

EXHIBIT L

DUNN ISAACSON RHEE LLP
KAREN L. DUNN (*pro hac vice*)
kdunn@dirllp.com
JESSICA E. PHILLIPS (*pro hac vice*)
jphillips@dirllp.com
KYLE N. SMITH (*pro hac vice*)
ksmith@dirllp.com
401 9th Street NW
Washington, DC 20004
Telephone: (202) 240-2900

CLEARY GOTTlieb STEEN &
HAMILTON LLP
ANGELA L. DUNNING (212047)
adunning@cgsh.com
151 University Avenue
Palo Alto, CA 94301-1632
Telephone: (650) 815-4131

Counsel for Defendant Meta Platforms, Inc.

COOLEY LLP
MARK R. WEINSTEIN (193043)
mweinstein@cooley.com
3175 Hanover Street
Palo Alto, CA 94304-1130
Telephone: (650) 843-5000

PHILLIP E. MORTON (*pro hac vice*)
pmorton@cooley.com
1299 Pennsylvania Avenue NW, Suite 700
Washington, DC 20004
Telephone: (202) 842-7800

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

CAMBRONNE, INC., et al.

Plaintiff,

v.

META PLATFORMS, INC.,

Defendant.

Case No. 3:26-cv-03725-VC

**DEFENDANT META PLATFORMS, INC.'S
SECOND SUPPLEMENTAL RESPONSES AND
OBJECTIONS TO PLAINTIFF'S FIRST SET OF
INTERROGATORIES**

Judge: Hon. Vince Chhabria

PROPOUNDING PARTY: PLAINTIFFS CAMBRONNE INC., LISA BARRETTA, PHILIP SHISHKIN, JANE ADAMS, MATTHEW SACKS, AND MICHAEL KOCHIN

RESPONDING PARTY: DEFENDANT META PLATFORMS, INC.

SET NUMBER: ONE

Pursuant to Federal Rules of Civil Procedure Rules 26 and 33, and Local Rule 33 of the United States District Court for the Northern District of California, Defendant Meta Platforms, Inc. (“Meta”) hereby objects and provides these supplemental responses to Plaintiffs Cambronne Inc., Lisa Barretta, Philip Shishkin, Jane Adams, Matthew Sacks, and Michael Kochin’s (“Plaintiffs”) First Set of Interrogatories (“Interrogatories”).

I. Responses to All Interrogatories

1. Meta’s responses to these Interrogatories are made to the best of Meta’s present knowledge, information, and belief. Said responses are at all times subject to such additional or different information that discovery or further investigation may disclose, and while based on the present state of Meta’s recollection, are subject to such refreshing of recollection, and such additional knowledge of facts, as may result from Meta’s further discovery or investigation. Meta reserves the right to make any use of, or to introduce at any hearing and at trial, information and/or documents responsive to these Interrogatories but discovered subsequent to the date of these responses, including, but not limited to, any such information or documents obtained in discovery herein.

2. To the extent that Meta responds to an Interrogatory by stating that Meta will provide information or documents that Meta deems to embody material that is private, business confidential, proprietary, trade secret, or otherwise protected from disclosure pursuant to Federal Rule of Civil Procedure 26(c) and/or Federal Rule of Evidence 501, Meta will only do so subject to the Protective Order (Dkt. 53) governing the unauthorized use or disclosure of such information or documents with a designated “CONFIDENTIAL” or “HIGHLY CONFIDENTIAL - ATTORNEYS’ EYES ONLY” or “HIGHLY CONFIDENTIAL – SOURCE CODE” distinction.

3. Meta reserves all objections or other questions as to the competency, relevance, materiality, privilege or admissibility as evidence in any subsequent proceeding in or trial of this

1 or any other action for any purpose whatsoever of Meta's responses herein and any document or
2 thing identified or provided in response to the Interrogatories.

3 **II. Objections to Instructions and Definitions**

4 Whether or not separately set forth in response to each Interrogatory, Meta makes these
5 objections to the following Instructions and Definitions.

6 1. Meta objects to the definition of "COPYRIGHT MANAGEMENT
7 INFORMATION" to the extent it calls for a legal conclusion. Specifically, a determination of
8 whether any particular data, text, or metadata constitutes "Copyright Management Information"
9 under 17 U.S.C. § 1202(c) is inherently legal in nature and calls for a legal conclusion. Meta is not
10 in a position to know whether any particular information qualifies as legally protected Copyright
11 Management Information and a determination of such is not within Meta's possession, custody, or
12 control. Meta further objects to this definition as vague and ambiguous, overbroad, unduly
13 burdensome, and disproportionate to the needs of the case to the extent it includes identifying
14 information for persons that "created" copyrighted works.

