NIGH GOLDENBERG RASO & VAUGHN

IN THE UNITED STATES DISTRICT COURT

FOR THE NORTHERN DISTRICT OF CALIFORNIA

IN RE: UBER TECHNOLOGIES, INC., PASSENGER SEXUAL ASSAULT LITIGATION	MDL No. 3084 CRB
This Order Relates To:	**PRETRIAL ORDER NO. 9: ORDER ON ESI PROTOCOL DISPUTES**
ALL ACTIONS	

The parties have filed competing proposed ESI protocols and briefs in support of their proposals. The Court resolves the parties' disputes below. The final ESI protocol will be entered as a stipulated order after the parties file a final version reflecting the Court's rulings in this Order.

Where the Court refers to section numbers from the parties' proposed ESI protocols, the Court will specify which party's document it is citing unless the disputed subject matter is address in the same section in both protocols.

I.   **DISCUSSION**

A.   **Definitions**

1.   **Definition of "Attachment"**

The parties disagree about how to define "attachment," at least in part as an outgrowth of a larger dispute about whether documents hyperlinked in electronic communications should be treated as "attachments." The resolution of this issue hinges on the outcome of the parties' dispute about the treatment of cloud-based documents. As

1   discussed in greater detail below, the Court will not resolve that issue at this juncture. See

2   Part I(J) below. This definition, too, should be revisited in accordance with the eventual

3   resolution of the cloud-stored documents issue.

2.   Definition of "Parent-Child"

5   Plaintiffs' proposed definition, contained in Section 2(t) of their proposed ESI

6   protocol, shall be adopted as follows: "'Parent-child' shall be construed to mean the

7   association between an attachment and its parent Document in a Document family."

B.   Cooperation

9   The parties disagree over certain introductory language that describes the parties'

10 obligations to cooperate. The core object of dispute is Uber's desire to include language

11 from Sedona Principle No. 6 to the effect that "responding parties are best situated to

12 evaluate the procedures, methodologies, and technologies appropriate for search, review,

13 and production of their own ESI[.]" The text on which the parties do not agree is

14 emphasized below:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 3: Cooperation	**Plaintiffs' Proposed Section 3: Cooperation**
The Parties are aware of the importance the Court places on cooperation and commit to cooperate in good faith throughout this Litigation consistent with this Court's Guidelines for the Discovery of ESI and this Court's Rules of Professional Conduct. ***The Parties further acknowledge that responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for search, review, and production of their own ESI, but that any such procedures must be consistent with the responding parties' obligation to make a reasonable and good faith effort to obtain the requested information via diligent search and reasonable inquiry, and any other duties***	The Parties are aware of the importance the Court places on cooperation and commit to cooperate in good faith throughout this Litigation consistent with this Court's Guidelines for the Discovery of ESI and this Court's Rules of Professional Conduct. ***The Parties will endeavor to cooperate in good faith and be reasonably transparent in all aspects of the discovery process, including the identification, preservation, and collection of sources of potentially relevant ESI, as well as propounding reasonably particular discovery requests, establishing proportional limits on the scope of potentially relevant and discoverable ESI, while endeavoring to identify and produce potentially relevant***

owed. The Parties additionally agree that while the Parties have, in the spirit of cooperation, detailed in this ESI Order their intended discovery-related processes and procedures, each Party is ultimately responsible for compliance with its discovery obligations under the Federal Rules of Civil Procedure, and may follow any such processes and procedures that satisfy obligations under those Rules to conduct discovery in a reasonable and proportional manner.	*and discoverable ESI, and maintaining security over the discovery in this Litigation.*

The Court will not require the parties' ESI protocol to restate Sedona Principle No. 6, nor will the Court require that the parties adopt the principle in the abstract. *See Klein v. Facebook, Inc.*, No. 20-cv-08570-LHK (VKD), 2021 U.S. Dist. LEXIS 175738, at *7 (N.D. Cal. Sep. 15, 2021). Nor will the Court adopt Plaintiffs' additional language in this section. Instead, the disputed paragraph shall read, in its entirety, as follows: "The Parties are aware of the importance the Court places on cooperation and commit to cooperate in good faith throughout this Litigation consistent with this Court's Guidelines for the Discovery of ESI and this Court's Rules of Professional Conduct."

C.   Documents Not "Reasonably Accessible"

Plaintiffs' proposal contains a paragraph describing a relatively detailed process for meeting and conferring about sources of ESI that the Producing Party determines are not "reasonably accessible," which includes a seven-day timeframe and a provision for bringing disputes to the Court if meeting and conferring is unsuccessful. Uber omits most of this language and simply says that the parties should meet and confer about such sources. The disputed language is as follows:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 7: Identification of Custodial and Non-Custodial Documents and ESI	**Plaintiffs' Proposed Section 7: Identification of Custodial and Non-Custodial Documents and ESI**

If the Producing Party determines that a source of ESI is not "reasonably accessible," pursuant to Fed. R. Civ. P. 26(b), the Parties will meet and confer as to the accessibility of the ESI.	The Parties agree that if the Producing Party determines a source is not "reasonably accessible" pursuant to Fed. R. Civ. P. 26(b) during the search and collection process it will provide sufficient information regarding the accessibility of the source to enable the Parties to confer in good faith within seven (7) days of this determination about whether such source or Document will be produced or methods by which the information can be produced. If the Parties disagree as to the accessibility of the source after a good faith meet and confer, the Party seeking discovery from the source may submit the issue to the Court or its designee in accordance with the Court's procedures. The Parties agree to take any unresolved disputes on same promptly to the Court or its designee.

Plaintiffs' proposal is preferable because it provides clearer guidance for resolving these disputes and a defined time period for doing so. The parties shall adopt the language in Plaintiffs' proposed Section 7.

