

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA

SARAH ANDERSEN, et al.,  
Plaintiffs,  
v.  
STABILITY AI LTD., et al.,  
Defendants.

Case No. [23-cv-00201-WHO](#) (LJC)

**ORDER ON ESI PROTOCOL;  
PROTECTIVE ORDER; DEPOSITION  
PROTOCOL**

Re: Dkt. Nos. 248, 249, 257

**I. INTRODUCTION**

Plaintiffs, a group of visual artists whose works were allegedly used to train a generative artificial intelligence (AI) model called Stable Diffusion, brought this putative class action on behalf other similarly situated artists against Defendants, five companies that created and/or used Stable Diffusion. Among other claims, Plaintiffs allege that Defendants have engaged in widespread copyright infringement. *See generally* ECF No. 238 (Second Amended Complaint). This matter was referred to the undersigned for discovery. ECF No. 250. The parties have been unable to agree on the terms of a stipulated electronically stored information protocol (ESI Protocol) and stipulated protective order (Protective Order) and have submitted competing proposals for both. The parties have also been unable to agree on the number of depositions of Defendants’ fact witnesses that Plaintiffs may conduct.

**II. ESI PROTOCOL**

Judge Orrick’s Standing Order specifies that parties who seek an order on ESI discovery protocols “shall use” one of the Northern District’s models, “unless good cause exists to depart from the model order.” Standing Order ¶ 5. The parties have agreed to use the Northern District’s Model Stipulated Order Regarding Discovery of ESI (Model ESI Order) as a starting point and

United States District Court  
Northern District of California

1 adapt it to fit their needs but disagree about the appropriate extent of modifications. In some  
2 instances, Defendants have advocated for certain language from the district’s Model Stipulated  
3 Order re Discovery of Electronically Stored Information for Patent Litigation (Patent Model ESI  
4 Order). Given the nature of the lawsuit, the Court anticipates that there will be a large volume of  
5 ESI in diverse forms. A detailed ESI protocol, similar to what Plaintiffs have proposed, is  
6 warranted. *See Harden v. Mendocino Coast District Hospital*, 17-cv-05554, 2019 WL 4256383,  
7 at \*1 (N.D. Cal. Sept. 9, 2019) (“[A]n ESI protocol has to be specific or it doesn’t mean  
8 anything.”).

9 **A. Paragraphs 1 and 2**

10 The Court rejects Defendants’ proposed addition to Paragraphs 1 and accepts Defendants’  
11 proposed language to Paragraph 2 that incorporates language from this district’s Model Patent  
12 Order. Thus, Paragraphs 1 and 2 shall read:

13 **1. PURPOSE**

14 This Order will govern discovery of electronically stored information  
15 (“ESI”) in this case as a supplement to the Federal Rules of Civil  
16 Procedure, this Court’s Guidelines for the Discovery of Electronically  
17 Stored Information, and any other applicable orders and rules.

18 **2. COOPERATION**

19 The parties are aware of the importance the Court places on  
20 cooperation and commit to cooperate in good faith throughout the  
21 matter consistent with this Court’s Guidelines for the Discovery of  
22 ESI. As in all cases, costs may be shifted for disproportionate ESI  
23 production requests pursuant to Federal Rule of Civil Procedure 26.  
24 Likewise, a party’s nonresponsive or dilatory discovery tactics are  
25 cost-shifting considerations.

26 **B. Paragraph 4**

27 The Court accepts the parties’ largely agreed upon Paragraph 4. *See* ECF No. 248-3 ¶ 4.  
28 As to Defendants’ modification to Plaintiffs’ addition to Paragraph 4, the Court expects that the  
“reasonable steps” the parties will take to implement litigation holds include establishing  
procedures to make sure that notices of the holds are easily understood, and received by the  
necessary parties and acted upon. *See* ECF No. 249-4 ¶ 4. Accordingly, Paragraph 4 has been  
modified slightly for clarity and shall read:

The parties have an obligation to take reasonable steps to preserve

1 discoverable information in the parties’ possession, custody, or  
2 control, which includes Metadata where applicable, consistent with  
the Federal Rules.

3 The parties that are business entities have issued, or immediately after  
4 this Order has taken effect shall issue, litigation hold notices to those  
5 internal departments, divisions, committees, teams, and individuals  
6 likely to possess potentially responsive information, and to persons or  
7 entities responsible for maintenance of non-custodial sources, likely  
8 to contain potentially responsive information. These business entities  
9 shall establish procedures, if they have not already done so, to ensure  
10 that the notices are easily understood, and have been received and  
11 appropriately acted upon.

12 **C. Paragraph 5 and its subparagraphs**

13 Plaintiffs and Defendants have presented competing proposals under the headings “Search”  
14 and “ESI Search,” respectively. Plaintiffs propose a series of five subparagraphs that address  
15 disclosures, search methodology, search terms, hit reports, and validation. Defendants propose a  
16 single paragraph that focuses on the selection of search terms and search strings. The Court  
17 adopts in part Plaintiffs’ approach because it sets parameters for a series of topics related to ESI  
18 search that will clarify and advance cooperation for purposes of ESI discovery.