15 2. Meta objects to the definition of "PLAINTIFFS' WORK(S)" as overbroad, unduly
16 burdensome and disproportionate to the needs of the case to the extent it puts the burden on Meta
17 to determine which works purportedly "belong" to PLAINTIFFS. Meta further objects to the term
18 "belong" as vague and ambiguous as it is unclear what form of possession or ownership is intended
19 by the use of this term. Without agreeing that PLAINTIFFS possesses or owns any valid copyright
20 in any specific works, Meta will interpret the term PLAINTIFFS' WORKS to refer to the specific
21 works identified in Exhibit A to the Interrogatories.

22 3. Meta objects to the definitions of "COMMUNICATIONS" and "DOCUMENT(S)"
23 to the extent these definitions are inconsistent with and otherwise seek to circumvent the custodian
24 and search term limits for electronic communications (including emails and other electronic
25 correspondence, and documents attached thereto), as provided in the ESI Order.

26 4. Meta further objects to the definitions of "COMMUNICATIONS" and
27 "DOCUMENT(S)" to the extent they seek to expand or alter the scope of the Federal Rules of Civil
28 Procedure, the Local Rules of this Court, or the ESI Order. Meta further objects to these definitions

1 to the extent they seek Meta to produce documents that are not reasonably accessible because of
2 undue burden or cost pursuant to Fed. R. Civ. P. 26(b)(2)(B), which are unduly burdensome to
3 preserve and search, or which otherwise, under the proportionality factors, should not be
4 preserved.

5 **5.** Meta objects to the definitions of “CORPORA” and “DATASET(S)” as overbroad,
6 unduly burdensome, and disproportionate to the needs of the case, including (but not limited to) the
7 extent they refer to non-textual materials that are not relevant to any party’s claims or defenses and
8 to the extent they refer vaguely to “other content.” Without any admission with respect to content
9 or nature, Meta will construe the term “DATASET(S)” to refer to Anna’s Archive, Bibliotik,
10 Books3, C4, Library Genesis (or LibGen), Sci-Hub, The Pile, and Z-Library.

11 **6.** Meta objects to the definition of “INFRINGE,” “INFRINGED,” and “INFRINGES”
12 to the extent it calls for a legal conclusion. Specifically, a determination of whether any conduct
13 constitutes copyright infringement is inherently legal in nature and calls for a legal conclusion.
14 Meta is not in a position to know whether any particular conduct infringes on a legally protected
15 work and a determination of such is not within Meta’s possession, custody, or control.

16 **7.** Meta objects to the definition of “LARGE LANGUAGE MODEL(S)” and “LLM
17 MODEL(S)” as vague and ambiguous, overbroad, unduly burdensome, and disproportionate to the
18 needs of the case. In particular, the phrase “any machine learning model, neural network,
19 generative artificial intelligence model, transformer model, or similar computational system trained
20 on large quantities of text or other data” is overly broad and encompasses tools and services that
21 are not relevant to any party’s claims or defenses. Meta further objects to this definition to the
22 extent that it purports to require Meta to produce documents or information concerning large
23 language models that were not publicly released and/or were not trained on corpuses of text that
24 allegedly include any of Plaintiff’s allegedly copyrighted works. Meta further objects that the
25 phrase “process, predict, generate, summarize, classify, or otherwise manipulate natural language
26 or content” renders the definition vague, ambiguous, confusing, and overly broad. Meta will
27 construe “LARGE LANGUAGE MODEL(S)” and “LLM MODEL(S)” to mean the publicly
28 released Llama 1, Llama 2, Llama 3, Llama 4, and Muse Spark models.

1 **8.** Meta objects to the definition of “META,” “YOU,” and “YOUR” as overbroad,
2 unduly burdensome, and disproportionate to the needs of the case. Meta will construe “META,”
3 “YOU,” or “YOUR” to mean Meta Platforms, Inc.

4 **9.** Meta objects to the definition of “PIRATED,” “PIRACY,” and “PIRATED
5 MATERIAL” to the extent it calls for legal conclusions and assumes disputed issues. Meta further
6 objects that the definition is phrased in a way that is argumentative, prejudicial, and lacks
7 foundation. Meta is not in a position to know whether any particular work is subject to copyright
8 protection or whether third-party repositories lacked authorization to host or possess such works.