D.   Search Queries and Methodologies

1.   Overview of Search Queries and Methodologies

The parties disagree over certain introductory language in Section 8 of their proposed ESI protocols:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8: Search Queries and Methodologies	**Plaintiffs' Proposed Section 8: Search Queries and Methodologies**
Pursuant to Fed. R. Civ. P. 26(f), and the ESI Guidelines and Section [7] above, the Parties will meet and confer, as appropriate, to discuss certain aspects of the discovery process, for example, the number of custodians, the identity of those custodians, keywords to be used as part of culling files, collection from non-custodial	Pursuant to Fed. R. Civ. P. 26(f), and the ESI Guidelines, the Parties shall meet and confer on the application, if any, of search or other filtering technologies, including search terms, file types, date ranges, transparent validation procedures and random sampling, predictive coding or other appropriate advanced technology,

files, file types, date ranges, validation procedures and random sampling, technology assisted review ("TAR") or other appropriate advanced technology. This process will be iterative. For the avoidance of doubt, Plaintiffs will disclose to Defendants their processes for preservation and collection of documents, including the sources from which such documents will be collected, and the parameters for search and review of documents. Plaintiffs will meet and confer with Defendants about these issues.	including systems used to track review status related to those advanced technologies, including systems used to track review status related to those advanced technologies. The Parties are expected to work in a cooperative, collaborative, and iterative manner, in order to reach agreement upon a reasonable search methodology to achieve an appropriate level of recall (the percentage of responsive Documents in the collection against which the search terms were run which include a search term). To the extent the Parties are unable to reach agreement on the application of, or procedures for, any search or filtering processes, the Parties shall raise such issues for resolution by the Court or its designee. The Parties recognize that as the litigation evolves, there may be a need to supplement earlier agreed methods or search terms to enhance or improve the identification of potentially relevant ESI.

The parties shall adopt Uber's proposed language for this portion of the protocol, with the following modifications:

> Pursuant to Fed. R. Civ. P. 26(f), and the ESI Guidelines ~~and Section [7] above~~, the Parties will meet and confer, as appropriate, to discuss certain aspects of the discovery process, for example, the number of custodians, the identity of those custodians, keywords to be used as part of culling files, collection from non-custodial files, file types, date ranges, validation procedures and random sampling, technology assisted review ("TAR") or other appropriate advanced technology. This process will be iterative. For the avoidance of doubt, Plaintiffs will disclose to Defendants their processes for preservation and collection of documents, including the sources from which such documents will be collected, and the parameters for search and review of documents. Plaintiffs will meet and confer with Defendants about these issues.

The Court strikes the reference to Section [7] because there are other sections of the ESI Protocol that address aspects of the discovery process that will be appropriate subjects for meet and confers. The Court also recognizes that terms such as "predictive coding" are

sometimes used interchangeably with TAR, though in recent years TAR appears to be the more commonly used term. *See* The Sedona Conference, TAR Case Law Primer, Second Edition, 24 Sedona Conf. J. 1, 7 n.1 (2023). Furthermore, as TAR methodologies evolve, they are categorized as TAR 1.0, TAR 2.0, and so on. The purpose of the prefatory language In Section 8 of the ESI Protocol is not to enumerate every topic for negotiation. Instead, the purpose is to require the parties to cooperate, consistent with the Federal Rules of Civil Procedure and our Court's guidelines, to determine the search queries and methodologies that will be used and how ESI discovery will be conducted more generally, in a manner that satisfies the "reasonable inquiry" under Federal Rule of Civil Procedure 26(g)(1).

2.   Technology Assisted Review (TAR)

The parties disagree over certain language in the description of the TAR methodology that Uber intends to use. The language on which the parties disagree is emphasized below.

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8(a)(1): Use of TAR by Uber Defendants	**Plaintiffs' Proposed Section 8(a)(1): Use of TAR by Uber Defendants**
As part of document review, the Uber Defendants intend to use TAR methodology known as TAR 2.0, which utilizes continuous active learning to classify and prioritize documents for attorneys to review. Specifically, the Uber Defendants intend to use Relativity Active Learning ("RAL") on a Relativity Server 12.1.537.3 platform provided by their vendor Lighthouse. Commonly, a TAR 2.0 methodology begins with ingesting document population into the TAR 2.0 software where the algorithm learns to distinguish relevant from non-relevant documents through attorney review of documents. The TAR 2.0 algorithm prioritizes the documents in the review	As part of document review, the Uber Defendants intend to use TAR methodology known as TAR 2.0, which utilizes continuous active learning to classify and prioritize documents for attorneys to review. Specifically, the Uber Defendants intend to use Relativity Active Learning ("RAL") on a Relativity Server 12.1.537.3 platform provided by their vendor Lighthouse. Commonly, a TAR 2.0 methodology begins with ingesting document population into the TAR 2.0 software where the algorithm learns to distinguish relevant from non-relevant documents through attorney review of documents. The TAR 2.0 algorithm prioritizes the documents in the review

queue *in a more efficient manner*. Attorney reviewers then review documents the TAR 2.0 model has prioritized as *most likely to be responsive*. As the review continues and reviewers code documents, the TAR 2.0 model continues to learn and prioritize likely responsive documents until a stopping point is reached and a validation is conducted.	queue *from most to least likely to be responsive*. Attorney reviewers then review documents the TAR 2.0 model has prioritized as most *likely to be responsive in descending order from most to least* likely to be responsive. As the review continues and reviewers code documents, the TAR 2.0 model continues to learn and prioritize likely responsive documents until a stopping point is reached and a validation is conducted.

The parties shall adopt Plaintiff's proposed language in this section.

3.   TAR and Search Terms

The parties disagree about whether the dataset to which the TAR methodology is applied will be pre-filtered with search terms—i.e., whether TAR processing will be "stacked" with the application of search terms. Plaintiffs request language foreclosing this approach, while Uber seeks to omit Plaintiffs' proposed language:

Uber's Proposal	Plaintiffs' Proposal **Plaintiffs' Proposed Section 8(a)(2): TAR and Search Terms**
[delete this paragraph]	TAR processing will not be "stacked" with the application of search terms, i.e., search terms will not be applied before or, unless agreed or ordered pursuant to Section 6.x below, after any application of TAR.