19 Plaintiffs’ first proposed subparagraph, Paragraph 5.a is adopted to the extent it is  
20 consistent with this district’s Guidelines for the Discovery of Electronically Stored Information  
21 (ESI Guidelines). The ESI Guidelines encourage the parties to discuss certain ESI-related topics  
22 as part of the required Rule 26(f) meet and confer conference, including the sources, scope, and  
23 type of ESI, as well as date ranges and other details. *See* ESI Guideline 2.02.

24 Plaintiffs’ proposed Paragraph 5.a also requires the disclosure of certain preservation-  
25 related information in writing, namely the specific reasons why categories or sources of  
26 discoverable information have not been preserved or should not be preserved. In contrast,  
27 Defendants’ proposed protocol does not address disclosures or discussions regarding the scope of  
28 preservation. It is well recognized that early discussion of the scope of preservation can reduce  
disputes and the need for court intervention. *See Sedona Principles, Third Edition, Principle 3*  
 (“As soon as practicable, parties should confer and seek to reach agreement regarding the  
preservation and production of electronically stored information.”); *Comment 3.a.* (noting that  
parties may seek to resolve early the identification of data sources that will be subject to

1 preservation, and the relevant time period, among other issues). The Court modifies Paragraph 5.a  
2 to align it with the ESI Guidelines' recommendations and Rule 26(b)(2)(B) of the Federal Rules of  
3 Civil Procedure. At the outset of the discovery process, this protocol does not require Defendants  
4 to detail in writing the specific reasons why categories or sources of discoverable information have  
5 not been preserved or should not be preserved. *See* ESI Guideline 2.01(e).

6 The parties also disagree on the number of custodians. Plaintiffs propose that each party  
7 identify the twenty-four custodians believed to most likely have discoverable ESI, while  
8 Defendants' counterproposal provides that "the parties will identify up to five (5) custodians  
9 believed to most likely have discoverable ESI." In this action there are eight individual named  
10 Plaintiffs<sup>1</sup> and multiple Defendants: Stability AI Ltd. and Stability AI, Inc. (collectively "Stability  
11 AI"), Midjourney, Inc., DeviantArt, Inc., and Runway AI, Inc. On one hand, given the number of  
12 individual named Plaintiffs, Plaintiffs' proposal raises the prospect of a significant, perhaps  
13 excessive number of custodians relative to the number and size of each Defendant and the issues  
14 involved in this case. At the same time, it is not clear from the record the number of employees at  
15 each defendant company, and whether each Defendant is a similarly sized company. On the other  
16 hand, Defendants' counter proposal appears overly restrictive. To resolve the number of  
17 custodians, the Court directs that a Requesting Party may request up to twelve custodians per  
18 Producing Party. Named Plaintiffs are collectively considered a single "Requesting Party," and  
19 "Producing Party." *See Nazemian v. NVIDIA*, 24-cv-01454-JST, Order Regarding Discovery of  
20 Electronically Stored Information (*Nazemian* ESI Order), ECF No. 98 at 3. Stability AI Ltd. and  
21 Stability AI, Inc., collectively Stability AI, is considered a single Requesting and Producing Party.

22 Thus, incorporating the Court's modifications discussed above, Paragraph 5.a, regarding  
23 disclosures shall state:

24 Consistent with the Parties' obligations under the Rules, the Parties  
25 will meet and confer to plan for discovery. This will include a  
26 disclosure of custodians, data sources, anticipated date ranges of  
relevant documents, and categories of documents or ESI that contain  
information relevant to this litigation.

27 \_\_\_\_\_  
28 <sup>1</sup> The Second Amended Complaint names ten individual Plaintiffs, though Kelly McKernan and  
Hawke Southworth later dismissed their claims and withdrew from the action. ECF No. 241.

1 The Parties shall endeavor to reach an agreement regarding the data  
2 sources subject to preservation and discovery. To advance that effort,  
3 the Parties shall disclose (a) categories or sources of relevant  
4 information that are not reasonably accessible because of undue  
5 burden or cost, and (b) categories of data or sources that a party  
6 believes could contain relevant information, but are not discoverable  
7 and should not be preserved under the proportionality factors. A  
8 Requesting Party may request up to twelve custodians per Producing  
9 Party. Named Plaintiffs are collectively considered a single  
10 “Requesting Party,” and a single “Producing Party.” Stability AI Ltd.  
11 and Stability AI, Inc., collectively Stability AI, is considered a single  
12 Requesting and a single Producing Party. The parties may jointly  
13 agree to modify these limits without the Court’s leave. The Parties  
14 may request inclusion of additional custodians as discovery  
15 progresses. The Court shall consider contested requests for additional  
16 custodians, based upon a showing of proportionality to the needs of  
17 the case as that Phrase is used in Rule 26 of the Federal Rules of Civil  
18 Procedure.

