A (Second Set, served January 2, 2026), B (Third Set, served March 10, 2026), C–E (Meta’s Second, Third, and Fourth Supplemental Responses to the Third Set, served August 4, August 29, and September 4, 2026), and F (Meta’s Third Supplemental Responses to the Second Set, served September 4, 2026). The parties conferred on these interrogatories more than a dozen times, including on August 17, 2026, September 8, 2026, September 9, 2026, and September 11, 2026. Meta’s last supplements were served on September 4, the final day of fact discovery, and Meta has not committed to a date for the further supplements it has promised; Civil Local Rule 37-3 requires that this dispute be presented today. EM’s position is that Meta’s responses remain deficient and it respectfully requests an order compelling further responses. Meta’s disagrees, but to compromise, Meta offers supplemental responses and solutions to address all of EM’s issues, as set forth below.

I. EM’S POSITION

“A party answering interrogatories has an affirmative duty to furnish any and all information available to the party” and “[i]nterrogatories must be answered separately and fully in writing under oath.” *Bryant v. Armstrong*, 285 F.R.D. 596, 612 (S.D. Cal. 2012) (quoting Fed. R. Civ. P. 33). Meta’s responses are not verified, so nothing Meta has said about its training data contents has been said under oath—and the answers are partial and internally inconsistent. EM requests an order requiring Meta to serve verifications by a date certain.

A. Interrogatory No. 19

Interrogatory No. 19 asks Meta to identify, for each version of Llama, which of EM’s works were included in any training dataset, the datasets in which each appeared, and the “source from which META obtained the work.” Ex. B at 6.

Meta has answered twice, in two different ways. On August 4, 2026, Meta served four spreadsheets reporting a vendor search that flagged every training-data record in which a ten-word sequence from an EM work matched “at least two different excerpts.” Ex. C at 8. On August 29 and September 4, Meta served lists reporting no search at all, stating Meta’s position on “knowledge, information and belief” as to which EM works are and are not in its training data. Exs. D at 8-9, E at 9. Meta says the August 4 exhibits are not superseded, so both stand.

HUESTON HENNIGAN LLP

*The search behind Meta's position on 261 books is undisclosed.*¹ Meta's August 4 disclosed search covered 335 of EM's 596 books and located EM content in 305 of them. Meta's responses describe no search covering the other 261, of which its list identifies 34:

	Books	Identified in training data
Covered by disclosed 8/4 search	335	305 (91%)
Not covered by any disclosed search	261	34 (13%)

Meta has never searched the SpiderMate data. Subparts (d) and (e) of the interrogatory reach Entrepreneur Magazine articles and content bearing EM's names—precisely what Meta's proprietary crawler collected from EM's websites. Meta has never run text-matching against that data, even though its response to Interrogatory No. 20, served the same day, confirms that Meta retains processed SpiderMate text from EM's domains. Hence the magazine results: of 152 issues spanning 2009 through 2026, Meta identifies only nine issues, and not a single article.

Meta removed 30 works its own search located, on the ground that the matched text was a block quote, excerpt, blogpost, or review. But Interrogatory No. 19 does not ask which works appear in full, and EM cannot evaluate any exclusion because the August 4 exhibits give the dataset for each match but no match count and no matched text.

The answers omit fields the interrogatory names: the version(s) of Llama for which each work was used (Meta says only “at least one”); the “source from which META obtained the work” (Meta gives the dataset where the work resides); and the author, ISBN, and publication date of every of the 309 works listed, 50 of which have no title either.²

Requested relief: EM requests an order directing Meta to (i) state in a verified response the search Meta applied to every EM book and magazine issue produced in this action, produce the results, and disclose the search parameters; (ii) apply that methodology to the SpiderMate data, without prejudice to EM's right to contest it; (iii) produce the match count and text for each work Meta's 8/4 search located and state the basis for excluding any such work from its 9/4 list; and (iv) identify the source from which Meta obtained each work, the version(s) of Llama for which it was used, and each work's title, author, ISBN, and publication date.

Meta has proposed, and describes below, a two-step process to define the universe of EM works at issue and to identify which Meta contends are and are not present in its training data. EM does not oppose that process and asks that the Court order it, on dates certain that permit use of the results in expert discovery, in addition to the relief above. Two clarifications: *First*, Interrogatory No. 19 is not limited to works Meta obtained from shadow libraries, and the resulting list of “Downloaded EM Works” would therefore exclude EM content Meta obtained by crawling EM's websites, which is the subject of subparts (d) and (e); *Second*, Meta's position that EM can check its results against the produced training data does not substitute for disclosure of the search parameters: reviewing the matches a search returned cannot disclose what the search was built not to find, and the ten-word, two-excerpt threshold determines that.

¹ Meta's counsel stated on September 11 that the text-matching was applied to the 261. That statement is in no response, is unverified, and describes neither the search nor its results.