Instead of Plaintiffs' proposed language, the relevant section of the ESI protocol shall read as follows: "TAR processing may be 'stacked' with the application of search terms. If search terms are to be applied, the parties shall meet and confer regarding the proposed search terms. The search terms may be agreed by the parties, or certain search terms may be ordered by the Court if the parties are unable to reach an agreement."

4.   TAR Sample Set and TAR Training Process

Plaintiffs' proposal describes a detailed process by which the TAR methodology will initially be applied to a sample set of documents and the documents will be reviewed

1   and coded for relevance, with input from both parties. *See* Plaintiffs' Proposed ESI

2   Protocol, § 8(a)(3). It then describes a process through which the TAR software will be

3   trained using the agreed upon, coded sample set. *See id.* § 8(a)(4). Uber's proposal omits

4   any discussion of this subject. Uber, with the support of its expert Maura Grossman,

5   argues that the training processes described by the Plaintiffs are unnecessary for the TAR

6   2.0 methodology they will use. *See* Grossman Decl. (dkt. 262-7) ¶¶ 15–17.

7   The Court agrees with Uber. TAR 2.0, unlike TAR 1.0, does not require a

8   preliminary training process or a sample set with which to carry out that process, so it is

9   unnecessary for the protocol to include any of Plaintiffs' proposed language on these

10 subjects. Plaintiffs' proposed Sections 8(a)(3) and 8(a)(4) shall be omitted.

5. Stopping Criteria

12   The parties disagree over the language describing the "stopping criteria" that will

13 dictate when the TAR is paused for validation of the results. The parties' proposals are as

14 follows, with key disputed language emphasized:

Uber's Proposal	**Plaintiffs' Proposal**
Uber's Proposed Section 8(a)(2)(i): Stopping Criteria	**Plaintiffs' Proposed Section 8(a)(5)(i): Stopping Criteria**
Once *two reasonably sized review batches are found to contain 10% or fewer documents marked responsive*, Defendants will pause the review and turn to validation. Defendants may extend the review past this point if they believe sufficient thoroughness has not been achieved. Defendants do not intend for the relevant batches to include index health documents.	Once *two or more consecutive review batches sequentially populated by the highest-ranking uncoded documents remaining in the project in order from highest to lowest scores and containing a total of at least 1,000 documents are found to contain 10% or fewer documents marked responsive*, Defendants will pause the review and turn to validation. Defendants may extend the review past this point if they believe sufficient thoroughness has not been achieved. Defendants do not intend for the relevant batches to include index health documents.

27 Plaintiffs' proposed language provides more definite guidance for the validation process,

28

1 while Uber's approach is more vague and therefore more likely to lead to disputes. The

2 parties shall adopt Plaintiffs' proposed language on this topic.

6.   Validation—Recall and Richness

4 The parties disagree over a phrase regarding what information Uber will have to

5 disclose as part of the TAR validation process:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8(a)(3)(vii): Validation	Plaintiffs' Proposed Section 8(a)(6)(vii): Validation
vii. The Uber Defendants will disclose the Recall and Prevalence, *once calculated as set forth above.*	vii. The Uber Defendants will disclose the calculated Recall and Richness *and the input quantities used to calculate Recall and Richness.*

The parties shall adopt Plaintiffs' proposed Section 8(a)(6)(vii). However, references to

"Richness" shall be replaced with the term "Prevalence" in this section and in all other ESI

Protocol provisions concerning validation. See, e.g., Grossman Decl. ¶ 24 (describing

steps to estimate "recall and prevalence").

7.   Validation—Further Review

Plaintiffs' proposal sets forth a process through which Plaintiffs will have the

opportunity to review documents from the TAR validation sample and independently

assess whether the process has accurately coded the documents. Plaintiffs include a

provision that "[i]f the recall estimate derived from the validation sample is below 80%" or

otherwise is "too limited," then the parties shall discuss remedial action. Plaintiffs'

Proposed ESI Protocol § 8(a)(6)(x). Uber argues that Plaintiffs' proposal on this topic

would require an excessive degree of transparency and input into Uber's search processes.

They also argue that the benchmarks Plaintiffs set (for example, the 80% figure) are

unreasonable and unlikely to be met by any conceivable TAR approach.

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8(a)(3)(viii)	Plaintiffs' Proposed Sections 8(a)(6)(viii)-(x)

viii. Defendants will determine, based on this validation test, whether further review or other targeted searches may be warranted, or whether further review would be disproportionate and the TAR 2.0 process can be concluded.

viii. Plaintiffs' Designated Reviewers shall have the opportunity to review all non-privileged documents in the Validation Sample, without any knowledge of how any individual documents were coded by the Uber Defendants, in order to perform a blind comparison of the provided Recall and Richness estimates.

ix. This review may take place (a) at such location or locations mutually agreed by the Parties, on a date and time to be agreed to by the Parties, or (b) via a secure web-based viewer on a date and time to be agreed to by the Parties. Any documents coded Not Responsive by the Uber Defendants to which Plaintiffs' Designated Reviewers are provided access as part of this review are provided for the limited and sole purpose of raising and resolving disagreements, if any, regarding the coding calls made by the Uber Defendants. Any such disagreements shall be recorded on a TAR Protocol Classification Dispute Log (the "Log"), which shall be in a form agreed upon by the Parties. Once Plaintiffs' Designated Reviewers complete their review of the Validation Sample, the Parties shall meet and confer to resolve any differences in coding designation. If resolution cannot be reached, the issue shall be submitted to the Court for resolution.

x. If the recall estimate derived from the validation sample is below 80%, or if the documents designated responsive in the part of the sample drawn from the null set indicate that the TAR tool's model of Responsiveness was too limited, e.g., if the responsive documents in the Validation Sample included novel or significant documents, then Plaintiffs and Defendants will discuss remedial action to locate an adequate proportion of the remaining relevant documents in the null set, including but not limited to: continuing

	reviewing from the prioritized queue; and training alternative predictive models focused on the relevant documents found in the elusion test. After Defendants disclose these metrics, the parties may meet and confer to discuss questions and issues relating to the TAR process.