19 Next, Plaintiffs’ proposed Paragraph 5.b. addresses search methodology by requiring a  
20 Producing Party to disclose certain information, namely search terms, keywords, date limitations  
21 and custodians, *prior* to using a search methodology to identify documents for production. ECF  
22 No. 249-4 at 4-5. Defendants’ proposed protocol does not contain an equivalent requirement. *See*  
23 ECF No. 248-3 at 4. The parties, however, have agreed that in responding to a request for  
24 production of documents under Federal Rule of Civil Procedure 34, they will “meet and confer  
25 about methods to search ESI[.]” *Id.*; ECF No. 249-4 at 4.

26 Plaintiffs’ proposed Paragraph 5.b does not dictate what methodologies a Producing Party  
27 must use or authorize the Requesting Party to choose the search methodolog(ies). The paragraph  
28 contemplates that by being informed of certain aspects of the search method(s), the Requesting  
Party will be able to independently evaluate the search method being used to collect ESI before it  
is deployed. The mere disclosure of this information does not run afoul of the Sedona principle  
observed in *NuVasive, Inc. v. Alphatec Holdings, Inc.*, No. 18-cv-0347, 2019 WL 4934477, at \*2  
(S.D. Cal. Oct. 7, 2019), which recognized that the responding party is generally best situated to  
evaluate procedures and methodologies appropriate for producing their own ESI. *Id.* The Court’s  
concern, however, is that the strict timing for the disclosure—prior to a party’s use of a search  
method—is an unnecessary, preemptive restraint on responding parties.

The Sedona Conference is a source for guidance related to ESI discovery, and its

1 commentary related to this issue is nuanced. On one hand, the Sedona Conference’s Comment 6.b  
2 explains, “as a general matter, neither a requesting party nor the court should prescribe or detail  
3 the steps that a responding party must take to meet its discovery obligations, and there should be  
4 no discovery on discovery, absent an agreement between the parties, or specific, tangible,  
5 evidence-based indicia (versus general allegations of deficiencies or mere “speculation”) of a  
6 material failure by the responding party to meet its obligations.” The Sedona Principles, *Third*  
7 *Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document*  
8 *Production*, 19 Sedona Conf. J. 1, 123 (2018). At the same time, the Sedona Conference and this  
9 district have recognized that search and retrieval parameters and techniques are appropriate topics  
10 for discussion at an early meet and confer session. *Id.* at 83; ESI Guideline 2.02 (stating that it  
11 may be useful for the parties to discuss, as part of their Rule 26(f) meet and confer conference  
12 “planned methods to identify discoverable ESI and filter out ESI that is not subject to discovery.”).

13 The Court expects cooperation relating to the search and production of ESI, and  
14 “[c]ooperation in reasonably limiting ESI discovery requests on the one hand, and in reasonably  
15 responding to ESI discovery requests on the other hand, tends to reduce litigation costs and delay.”  
16 ESI Guideline 1.02; *see also Comment 3.b.*, 19 Sedona Conf. J. at 76 (“Cooperation among  
17 counsel can enhance the meet and confer process, reduce unnecessary delay and expense  
18 associated with non-merit issues, and foster the overriding objectives of Rule 1.”). Pursuant to the  
19 Federal Rules of Civil Procedure, the parties are required to confer on and prepare a discovery  
20 plan, and their conference must address ESI issues. Fed. R. Civ. P. 26(f)(3)(C). The Northern  
21 District’s accompanying Checklist for Rule 26(f) meet and confers regarding ESI guides the  
22 parties to discuss “search method(s), including specific words or phrases or other methodology,  
23 that will be used to identify discoverable ESI and filter out ESI that is not subject to discovery,” as  
24 well as the “time period during which discoverable information was most likely to have been  
25 created or received” and custodians.<sup>2</sup> The Checklist does not specify “the timing of the discussion  
26 of the discussion about these topics, [which] may depend on the nature and complexity of the  
27

28 <sup>2</sup> [https://cand.uscourts.gov/filelibrary/1118/ESI\\_Checklist-12-1-2015.pdf](https://cand.uscourts.gov/filelibrary/1118/ESI_Checklist-12-1-2015.pdf)

1 matter.”

2 To strike the appropriate balance, the Court modifies Plaintiffs proposed Paragraph 5.b as  
3 follows:

4 At the appropriate time, a Party that intends to use or is using a search  
5 methodology (if any) to identify documents for production, shall  
6 disclose to the Requesting Party, the search terms, keywords, date  
7 limitations, and custodians, and shall specify what custodial and non-  
8 custodial sources that the search methodology will or was run against.  
9 The custodial sources will presumptively be those identified in the  
disclosures set forth in Paragraph 5.a. The parties shall discuss search  
methods at an early meet and confer conference, including the  
retrieval parameters identified in the preceding sentence, to engage in  
meaningful cooperation that reduces the costs and risk associated with  
ESI production.

10 While the early meet and confer conference may not entail an  
11 exhaustive discussion of every search to be deployed, the Requesting  
12 Party may, at a later point in the discovery process, request a meet and  
13 confer to address the adequacy of a search method or search term, if  
14 there is a specific, tangible concern that the Requesting Party has  
15 identified to the Producing Party. In such instances, the Producing  
16 Party shall provide to the Requesting Party any information that it has  
17 not already provided that is necessary to evaluate the effectiveness of  
18 the method or the search, such as the search term hit reports pursuant  
19 to 5.d.