² Rule 33(d) does not excuse any of this. It applies only where the burden of deriving the answer is substantially the same for both parties, and Meta built this search and ran it against corpora it alone controls, which Meta describes as requiring “considerable engineering effort.” Meta cannot invoke the rule selectively, answering for the works it chose to search and leaving the rest to EM.

HUESTON HENNIGAN LLP

B. Interrogatory No. 20

Interrogatory No. 20 asks Meta to identify content from EM's websites used to train its models, the datasets containing it, and "the approximate volume of such content (measured in tokens or documents)." Ex. B at 6. Meta answered by searching dataset URL metadata for EM's domains and producing lists of matching URLs. Three components remain outstanding: the processed text Meta's September 4 response confirms it "maintains" for content its crawler collected from entrepreneur.com, which Meta agreed on September 8 to produce; whether crawl data from the EM domains was used for any post-training purpose; and the volume component.

Requested relief: Meta's proposal on Interrogatory No. 20, set forth below, is acceptable to EM. EM asks the Court to order Meta to serve, by September 18, 2026, a verified response stating whether data from EM domains was used in post-training and if so how, and stating the approx. number of EM-domain webpages used for training, and to produce by the same date the processed text associated with the SpiderMate records previously produced, linked to crawl_id.

C. Interrogatory No. 12

Interrogatory No. 12 asks Meta to identify its uses of retrieval-augmented generation and to "describe the corpora indexed." Ex. A at 4. The corpora matter because Meta AI outputs can contain EM content fetched at the retrieval stage: Meta AI "conducts one or more rounds of web searches for additional information," subject to Meta's legal, business, or robots.txt blocklists, and then synthesizes those results into a response. Ex. F at 17. Meta identified the engines it uses but never the corpora indexed.

Requested relief: Meta's proposal on Interrogatory No. 12, set forth below, is acceptable to EM. EM asks the Court to order Meta to serve, by September 18, 2026, a verified supplemental response describing the content indexed by MASE, stating whether content from the EM domains is included in or retrievable through MASE, and stating whether any EM domain appears on any legal or business blocklist or is subject to robots.txt restrictions.

II. META'S POSITION

A. Interrogatory No. 19

Meta created and deployed a complex system for searching its training datasets for EM's works, which involves performing text matching between the text of EM's works and the text of the training datasets. Developing this technique required considerable engineering effort and iteration, but it achieves a result that, Meta believes, reliably locates EM's works to the extent they appear in the at-issue training datasets. Meta did all of this work using materials that both sides produced in discovery. EM could have done these searches itself if it was not willing to rely on Meta's results.

Meta has worked with EM in good faith over several weeks in an effort to resolve issues concerning identification of EM's works in the training datasets, but it has now become clear that EM will likely not be satisfied with any search unless it shows that nearly all of EM's books and magazines in Meta's training datasets (even though EM appears to concede above that such an outcome is unlikely). Meta does not agree with EM's characterizations and arguments about Meta's search methodologies, but rather than engage with those arguments, Meta believes its pages are best devoted to identifying a path forward for finally resolving EM's disputes for the remainder of this case.

HUESTON HENNIGAN LLP

What Meta Can Agree To In EM's Proposal: Meta is agreeable to a large portion of EM's request for relief for Interrogatory No. 19, subject to the understandings and explanations set forth below.

First, to resolve outstanding issues related to magazine searches, Meta will provide the following for **item (ii)** in EM's request for relief. Based on several meet and confers with EM's counsel, Meta understands that EM is asking that Meta have its vendor search the SpiderMate processed text data collected from the EM Identified Web Domains against the text-searchable magazine files produced by EM in this case. Meta will engage its vendor to perform that search and Meta will provide EM for each webpage, a list of magazine issue(s) that contain text matching one or more of the EM webpages in the SpiderMate processed data. The parties have agreed that Meta will not be obligated to match individual articles within a magazine issue with particular EM webpages. Meta cannot provide a date certain for these results as it requires Meta's vendor to engineer a solution to search the SpiderMate processed text data, which is formatted differently than prior searches. Meta expects this process to take several weeks and will keep EM apprised of its progress.

Second, for **item (iv)** Meta has already provided EM with information linking training data files, dataset sources and specific Llama models in prior interrogatory responses. To make it easier for EM to reference, Meta will aggregate this information into a single table collecting all of this information in one place and provide that to EM. Using this table, EM will be able to quickly associate each identified work with a particular dataset source and a particular Llama model. Meta expects to provide this table to EM by Friday, September 18th.

Additionally for item (iv), Meta will add information about the title, author, ISBN and publication date to each of the works identified in its forthcoming supplemental interrogatory responses proposed below, but in order to do so EM will need to provide a complete list associating each of the Bates numbers of each work it has produced for searching with its title, author, ISBN and publication date data. EM has only provided partial lists of this information to Meta, which has made it difficult for Meta to provide this detail in Meta's supplemental responses absent extensive manual review. EM has this information in its possession and should provide it promptly.