The review process that Plaintiffs propose establishes quality-control and quality-assurance procedures to validate Uber's production and ensure a reasonable production consistent with the requirements of Federal Rule of Civil Procedure 26(g). *See* Forrest Decl. ¶ 41; Luhana Decl. Ex. 8 (dkt. 261-9) (*In re: 3M Combat Arms Earplug Prods. Liability Litig.* TAR Protocol). However, some adjustment to the provision that appears to set an 80% recall requirement is needed. The parties shall adopt Plaintiffs' proposed Sections 8(a)(6)(viii)–(x) with this modification to Section 8(a)(6)(x). The following sentences shall be added at the end of Section 8(a)(6)(x):

> If the validation protocol leads to an estimate lower than 80%, or even lower than 70%, this lower recall estimate does not necessarily indicate that a review is inadequate. Nor does a recall in the range of 70% to 80% necessarily indicate that a review is adequate; the final determination of the quality of the review will depend on the quantity and nature of the documents that were missed by the review process.

8.   Disclosures

The parties disagree on certain language in the section on disclosures to be made about the TAR process after it is complete:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8(b): Disclosures	**Plaintiffs' Proposed Section 8(b): disclosures**

Once the TAR process is complete in addition to above, Defendants intend to disclose various metrics regarding the TAR 2.0 methodology utilized, including the following: (i) the total TAR population, (ii) the total population produced, (iii) the total population not produced, (iv) the total population not reviewed, (v) the size of the validation set used to verify the TAR 2.0 results, and (vi) a summary of the validation process. The summary of the validation process will include the following figures from the Validation Sample: (a) the number of documents within the sample that were previously coded relevant; (b) the number of documents within the sample that were previously coded not relevant; (c) the number of unreviewed documents within the sample. The summary of the validation process will also include the number of actual responsive documents identified in (a), (b), and (c) during the validation process. After Defendants disclose these metrics, the parties may meet and confer to discuss *reasonable* questions and issues relating to the TAR process. *If the volume of documents intended for TAR review becomes so large that it is necessary to run multiple TAR projects to ensure the smooth operation of the technology, the Producing Party will disclose that and will provide reasonable transparency into the TAR workflow.*	Once the TAR process is complete in addition to above, Defendants intend to disclose various metrics regarding the TAR 2.0 methodology utilized, including the following: (i) the total TAR population, (ii) the total population produced, (iii) the total population not produced, (iv) the total population not reviewed, (v) the size of the validation set used to verify the TAR 2.0 results, and (vi) a summary of the validation process. The summary of the validation process will include the following figures from the Validation Sample: (a) the number of documents within the sample that were previously coded relevant; (b) the number of documents within the sample that were previously coded not relevant; (c) the number of unreviewed documents within the sample. The summary of the validation process will also include the number of actual responsive documents identified in (a), (b), and (c) during the validation process. *Defendants will also produce all non-privileged responsive documents in the Validation Sample.* After Defendants disclose these metrics, the parties may meet and confer to discuss questions and issues relating to the TAR process.

The parties shall adopt Plaintiffs' proposed language on this topic.

9. Key Word Search

The parties disagree about the provision that should govern the use of search terms to find relevant ESI, where such terms are necessary in addition to or in combination with TAR. The competing proposals are as follows:

12

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 8(c): Key Word Search	Plaintiffs' Proposed Section 8(c): Key Word Search
If the Producing Party is identifying responsive ESI using search terms, the Parties will meet and confer about search terms in English and any other languages used in the Producing Party's documents. To facilitate the meet and confers, the Producing Party shall make disclosures reasonably necessary for the Requesting Party to assess the proposed terms and resolve any disputes.	If the Producing Party is identifying responsive ESI, which is not already known to be responsive, using search terms, the Parties will meet and confer about search terms in English and any other languages used in the Producing Party's documents. The Parties will meet and confer about information to improve the effectiveness of the search terms, such as providing a list of relevant English and foreign language company terminology (or equivalent) and all relevant project and code names, code words, acronyms, abbreviations, and nicknames, if any. Before implementing search terms, the Producing Party will disclose information and meet and confer within seven (7) days of the ESI Protocol being entered regarding the search platform to be used, a list of search terms in the exact forms that they will be applied (i.e., as adapted to the operators and syntax of the search platform), significant or common misspellings of the listed search terms in the collection to be searched, including any search term variants identifiable through a Relativity dictionary search with the fuzziness level set to 3, any date filters, or other culling methods, after which the Receiving Party may propose additional terms or culling parameters. Use of search terms will be validated post-review using comparable methodology and metrics to those set out in Disclosures (a) and (c) above.

The parties shall adopt Uber's proposal as modified with certain additional language, some of which is proposed by Plaintiffs. Thus, Section 8(c) shall read as follows:

If the Producing Party is identifying responsive ESI, which is not already known to be responsive, using search terms, the Parties will meet and confer about search terms in English and any other languages used in the Producing Party's documents. Before implementing search terms, the Producing Party will disclose information and meet and confer within seven days of the ESI Protocol being entered, or on a date agreed upon by the parties, regarding the search platform to be used, a list of search terms in the exact form that they will be applied (i.e., as adapted to the operators and syntax of the search platform), significant or common misspellings of the listed search terms in the collection to be searched, including any search term variants identifiable through a Relativity dictionary search with the fuzziness level set to 3, any date filters, or other culling methods. At the same time the Producing Party discloses the search terms, unless the Receiving Party agrees to waive or delay disclosure, the Producing Party shall disclose the unique hits, hits with families, and the total number of documents hit. Within seven days after the Producing Party discloses its list of search terms and related information, the Receiving Party may propose additional or different search terms or culling parameters and may propose a limited number of custodians for whom, across their email and other messages, the Receiving Party requests that no search term pre-culling be used prior to TAR 2.0. At the same time the Receiving Party discloses its proposals, it may request that the Producing Party provide hit reports, and the Producing Party must promptly respond with that information but may also provide other information. The parties must confer within 14 days after the Receiving Party's proposal to resolve any disputes about the search terms. The parties may agree to extend this deadline, but no extension may be more than 14 days without leave of the Court. Use of search terms shall be validated post-review using comparable methodology and metrics to those set out in Disclosures (a) and (c) above.