20 At the early meet and confer conference where the parties address ESI  
21 search methods, the parties shall also discuss what method(s) will be  
22 used for documents that are reasonably believed to be responsive and  
23 for which text-based search technologies are fundamentally in  
24 effective, such as images, spreadsheets, hard copy documents, or  
25 certain foreign language documents where the parties do not have  
26 suitable search terms in such language. The Producing Party shall  
27 disclose to the Requesting Party the specific method(s) it intends to  
28 use to search for such documents, and the Requesting Party may make  
a reasonable request for information that relates to the method(s)  
effectiveness.

Should a Producing Party wish to use any method or methodology not  
specifically addressed by this Order to exclude documents from any  
review or production, notice to the Receiving Party must be made  
prior to the use of any such method or methodology. If a Party intends  
to use Technology Assisted Review (“TAR”) (including Continuous  
Active Learning), the Parties will agree to a separate TAR protocol to  
govern that process. If the Parties are unable to reach agreement on a  
TAR protocol, they will submit any disputed issues to the Court or its  
designee for resolution.

Plaintiffs’ proposed Paragraphs 5.c and 5.d address search terms and hit reports,  
respectively. In Paragraph 5.c Plaintiffs propose that the parties shall meet and confer to discuss

1 the Producing Party’s proposed search terms and methodology, and prior to or during such meet  
2 and confers, the Producing Party shall provide all associated information required by the Court.  
3 Defendants do not appear to object to meeting and conferring regarding search terms: they have  
4 agreed to meet and confer about methods to search ESI; they propose a sentence stating, “The  
5 Court encourages the parties to confer on a process to test the efficacy of the search terms[;]” and  
6 they agree that the parties may agree jointly to modify limits on the number of search terms that  
7 the Court may impose. Defendants appear amenable to meeting and conferring regarding the  
8 selection of search terms and to discuss their efficacy. For clarity, the Court adopts Plaintiff’s  
9 proposed language in Paragraph 5.c requiring that the parties shall meet and confer to discuss the  
10 Producing Party’s proposed search terms and methodology. The Court also adopts Defendant’s  
11 proposed language, with certain modifications, concerning the tailoring of search terms.

12 The parties’ proposals diverge with respect to establishing a requirement that the  
13 Producing Party provide search term hit reports in certain circumstances, and limiting the number  
14 of search terms. Plaintiffs’ proposed Paragraph 5.c provides that prior to or during meet and  
15 confers addressing search terms, “the Producing Party shall provide all associated information  
16 required by this Order.” The proposed Paragraph 5.b states, “At all meet and confers where the  
17 topic is adequacy of search terms to identify documents subject to production, the Producing Party  
18 shall be required to provide search term hit reports pursuant to paragraph 5.D, below.”  
19 Defendants’ proposed protocol does not address hit reports.

20 The effectiveness of meet and confers depends on the parties’ candor, and this is especially  
21 true when the parties are addressing potential search terms. Defendants are evidently open to  
22 engaging in a process to test the efficacy of the search terms, but their proposal does not offer an  
23 alternative to Plaintiffs’ position that the Producing Party provide hit reports containing certain  
24 data points. Accordingly, the Court accepts Plaintiffs’ proposal that hit reports be provided as part  
25 of the meet and confer process regarding search terms.

26 Turning to the number of search terms, Defendants propose that each requesting party limit  
27 its ESI production requests to five search terms per custodian per party—the same limited  
28 imposed by the Patent Model ESI Order. ECF No. 248-3 at 4. Plaintiffs do not propose to limit

United States District Court  
Northern District of California

1 the number of search terms that a requesting party may use. Defendants, however, have not  
2 persuaded the Court that the default limit imposed by the Patent Model ESI Order ought to apply  
3 to this case. Instead, the Court sets a limit of fifteen search terms per custodian per Party, which  
4 was the limit in the *Nazemian* ESI Order.

5 In light of the Court’s determinations above, Paragraph 5.c shall read as follows:

6 Where the Parties agree that potentially responsive ESI shall be  
7 searched using search terms, the Parties shall meet and confer to  
8 discuss the Producing Party’s proposed search terms and  
9 methodology. Prior to or during such meet and confers, the Producing  
10 Party shall provide all associated information required by this Order.  
At all meet and confers where the topic is adequacy of search terms  
to identify documents subject to production, the Producing Party shall  
provide search term hit reports pursuant to Paragraph 5.d, unless the  
parties agree otherwise.

11 Each Requesting Party shall limit its ESI production requests to a total  
12 of fifteen search terms per custodian per party. The search terms shall  
13 be narrowly tailored to particular issues. Indiscriminate terms, such  
14 as the producing company’s name or its product name, which are  
15 inappropriate unless combined with narrowing search criteria that  
16 sufficiently reduce the risk of overproduction. A conjunctive  
17 combination of multiple words or phrases (e.g., “computer” and  
“system”) narrows the search and shall count as a single search term.  
A disjunctive combination of multiple words or phrases (e.g.,  
“computer” or “system”) broadens the search, and thus each word or  
phrase shall count as a separate search term unless they are variants  
of the same word. Use of narrowing search criteria (e.g., “and,” “but  
not,” “w/x”) is encouraged to limit the production.