What Meta Proposes to Finally Resolve The Works Identification Issue: Meta cannot agree to items (i) and (iii) in EM's request for relief as stated. Items (i) and (iii) do not provide any path to reaching a resolution on these disputes; they will only engender more questions and disputes because they demand that Meta continue to explain searches it already ran, disclose unspecified "search parameters" that EM would presumably raise questions about, and provide results from prior searches that EM will continue to quibble with. EM's proposal is a recipe for more disputes and more court intervention.

Indeed, after hours of meeting and conferring with EM, Meta has determined, and EM's counsel has so much as admitted, that EM will not be satisfied until Meta agrees that all of EM's books and magazines were used to train Meta's Llama models, a fact that even EM appears to agree would not be true. Absent such an agreement from Meta, it has also become apparent from discussions and EM's request for relief above that EM intends to continue dragging this issue out for the next several months, continuing to ask for more information about Meta's searches and continuing to complain about Meta's works identification efforts at every CMC and discovery

HUESTON HENNIGAN LLP

conference. Providing more disclosures about searching methodologies and more lists of search results serves no purpose other than to extend this dispute without finally addressing the core issue in Interrogatory No. 19 - identifying which EM works Meta trained its Llama models on.

Continuing to litigate the details of Meta's text searching strategy is not a good use of the time and resources of the Court and the parties. Accordingly, Meta makes the following two-step proposal that it has already begun work on that will finally resolve this dispute once and for all.

Step One: Identify the Universe of Downloadable and Downloaded EM Works. First, Meta will determine which of the EM works were actually available in the so-called "shadow libraries" that Meta downloaded and will compare that list of works against Meta's records of what it actually downloaded. This will result in a list of "Downloaded EM Works" that will constitute the complete list of works from the "shadow libraries" that Meta could have possibly used for training its Llama models. Meta will provide this list of "Downloaded EM Works" to EM promptly upon its completion.

Step Two: Identify the Downloaded EM Works that Meta Did and Did Not Find In Its Training Data Searches. Next, Meta will supplement and verify its interrogatory response to identify (1) the Downloaded EM Works that Meta contends are present in the training data ("EM Included Works"), (2) the Downloaded EM Works that Meta contends are not present in the training data ("EM Excluded Works"), and (3) briefly explain why Meta contends the "EM Excluded Works" are not present in the training data based on its searches. Like Meta's prior interrogatory responses, Meta will link the "EM Included Works" to particular training datasets, using the table Meta is providing for items which can be associated with particular downloaded datasets and Llama models using the table Meta is providing in response to item (iv) above.

Meta will additionally provide EM with the text matching results for the EM Excluded Works, and will link those text matching results to particular training datasets and specific training data records, so that EM can check Meta's work using the training data that Meta previously produced to EM and assess for itself whether it agrees with Meta's conclusions. If EM disagrees and believes that one or more of the EM Excluded Works are present in the training data searches, EM is free to assert those works in the litigation provided EM has a good-faith basis for doing so after reviewing the results.

This simple procedure will provide certainty over the scope of EM Works that are being asserted in this case while also avoiding on-going disputes over search parameters and future motions over how Meta searched for its works. Meta expects this two-step process will take several weeks to finish and will keep EM apprised of its progress.

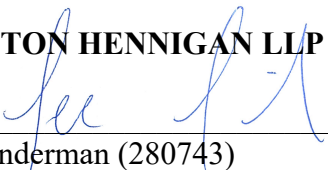
B. Interrogatory No. 20

As a compromise to resolve the dispute over this interrogatory, Meta has agreed to provide EM with (i) the processed text associated with the SpiderMate records previously produced, which will be linked to the "crawl_id" where applicable, (ii) whether data related to the identified EM web domains was used in post-training and if so, what, and (iii) Meta will provide the approximate number of EM domain webpages used for training has identified after a reasonable inquiry. Meta expects to provide this verified supplemental response by Friday, September 18th.

HUESTON HENNIGAN LLP**C. Interrogatory No. 12**

As a compromise to resolve the dispute over this interrogatory, Meta has agreed to provide EM with a supplemental interrogatory response (i) generally describing the types of content that is indexed by Meta AI's MASE search engine, (ii) identify whether content from EM's identified web domains is included in or retrievable through MASE, and (iii) identify whether any of EM's identified web domains are included on any legal or business blocklist, and identify if there are any robots.txt restrictions in place for those same domains. Meta expects to provide this verified supplemental response by Friday, September 18th.

Respectfully submitted,

HUESTON HENNIGAN LLPBy: s/ 
Lee Linderman (280743)*Attorneys for Plaintiff Entrepreneur
Media, LLC***COOLEY LLP**By: s/ Phillip E. Morton
Phillip E. Morton (*pro hac vice*)*Attorneys for Defendant Meta Platforms, Inc.*