E.   End-to-End Validation of Defendants' Search Methodology and Results

Plaintiffs propose parameters for the parties to meet and confer regarding procedures to validate the effectiveness of Uber's search methods. Uber objects to the inclusion of this section and would omit it entirely. But in light of Uber's obligation under Rule 26(g) to certify complete production, it is appropriate that Uber demonstrates to Plaintiffs that it has made a reasonable inquiry as to the completeness of its production. In light of the anticipated volume and methods of ESI to be searched, Plaintiffs propose a reasonable process for Uber to do so. Similar language has been included in ESI orders in other MDLs. See, e.g., TAR Protocol, In re: Volkswagen "Clean Diesel" Marketing, Sales Practices, and Prods. Liab. Litig., 15-md-02672-CRB, Dkt. 2173 (N.D. Cal. Nov. 7, 2016).

Accordingly, the parties shall adopt Plaintiffs' proposed Section 9, which is as follows:

> The Parties shall participate in an iterative and cooperative approach in which the Parties will meet and confer regarding reasonable and appropriate validation procedures and random sampling of Defendants' Documents (both of relevant and non-relevant sets and of the entire collection against which search terms were run or TAR or other identification or classification methodology was used), 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 manual review was applied which were classified as responsive after Defendants search, TAR and review processes) has been achieved and ensure that the Defendants' search, classification and review methodology was effective and that a reasonable percentage of responsive ESI was identified as responsively being omitted.

F.   Unsearchable Documents

The parties disagree about the language that should govern documents that cannot be searched through text-based means, such as images, spreadsheets, or videos:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 9: Unsearchable Documents	**Plaintiffs' Proposed Section 10: Unsearchable Documents**
To the extent that responsive documents, such as images or spreadsheets, cannot be located through text-based technology, the parties will meet and confer about conducting targeted collections through other means.	Documents which are reasonably believed to be responsive and for which text-based search technologies are fundamentally ineffective, such as images, video, certain spreadsheets, certain hard copy documents, certain documents from noncustodial sources, or certain foreign language documents where the Parties do not have suitable search terms in such language, must be reviewed without culling by search terms, predictive coding, or other technologies that rely primarily on text within the document. Prior to the production of such unsearchable items, the Producing Party may conduct a page-by-page review for responsiveness, confidentiality, privilege, and other protections.

It is more efficient for the Court to direct the parties now regarding unsearchable

1 documents, rather than leaving the issue for a later meet and confer process. Accordingly,

2 the parties shall adopt Plaintiffs' proposed Section 10.

G.   Non-Traditional ESI

Section 10 of Uber's proposed ESI protocol states that, while the ESI protocol is
intended to "address the majority" of ESI handled in this matter, the parties may "come
into contact with more complex, non-traditional or legacy data sources, such as ESI from
social media, ephemeral messaging systems, collaboration tools, data formats identified on
a mobile or handheld device, and modern cloud sources." Uber's Proposed ESI Protocol
§ 10. If that occurred, the parties would agree to "take reasonable efforts to appropriately
address the complexities introduced by such ESI." Plaintiffs' proposal does not include a
comparable provision. Insofar as Uber's proposed language implies that the listed types of
ESI are exempt from the ESI protocol or otherwise due for special treatment, it would
leave too many sources of ESI outside the bounds of the protocol—it would create an
exception that could swallow the rule. Uber's proposed Section 10 shall be excluded.

H.   Reassessment

Section 11 of Plaintiffs' proposed ESI protocol provides for the "reassessment" of
search methods after the search process has been completed, if one of the parties or the
Court "perceive[s] the need" to do so. It further notes that "the time, cost, and/or other
resources expended in connection with ineffective methodologies and/or processes shall be
deemed irrelevant to the issues of reasonableness and proportionality for additional efforts
required." Pls.' Proposed ESI Protocol § 11. Uber's proposal does not contain a
comparable section. The Court agrees with Uber that this section is unnecessary.
Moreover, the Court will not predetermine that any issue related to reasonableness or
proportionality is irrelevant. Accordingly, Plaintiffs' proposed Section 11 shall be
excluded.

I.   Deduplication

The parties disagree over whether certain language should appear in their proposed
sections on "deduplication." The parties' competing proposals are as follows, with

disputed language emphasized:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 12: Deduplication	**Plaintiffs' Proposed Section 13: Deduplication**
The original file paths of a Document prior to deduplication will be populated in the "ALL FILE PATHS" metadata field, separated by semicolons. Hard-Copy Documents shall not be eliminated as duplicates of ESI.	The original file paths of a Document prior to deduplication will be populated in the "ALL FILE PATHS" metadata field, separated by semicolons, ***in the order corresponding to the order of names in ALL CUSTODIANS***. Hard-Copy Documents shall not be eliminated as duplicates of ESI.

Plaintiffs' expert Douglas Forrest attests that the sorting contemplated in the emphasized language above is typical in ESI protocols, and nothing in evidence submitted by Uber addresses this specific issue or contradicts Forrest's assertion. See Forrest Decl. ¶¶ 74–78 (dkt. 261-7). The parties shall adopt Plaintiffs' proposed language in this section of the ESI Protocol.