18 If, after disclosure of the Producing Party’s proposed search method,  
19 search parameters, and search terms, and after a reasonable meet-and-  
20 confer process, a Requesting Party has a specific, tangible, evidence  
21 based reason to believe that the Producing Party’s proposals regarding  
22 search, retrieval, and production would result in deficiencies in  
23 production, the Requesting Party may object to the proposal and make  
24 requests for different or additional search methods, parameters, or  
search terms. Such requests shall only be made after the Parties have  
met and conferred as to the alleged deficiencies identified by the  
Requesting Party. If the issue is not resolved within fourteen (14)  
calendar days after the objection, either Party may thereafter  
immediately submit the dispute to the Court or its designee for  
resolution.

25 Plaintiffs’ proposed Paragraph 5.d is adopted in full.

26 Plaintiffs’ proposed Paragraph 5.e states that the parties “will meet and confer about  
27 methods to validate each Producing Party’s review process[.]” ECF No. 249-4 at 6. Defendants’  
28 proposed ESI protocol does not address validation. It is appropriate for the ESI protocol to

1 meaningfully address validation procedures, and the Court modifies Plaintiff’s proposed  
2 Paragraph 5.e as follows, modified from *In re Social Media Adolescent Addiction*, 2024 WL  
3 1786293, at \*6-7:

4           Each Producing Party shall take reasonable steps to validate its own  
5 review process (i.e., using quality control measures to determine  
6 whether its production is missing relevant ESI or contains substantial  
7 amounts of irrelevant ESI) and shall take necessary adjustments to its  
8 process. After a Requesting Party reviews a Producing Party’s  
9 production, the Requesting Party may request additional information  
10 regarding the validation method(s) used by the Producing Party and  
11 the Producing Party shall disclose the level of end-to-end recall (the  
12 percentage of responsive documents in the collection that were  
13 identified as responsive by the Producing Party’s methodology). If  
14 there are remaining disputes after the Producing Party discloses the  
15 level of end-to-end recall, the parties shall meet and confer in good  
16 faith to resolve any remaining disputes. Such meet and confer efforts  
17 shall include a discussion of the tool and parameters used to calculate  
18 recall levels.

12           **D. Paragraph 6**

13           Plaintiffs propose adding a paragraph addressing “Enterprise messaging, mobile device  
14 data, and other communication tools.” ECF No. 249-4 ¶ 6. The Court recognizes that the parties’  
15 messaging platforms used during the course of business (Slack, Microsoft Teams, Telegram, etc.)  
16 and mobile devices may contain responsive ESI and accordingly approves the addition of  
17 Plaintiffs’ Paragraph 6. *Id.*

18           **E. Paragraph 7 and 8**

19           The Court approves Plaintiffs’ proposed edits to Paragraph 7. ECF No. 249-4 ¶ 7. It  
20 likewise approves the parties’ agreed-upon changes to Plaintiffs’ Paragraph 8 (Defendants’  
21 Paragraph 7) addressing phasing ESI production. *Id.* ¶ 8; 248-3 ¶ 7.

22           **F. Paragraph 9**

23           Plaintiffs, at Paragraph 9, and Defendants, at Paragraph 8, propose adding specifications  
24 about deduplication. Defendants’ proposal includes more permissive language (“the Producing  
25 Party *shall take reasonable steps to produce* only one such identical copy”) while Plaintiffs’  
26 proposal is more direct (“the Producing Party *shall produce* only one such identical copy”) but the  
27 proposals are otherwise the same. ECF No. 249-3 ¶ 8. The Court directs the parties to use the  
28

1 language in Defendants' proposed Paragraph 8.

2 **G. Paragraphs 10 and 11**

3 Plaintiffs propose protocols for handling embedded files (Paragraph 10). Defendants  
4 object to producing embedded files, asserting that such an obligation would be unduly  
5 burdensome. Defendants, however, do not explain why such a production would be unduly  
6 burdensome when embedded files are included in ESI production in several other comparable  
7 cases. *See e.g., Nazemian ESI Order at 7-8; In re Uber Techs., Inc., Passenger Sexual Assault*  
8 *Litig.*, No. 23-md-03084-CRB-LJC, 2024 WL 1772832, at \*6 (N.D. Cal. Apr. 23, 2024); *In re*  
9 *Facebook, Inc. Consumer Priv. User Profile Litig.*, No. 18-md-2843 VC (JSC), 2021 WL  
10 10282215, at \*30 (N.D. Cal. Sept. 29, 2021). The Court grants Plaintiffs' request to add  
11 Paragraph 10, requiring the production of embedded files.