J.   Cloud Stored Documents

One of the parties' central areas of dispute is the treatment of cloud-based documents, such as Google Docs, that are incorporated into emails or other communications by hyperlink. In essence, the parties' competing proposals on this topic reflect disputes over (1) whether the Producing Party will have to identify the metadata associated with the email and hyperlinked documents; (2) whether the Producing Party will have to produce hyperlinked documents along with the communications that link to them; and (3) whether it will have to produce contemporaneous versions of those documents, or simply have to produce whatever the current version is at the time of production. Plaintiffs argue that Uber can and should use a program called MetaSpike to achieve the types of output Plaintiffs want, while Uber argues that MetaSpike is not a feasible or proportional solution, and that the "Google Parser" tool developed by Lighthouse, Uber's e-discovery vendor, should be used. The proposed language is as follows:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 17: Cloud Stored Documents	**Plaintiffs' Proposed Section 18(a)-(c): Cloud Stored Documents**
Uber will make reasonable and proportionate efforts to preserve the metadata relationship between email messages with links to files on Google Drive, to the extent Uber's vendor for processing and managing the documents to be reviewed and produced in this action possesses technology that enables it to maintain such a relationship. Defendants may use Lighthouse's "Google Parser" for this purpose. Notwithstanding that Uber agrees to make reasonable and proportionate efforts in this regard, because of technological limitations inherent in the processing of emails containing embedded links, it shall not be presumed that all emails containing links to files on Google Drive will be produced with a metadata relationship between the parent email and the linked document. To the extent the Receiving Party believes that there is a lack of a metadata relationship between a specific email message and a specific linked document, the Receiving Party may notify the Producing Party and request that the particular linked file be extracted and produced or identified. To the extent that the linked file in question is nonprivileged, and is relevant to either Party's claims or defenses and the efforts required to search for it would be proportional to the needs of the case, the Producing Party shall use reasonable and proportionate efforts to collect and produce/identify the document that was linked in the original email. The Parties agree to meet and confer to resolve any disagreements as to what constitutes	a) Metadata Preserved. Uber shall preserve the metadata relationship between email messages with links to files on Google Drive. Uber shall preserve and produce (including, if necessary, as custom fields) all metadata collected with respect to all cloud-stored documents. That includes, but is not limited to, all metadata output by Google Vault when exporting a matter. Thus, the metadata exported from Google Vault pertaining to each document shall be preserved and produced as metadata for the same document within the load file of any production containing any such document. b) Hyperlinked/URL-Referenced Documents. Producing party shall make all reasonable efforts to maintain and preserve the relationship between any message or email and any cloud-hosted document hyperlinked or referenced within the message or email. Thus, for instance, where a collected email links to or references by URL a document on Google Drive (or housed within Google vault,) the metadata for that message or email shall include the URLs and Google Document ID of all hyperlinked documents, and if a hyperlinked document was produced. c) Contemporaneous Versions of Hyperlinked/URL-Referenced Documents. Uber shall produce the contemporaneous document version, i.e., the document version likely present at the time an email or message was sent, of Google Drive documents referenced by URL or hyperlinks in emails or messages such as Slack. If Uber contends that it is unable to

18

reasonable and proportionate discovery efforts.	meet this requirement through commercial or vendor software, Google APIs or through other reasonable manual or automated means, then Plaintiffs and Defendants shall meet and confer to discuss solutions. This will not exempt Uber from producing the applicable version of any document so referenced by URL or hyperlink.

Uber asserts that complying with Plaintiffs' proposal will not only be unduly expensive, but that existing technology (including MetaSpike) does not permit it to do so. For support, it points to a number of recent ESI protocols that have declined to treat hyperlinked documents and traditional email attachments the same way. Plaintiffs and their expert disagree that their proposal is infeasible, and they point to several other recent MDL ESI protocols that have adopted their basic approach of treating hyperlinked documents as attachments.

This is a difficult, and highly technical, area of dispute. With respect to Plaintiffs' demand that Uber produce the contemporaneous version of hyperlinked/URL-referenced documents, the evidence that Uber has introduced in support of its position—principally, the declarations of Philip Favro, an e-discovery expert, and Jake Alsobrook, a representative of Uber's e-discovery vendor—speaks generally about the difficulties of automated production of hyperlinked documents, and particularly old versions of those documents, in the Google Workspace environment. See Favro Decl. (dkt. 262-8); Alsobrook Decl. (dkt. 262-9).

Google Vault is a primary concern because much of Uber's ESI is located in Google Vault due to is retention policy. The parties' declarants agree that Google Vault provides functionality to enable users to preview and export earlier versions. *See* Favro Decl. ¶ 22 (dkt. 262-8); Forrest Decl. ¶ 73 (dkt. 261-7). Favro, however, represents that Google Vault does not offer a "scalable process" to enable users to capture both the current version of a document, along with the version contemporaneously exchanged by email. On the other hand, Forrest states broadly that "Google Vault also has an API that should be

1   explored," and he proposes that "macro recorders may enable automation and should also

2   be considered" to the extent that Google Vault requires manual steps to recover a

3   document. None of Uber's declarants specifically address whether a macro is feasible to

4   automate to some extent the process of collecting the contemporaneous versions of

5   hyperlinked/URL-referenced documents within Google Vault.

6   Plaintiffs also assert that Uber can access contemporaneous versions of Google

7   documents sent with e-mails by using MetaSpike's Forensic Email Collector (FEC). ECF

8   No. 261 at 15. Neither Favro, nor Alsobrook, specifically address FEC or the feasibility of

9   deploying it in Uber's data environment or systems. However, a single paragraph in the

10   declaration of Uber's counsel, Caitlin E. Grusaukas, states that in connection with

11   negotiations over the JCCP protocol, unnamed counsel for Uber spoke directly with an

12   unnamed person at Metaspike, and "[o]n information and belief, Metaspike confirmed to

13   Uber's counsel that its FEC software program cannot access items stored in the document

14   retention and archiving system, Google Vault, which Uber uses for Google Workspace

15   data." Grusaukas Decl. ¶ 6 (dkt. 262-1). The vagueness of this paragraph—and the fact

16   that it is made on information and belief, even though it is one Uber attorney's description

17   of a conversation had by another Uber attorney—makes it unhelpful. *See also* Exh. D to

18   Grusaukas Decl. (dkt. 262-5 at 29) ("Metaspike's documentation indicates that FEC only

19   collects Google Drive documents as well."). Nor does Plaintiff's expert clarify the matter.