12 Plaintiffs propose protocols for handling the production of hyperlinked documents  
13 (Paragraph 11). Defendants' arguments misapprehend Plaintiffs' proposal. Defendants claim that  
14 Plaintiffs' proposal would require the production of any hyperlinks that appear in any document or  
15 communication, along with metadata connecting the document in native format to the document.  
16 Plaintiffs, however, appear to acknowledge, to some extent, that technological limitations will  
17 restrict which hyperlinks may be "automatically produced," and propose that upon a reasonable  
18 request by the Receiving Party, the Producing Party will produce such hyperlinked documents.  
19 ECF No. 249-4 at 8. In addition, Plaintiffs' proposal calls upon the Producing Party to use  
20 "available technology" to match hyperlinked documents with email or familial documents, and the  
21 production of related metadata, e.g. BEGATTACH and ENDATTACH, would only be required  
22 where the technology is available to extract such information. Plaintiffs' proposal does not appear  
23 as onerous as Defendants contend, though it leaves out important details, such as how many  
24 hyperlinks may be requested by a receiving party.

25 To enhance the ESI protocol's clarity, the Court adopts, with some modifications, the  
26 provisions for hyperlinks, the production of hyperlinked documents, and hyperlink metadata from  
27 *Nazemian ESI Order at 8-9 (Section 16)*. Accordingly, Paragraph 11 and its subparagraphs shall  
28 read as follows:

**11. HYPERLINKS**

1 Except as provided in this paragraph and its subparagraphs, the  
2 Producing Party shall not be required to produce documents solely  
3 because they are hyperlinked within the Producing Party’s  
4 production. For purposes of this protocol, “hyperlinked documents”  
include but are not limited to documents shared via Google  
Workspace, Microsoft Office’s “Share Documents via Link” feature,  
or similar cloud-based document storage and sharing platforms.

5 The Parties agree to meet and confer in good faith to resolve any  
6 disputes concerning the appropriateness of production of the  
7 requested documents and metadata. Nothing in this paragraph and its  
8 subparagraphs shall prevent the Producing Party from producing  
hyperlinked documents that are relevant or responsive to discovery  
requests in addition to the documents produced by specific  
request by the Receiving Party.

**a. PRODUCTION OF HYPERLINKED DOCUMENTS**

9 The Parties may request the production of the current version or  
10 contemporaneous version of a responsive and non-privileged  
11 document or communication that is within the Producing Party’s  
12 possession, custody, or control and is hyperlinked within the  
13 Producing Party’s production by providing to the Producing Party a  
14 list of hyperlinks and corresponding Bates numbers in which the  
15 hyperlinked document is referenced. If the hyperlinked document has  
16 already been produced, the Producing Party may alternatively provide  
17 the Receiving Party with the document’s Bates number. Each Party  
18 (with Named Plaintiffs constituting a single Party, and Stability AI  
19 entities constituting a single Party) may make up to two hundred such  
20 requests. The parties agree to meet and confer in good faith to resolve  
any disputes concerning the appropriateness of the production of the  
requested documents and communications and any reasonable request  
to increase the limit on the total number of hyperlinked documents  
and communications to be produced. If a Requesting Party believes  
that more requests are warranted and the Receiving Party has not  
agreed to produce additional hyperlinked documents, the Requesting  
Party may initiate the joint letter-brief process pursuant to the Court’s  
dispute resolution procedures set forth in Standing Order, Section F.5.

**b. PRODUCTION OF HYPERLINK METADATA**

21 The Parties acknowledge that the current version of the hyperlinked  
22 document (i.e., the version produced pursuant to the paragraphs  
23 concerning hyperlinks) may differ from the version of the document  
24 at the time the hyperlink was sent. The Parties also acknowledge that  
25 the metadata of such document(s) may not reflect the metadata for  
26 requested documents at the time the documents were hyperlinked.  
27 For hyperlinked documents produced pursuant to these paragraphs  
28 concerning hyperlinks, the Producing Party will use available  
technology to match hyperlinked documents with the email or  
familial document to which the documents were attached in Native  
format, and the Producing Party will use available technology to  
provide the BEGATTACH and ENDATTACH metadata fields to  
show the family relationship.

**H. Paragraph 12**

At Paragraph 12, Plaintiffs propose adding specifications about how a producing party must collect and produce structured data, including requirements that the producing party inform the receiving party about all possible formats for the structured data, and, if parties cannot agree on the method of production, the producing party “make the information available for inspection on this system where it is currently maintained, or otherwise provide direct access to the data.” ECF No. 249-4 ¶ 12. Defendants argue that such process could provide the receiving party “unfettered access to their IT infrastructure with no specific showing of burden, need, or proportionality[.]” ECF No. 248 at 5. Plaintiffs recognize that “it is important for the parties to discuss at the outset how [structured data] should be handled.” ECF No. 249-1 (Saveri Decl.) ¶ 12. It is unclear if the parties have engaged in any meaningful discussion about the production of structured data. The parties are therefore ordered to meet and confer to develop a stipulated protocol for producing and reviewing structured data.