20   In one statement Forrest seems to concede that FEC is not able to access emails stored in

21   Google Vault. *See* Forrest Decl. ¶ 72(a). In another Forrest suggests that, depending on a

22   variety of factors, FEC may be deployed at some scale to retrieve emails and linked

23   documents in Google Vault. *See id.* ¶ 72(c).

24   In other complex litigation, more detailed information has been requested and

25   provided to explain why such tools are not feasible. See Declaration of Sam Yang, In re:

26   Meta Pixel Healthcare Litigation, No. 22-cv-03580-WHO (VKD), Dkt. No. 265 (N.D. Cal.

27   June 1, 2023); id., Declaration of Jamie Brown, In re: Meta Pixel Healthcare Litigation,

28   No. 22-cv-03580-WHO (VKD), Dkt. No. 266 (N.D. Cal. June 1, 2023); see also Third

United States District Court
Northern District of California

1   Order re Dispute re ESI Protocol, In re: Meta Pixel Healthcare Litigation, No. 22-cv-

2   03580-WHO (VKD), Dkt. No. 267 (N.D. Cal. June 2, 2023) (ruling that "the commercially

3   available tools plaintiffs suggest may be used for automatically collecting links to non-

4   public documents have no or very limited utility in Meta's data environments or systems").

5   Accordingly, and in recognition of the challenging nature of hyperlinks, Uber shall direct

6   an employee with knowledge and expertise regarding Google Vault and Uber's data and

7   information systems to investigate in detail the extent to which Google Vault's API, macro

8   readers, Metaspike's FEC or other programs may be useful to automate, to some extent,

9   the process of collecting the contemporaneous version of the document linked to a Gmail

10   or other communication within Uber's systems, whether the email or communication is

11   stored in Google Vault, or outside. This investigation shall not be limited to documents

12   referenced by URL or hyperlinks in emails or Google documents stored in Google Vault,

13   but shall also include other cloud-based messages such as Slack. Uber's designated

14   employee may consult with Uber's e-discovery experts. Likewise, Plaintiffs shall also

15   more thoroughly investigate these potential solutions.

16   The parties shall complete their investigation by March 22, 2024, and meet and

17   confer with by March 27, 2024 regarding the hyperlinks issue. The parties should also

18   discuss related portions of the ESI protocol, such as the definition of "attachment," the

19   metadata categories in Appendix 2, and Sections 1(e), 4, and 14 of Appendix 1. If there is

20   still disagreement on these issues, the parties may submit a discovery letter pursuant to

21   procedure established under Pretrial Order No. 8. If the parties submit a discovery letter to

22   resolve these issues, Uber's employee designated to conduct its investigation shall submit

23   a declaration supporting its positions, and Plaintiff's expert(s) and/or e-discovery vendors

24   shall do the same. Uber may also submit declarations from its experts and e-discovery

25   vendors.

26   **K.   Continuing Obligations**

27   The parties disagree over certain language in a section regarding the parties'

28   continuing obligations to meet and confer about discovery. The competing proposals are

1   as follows, with disputed language emphasized:

2

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed Section 19(i): Continuing Obligations	**Plaintiffs' Proposed Section 20(i): Continuing Obligations**
The Parties will continue to meet and confer regarding ***ESI*** issues as reasonably necessary and appropriate. This Order does not address or resolve any objections to the Parties' respective discovery requests.	***The Parties recognize that discovery shall be an iterative and cooperative process.*** The Parties will continue to meet and confer regarding ***any*** issues as reasonably necessary and appropriate. This Order does not address or resolve any objections to the Parties' respective discovery requests.

The parties shall adopt Uber's proposed language on this topic as modified by the Court.

The text of this section shall read: "The parties will continue to meet and confer regarding

discovery issues as reasonably necessary and appropriate. This Order does not address or

resolve any objections to the Parties' respective discovery requests."

L.   Appendix 1: Production Format[1]

1.   Family Relationships

The parties agree on the following text, which shall be included in the ESI Protocol:

> Family relationships (be that email, messaging applications, or otherwise) will be maintained in production. Attachments should be consecutively produced with their parent. Objects, documents or files embedded in documents, such as OLE embedded objects (embedded MS Office files, etc.), or images, etc., embedded in RTF files, shall be extracted as separate files and treated as attachments to the parent document. Chats from programs like Slack and HipChat should be produced in families by channel or private message.

Plaintiffs further propose the following language:

> "Attachments" shall be interpreted broadly and includes, e.g., traditional email attachments and documents embedded in other documents (e.g., Excel files embedded in PowerPoint files) as well as modern attachments, internal or non-public documents linked, hyperlinked, stubbed or otherwise pointed to within or as part of other ESI (including but not limited to email, messages,

---

[1] Because certain disputes about language in Appendix 1 depend on the resolution of the cloud-based documents issues, the Court does not address them here. See Part I(J) above.

22

comments or posts, or other documents).

The Court will resolve whether to include this text in the ESI protocol after it decides the disputes raised in Part I(J) above. At that time, the Court will also address the disputes concerning Metadata Fields and the Production of Family Groups and Relationships.