**I. Paragraph 13**

The Court accepts Plaintiffs’ addition of Paragraph 13 to the extent that Defendants have not proposed deleting certain text. ECF Nos. 249-4 ¶ 4; 248-3 ¶ 9. The examples of how Model Data may be produced, set forth in Plaintiffs’ proposed Paragraph 13, do not obligate Defendants to produce Model Data in those formats and the inclusion of this sentence do not waive any arguments Defendants might wish to make about producing Model Data. The parties may discuss those examples during their meet and confer conference regarding model data. Paragraph 13 shall read as follows:

To the extent such data is not addressed in this protocol, the Parties will meet and confer to address the production of data related to the Models at issue. Given the size, structure and format of such data, the Parties agree to discuss how Model Data, if any, should be produced.

**J. Paragraphs 14 and 15**

Plaintiffs propose adding detailed specifications about what information must be included in privilege logs and a procedure for challenging asserted privileges. ECF No. 249-4 ¶¶ 14-15. The Court agrees that not establishing clear requirements for privilege logs and privilege log disputes at the outset of discovery will likely invite future disputes. Defendants have objected to

1 Plaintiffs’ proposed specifications for the privilege logs, but these requirements are in line with  
2 privilege log requirements in other cases, *see e.g.*, *Nazemian* ESI Order at 9-11; *In re Uber Techs.,*  
3 *Inc., Passenger Sexual Assault Litig.*, No. 23-md-03084-CRB-LJC, at ECF No. 396 (¶¶ 14-17),  
4 and Defendants fail to cite to any specific authority that discourages against these provisions. The  
5 Court accepts Plaintiffs’ proposal for privilege logs and the procedure for challenging asserted  
6 privileges at Paragraphs 14 and 15, respectively.

7 **K. Paragraph 16**

8 The parties agree regarding “DOCUMENTS PROTECTED FROM DISCOVERY”. *See*  
9 Plaintiffs Paragraph 16, ECF No. 249-4 at 12; Defendants’ Paragraph 11, ECF No. 248-3 at 6.

10 **L. Appendices**

11 Plaintiffs have proposed two appendices: Appendix A, which sets forth technical  
12 specifications for the production of ESI, and Appendix B, which addresses metadata fields.  
13 Defendants object to both, though they only provide a counterproposal for metadata fields.  
14 Defendants make only a conclusory argument that Plaintiffs’ proposal for technical specifications  
15 for ESI production is onerous. Therefore, the Court adopts Plaintiffs’ proposed Appendix A, ECF  
16 No. 249-4 at 15-17.

17 With regard to the appendix for metadata fields, given Defendants’ representation that  
18 complying with Plaintiffs’ proposal would require manual entry, be “technically onerous[,]” and  
19 likely to result in error, the Court directs the parties to use Defendants’ metadata fields. ECF Nos.  
20 248 at 5; 248-3 at 9-11. This ruling is made without prejudice and Plaintiffs may renew their  
21 request for the additional fields upon showing that their need outweighs the burden on Defendants.

22 **III. PROTECTIVE ORDER**

23 The parties’ only outstanding dispute regarding their protective order is the definition of  
24 “source code.” *See* ECF No. 248-5 at 4. Plaintiffs define source code as “the human readable set  
25 of instructions written in a programming language that defines the operations of a software  
26 application, including the algorithms, logic, and functions that shape the program’s behavior.”  
27 ECF No. 249-3 § 2.18. Defendants seek to use the definition provided in the Patent Model  
28 Protective Order, which defines source code as “computer code and associated comments and

1 revisions histories, formulas, engineering specifications, or schematics that define or otherwise  
2 describe in detail the algorithms or structure of software or hardware designs, disclosure of which  
3 to another Party or Non-Party would create a substantial risk of serious harm that could not be  
4 avoided by less restrictive means.” ECF No. 248-2 § 2.9; *Model Protective Order for Litigation*  
5 *Involving Patents, Highly Sensitive Confidential Information and/or Trade Secrets*, U.S. District  
6 Court, Northern District of California, [https://www.cand.uscourts.gov/forms/model-protective-](https://www.cand.uscourts.gov/forms/model-protective-orders/)  
7 [orders/](https://www.cand.uscourts.gov/forms/model-protective-orders/) (last visited Jan. 22, 2025). Although the parties are not required to use the Patent Model  
8 Protective Order definition, the Court agrees with Defendants that using the broader definition  
9 provided in the Patent Model Protective Order is appropriate here. *See In re Meta Pixel*  
10 *Healthcare Litigation*, No. 22-cv-03580, 2023 WL 3483296, at \*1 (N.D. Cal. May 16, 2023)  
11 (adopting a definition of “source code” that “generally tracks” the language in the Patent Model  
12 Protective Order). The Court recognizes that “comments and revisions histories, formulas,  
13 engineering specifications, or schematics” that define or describe code may be as sensitive—and  
14 thus warrant the same safeguards—as the code itself. ECF No. 248-2 § 2.9. The parties are  
15 ordered use the Patent Model Protective Order’s definition of source code.

#### 16 **IV. DEPOSITION PROTOCOL**

17 The parties submitted an additional dispute regarding their deposition protocol. ECF No.  
18 257. Plaintiffs request to conduct up to 60 depositions of Defendants’ fact witnesses. Defendants  
19 object, countering that each side should be entitled to take up to twelve depositions.