2.   Redactions

The parties disagree over language governing redactions for relevance made to documents that otherwise contain relevant information:

Uber's Proposal	Plaintiffs' Proposal
Uber's Proposed App. 1, Section 17: Redactions	**Plaintiffs' Proposed App. 1, Section 17: Redactions**
Redactions for relevance may only be made where necessary to protect particularly sensitive or proprietary confidential information, and the Parties agree to meet and confer regarding any disputes over the propriety of relevance redactions for any particular document(s) or category(ies) of documents. For redacted items which were originally ESI, all metadata fields will be provided and will include all non-redacted data unless such metadata contains privileged information or information otherwise protected from disclosure. Redacted documents shall be identified as such in the load file provided with the production. A document's status as redacted does not relieve the producing party from providing all of the discoverable metadata required herein.	Other than as permitted by this Order or the Protective Order entered in this Action, no redactions for relevance may be made within a produced document or ESI item. The Parties agree to meet and confer on a case-by-case basis if a Party believes there is a good faith basis to permit limited redaction by agreement of the Parties of highly sensitive, non-relevant information within a Document that contains other relevant information. Any redactions shall be clearly indicated on the face of the document, with each redacted portion of the document stating that it has been redacted and the type of the redaction, and a metadata field shall indicate that the document contains redactions and the type of the redaction (e.g., "Privacy" or "Privilege"). Where a responsive document contains both redacted and non-redacted content, the Parties shall produce the non-redacted portions of the document and the OCR text corresponding to the non-redacted portions.

The parties shall adapt Plaintiffs' proposed language in Appendix 1, Section 17.

3. Mobile and Handheld Device Documents and Data

Plaintiffs propose to include a section concerning the production of data on mobile devices. Uber omits any such provision, and—as already discussed—its proposed protocol contains a separate section regarding "non-traditional ESI," including data from mobile devices. See Part G above. The Court has rejected Uber's non-traditional ESI language, and it similarly finds that Plaintiffs' proposed Section 22 of Appendix 1 is reasonable. Accordingly, Appendix 1 shall contain the following language:

> If responsive and unique data that can reasonably be extracted and produced in the formats described herein is identified on a mobile or handheld device, that data shall be produced in accordance with the generic provisions of this protocol. To the extent that responsive data identified on a mobile or handheld device is not susceptible to normal production protocols, the Parties will meet and confer to address the identification, production, and production format of any responsive documents and data contained on any mobile or handheld device.

4. Parent-Child Relationships

The parties disagree over certain language relating to the production format of documents in a parent-child relationship:

Uber's Proposal Uber's Proposed App. 1, Section 22: Parent-Child Relationships	Plaintiffs' Proposal Plaintiffs' Proposed App. 1, Section 23
The Parties shall use methods of collection and processing that preserve the integrity of document metadata. Except for hyperlinked documents, the parties shall use methods of collection and processing that preserve the parent-child relationships such as the association between attachments and parent documents, or between embedded documents and their parents, or between documents. For documents where the parent-child relationship is produced, all document family relationships shall be produced together and children files should follow parent files in sequential Bates number order. For the avoidance of doubt, a hyperlinked document, such as a cloud-	If responsive, Parent-child relationships that have been maintained in the ordinary course of business shall be preserved for both ESI and hard copy Documents. For example, for electronic production of a hard copy folder with Documents contained in the folder, the cover/title of the folder shall be produced first, with all contents of the folder in sequential Document order behind the containing folder. For email families, the parent-child relationships (the association between emails and attachments) should be preserved, i.e., email attachments should be consecutively produced with the parent email record.

| based document in Google Drive, is not part of parent-child relationship. | |

The parties shall adopt Plaintiffs' proposed language in this section of Appendix 1.

M.   Appendix 2: Metadata

The parties disagree over the inclusion of certain metadata information, particularly with respect to Google Workspace documents. The Court proposes the following resolution, which the Court believes to be reasonable based on the information before it:

1.   Exclude certain fields proposed by Plaintiffs, including (1) "ParticipantPhoneNumbers" and "OwnerPhoneNumbers," which relate to a service, Google Voice, that Uber does not use; (2) "Rfc822MessageId" and "Account," which the Alsobrook Declaration, unrebutted by Plaintiffs, states are duplicative of certain other fields; and (3) "LINKGOOGLEDRIVEURLS," producing which would impose a substantial burden on Uber's vendor and which appears to serve essentially the same function as another field, "LINKGOOGLEDRIVEDOCUMENTIDS," the inclusion of which is undisputed. See Ciaramitaro Decl. (dkt. 261-22); Alsobrook Decl. ¶¶ 13–15 (dkt. 262-9).

2.   Include the remainder of the metadata fields proposed by Plaintiffs, which are generated by Google Workspace applications and which, although they may not be standard information provided by Uber's vendor, can likely be provided without undue burden.
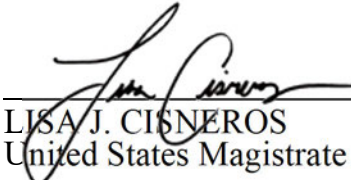
The Court recognizes that the parties have devoted relatively little attention to the metadata field disputes, either in their briefing or their expert declarations. It may also be the case that the utility of certain metadata fields depends on the outcome of cloud-stored documents disputes. Accordingly, the parties may negotiate the metadata issues when they meet and confer about the cloud-stored documents. But in the absence of complete agreement by the parties on an alternative approach or an otherwise compelling justification, the parties shall adopt the solution proposed above.

II.   CONCLUSION

The Court's rulings on the disputed ESI terms above will be incorporated by the Parties into a final Stipulated and [Proposed] ESI Protocol to be submitted to the Court at the appropriate time. The parties shall comply with the instructions regarding further investigations and meet-and-confer regarding the cloud-stored documents issues, metadata fields, and related provisions of the ESI protocol, including completing their investigations by March 22, 2024, and meeting and conferring by March 27, 2024. See Parts I(J), I(M) above. If any disputes remain after March 27, 2024, the parties shall promptly continue to meet and confer and resolve the outstanding issues by no later than April 3, 2024. If disagreements remain, the parties shall begin the process of preparing a joint letter following the procedures set forth in Pretrial Order No. 8, paragraph 3, and the parties' joint discovery letter is due April 12, 2024.

IT IS SO ORDERED.

Dated: 3/15/2024

LISA J. CISNEROS
United States Magistrate Judge

26