20 Federal Rule of Civil Procedure 30(a) provides that a party may take up to ten depositions  
21 without leave of the court. Courts “may alter the limits ... on the number of depositions ... or on  
22 the length of depositions under Rule 30.” Fed. R. Civ. P. 26(b)(2)A). In determining whether to  
23 permit more than ten depositions, the court must consider “whether: (i) the discovery sought is  
24 unreasonably cumulative or duplicative, or is obtainable from some other source that is more  
25 convenient, less burdensome, or less expensive; (ii) the party seeking discovery has had ample  
26 opportunity by discovery in the action to obtain the information sought; or (iii) the burden or  
27 expense of the proposed discovery outweighs its likely benefit[.]” *Authentec, Inc. v. Atrua Techs.,*  
28 *Inc.*, No. C 08-1423 PJH, 2008 WL 5120767, at \*1 (N.D. Cal. Dec. 4, 2008). Although “courts

1 will generally not grant leave to expand the number of depositions until the moving party has  
2 exhausted the ten depositions permitted as of right” (*id.*), the Northern District regularly departs  
3 “from this rule where the complexity of the case clearly warranted more than ten depositions.”  
4 *Sayce v. Forescout Techs., Inc.*, No. 20-cv-00076, 2024 WL 3579002, at \*1 (N.D. Cal. July 11,  
5 2024); *see Del Campo v. American Corrective Counseling Servs., Inc.*, No. C-01, 2007 WL  
6 3306496 (N.D. Cal. Nov. 6, 2007); *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417, 2024 WL  
7 4502099, at \*2 (N.D. Cal. Oct. 15, 2024) (permitting Plaintiff to take twenty-five fact depositions  
8 in copyright dispute regarding generative AI). “A party seeking to exceed the presumptive  
9 number of depositions must make a particularized showing of the need for the additional  
10 discovery.” *Authentec*, 2008 WL 5120767, at \*1.

11 Plaintiffs argue they are entitled to sixty depositions because of “the number of corporate  
12 defendants, the technical nature of the infringement, and the vast scale of the alleged misconduct.”  
13 ECF No. 257 at 1. They reason that taking sixty depositions, or fifteen per defendant, is  
14 reasonable given the scope of the case. *Id.* at 2. This will permit them to take twenty to twenty-  
15 five depositions of “core technical witnesses”, fifteen to twenty depositions of witnesses in  
16 corporate leadership roles (including “executive leadership, project managers, and strategic  
17 decision-makers”), and fifteen to twenty depositions of Defendants’ “business operations  
18 personnel.” *Id.* at 2-3. Defendants object, arguing that Plaintiffs’ request “lacks any particularized  
19 justification” and that Plaintiffs’ “proposed breakdown of total deponents...is arbitrary and  
20 untethered to any document, filing, or fact specific to any Defendant.” *Id.* at 4. They contend that  
21 most of the Defendant companies are “early-stage, lean startups” with relatively few employees,  
22 and permitting Plaintiffs to depose “extraneous employees” would result in duplicative testimony  
23 and pose a significant burden on Defendants. *Id.* at 4-5.

24 The Court agrees with Plaintiffs’ overall position that exceeding the ten-deposition limit is  
25 warranted here given the “complexity of the case[.]” but is mindful of the potential burden on  
26 Defendants and the risk that allowing Plaintiffs sixty depositions would be unreasonably  
27 cumulative or duplicative. *Sayce*, 2024 WL 3579002, at \*1; *see* Fed. R. Civ. P. 26(b). The Court  
28 finds it appropriate at this time to grant Plaintiffs leave to take thirty depositions, to be allocated at

United States District Court  
Northern District of California

1 Plaintiffs’ discretion. Plaintiffs may renew their request for additional depositions if they  
2 determine they are necessary after taking the first thirty depositions. The thirty-deposition limit  
3 does not include depositions of non-parties, expert witnesses, of former employees. This limit  
4 does not change the parties’ agreement regarding 30(b)(6) depositions at ¶ 22 of the Stipulated  
5 Order Entering Deposition Protocol. ECF No. 257-1 ¶ 22.


6 **V. CONCLUSION**

7 Having resolved the disputes discussed above concerning the terms of the ESI protocol, the  
8 parties are ordered to incorporate the Court’s rulings from this Order into a final Stipulated and  
9 [Proposed] ESI Protocol, to be filed with the Court by March 31, 2025. Before March 31, 2025,  
10 the parties shall also meet and confer regarding the outstanding issues related to structured data  
11 and model data. If the parties are able to reach an agreement to resolve these issues, they shall  
12 include the necessary provisions in the final Stipulated and [Proposed] ESI Protocol. If parties are  
13 unable to reach an agreement they shall file by April 7, 2025 a joint discovery letter concerning  
14 their respective proposals, following the procedure in Section F.5 of the undersigned’s Standing  
15 Order.

16 The parties shall file a Stipulated Protective Order, consistent with this Order, by March  
17 31, 2025.

18 **IT IS SO ORDERED.**

19 Dated: 3/19/2025

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