9 **10.** Meta objects to the definition of “SHADOW LIBRARY” and “SHADOW
10 LIBRARIES” to the extent it calls for legal conclusions and assumes disputed issues. Meta further
11 objects that the definition is phrased in a way that is argumentative, prejudicial, and lacks
12 foundation. Meta further objects to this definition as overbroad, unduly burdensome, and
13 disproportionate to the needs of the case, including (but not limited to) the extent it refers to non-
14 textual materials that are not relevant to any party’s claims or defenses and to the extent it refers
15 vaguely to “media and similar materials.” Meta is not in a position to know whether any particular
16 work is subject to copyright protection or whether third-party repositories lacked authorization to
17 host or possess such works. Meta further objects to this definition to the extent that it incorporates
18 the objectionable portions of the term “DATASETS.” Without any admission with respect to
19 content or nature, Meta will construe the term “SHADOW LIBRARY” to refer to certain of the
20 third-party datasets identified by Plaintiff, namely Anna’s Archive, Bibliotik, Books3, Library
21 Genesis (or LibGen), Sci-Hub, The Pile, and Z-Library.

22 **11.** Meta objects to the definition of “WEB SCRAPING” as vague, ambiguous,
23 overbroad, unduly burdensome, and disproportionate to the needs of the case. In particular, the
24 extensive list of undefined terms, including “extraction, collection, copying, downloading,
25 crawling, indexing, harvesting, or retrieval” and “aggregation, storage, processing” are susceptible
26 to multiple definitions, rendering this definition unintelligibly ambiguous. Meta further objects to
27 the phrase “images, metadata, or other content” as overbroad and disproportionate to the needs of
28 the case, to the extent it refers to training data other than text data that is not relevant to any party’s

1 claims or defenses. Meta further objects to this definition to the extent that it incorporates the
2 objectionable portions of the terms “SHADOW LIBRARIES” and “LARGE LANGUAGE
3 MODELS.” Meta will construe “WEB SCRAPING” to mean the automated collection of text data
4 from the public internet for use in training the Meta Language Models (as construed above).

5 **III. General Objections**

6 Meta hereby incorporates the following general objections into each of its specific
7 responses and objections to the individual Interrogatories as if fully set forth therein:

8 **1.** Meta incorporates by reference each and every General Objection set forth below
9 into each and every specific Response. From time to time, a specific Response may repeat a
10 General Objection for emphasis or some other reason. The failure to repeat any General Objection
11 in any specific Response shall not be interpreted as a waiver of any General Objection to that
12 Response.

13 **2.** Any information provided in response to any individual Interrogatory is made
14 without waiver of, and subject to, the General Objections and any additional objections asserted in
15 response to that specific Interrogatory.

16 **3.** Meta objects to the Interrogatories to the extent that they impose any differing or
17 additional obligations from those required by the Federal Rules, the Local Rules, or any other
18 applicable rule, law, or doctrine, the ESI Order and the Protective Order entered in this action (the
19 “Applicable Laws, Rules, and Agreements”), on the ground that such a requirement exceeds the
20 scope of permissible discovery and attempts to impose an unreasonable burden and expense on
21 Meta. To the extent any Interrogatory conflicts with, or purports to impose additional or different
22 duties and obligations, Meta will respond in accordance with the Applicable Laws, Rules, and
23 Agreements.

24 **4.** Meta objects to the Interrogatories to the extent that they are overly broad, vague,
25 ambiguous, and unduly burdensome; or seek documents and information that are neither relevant
26 nor proportional to the needs of this action.

27 **5.** Meta objects to the Interrogatories to the extent they seek information or documents
28 or seek to impose a search for information or documents: (a) not within Meta’s possession, custody,

1 or control; (b) not maintained by Meta in the normal course of business; (c) already in the
2 possession, custody, or control of Plaintiff; (d) readily available through public sources or from
3 sources that are more convenient, less burdensome, or less expensive; or (e) from sources that are
4 more readily available to Plaintiff than to Meta. Meta further objects to the extent the burden or
5 expense of the proposed discovery outweighs its likely benefit.

6 **6.** Meta objects to the Interrogatories to the extent that they seek documents and
7 information protected from disclosure by privileges and other protections from production,
8 including, without limitation: (a) the attorney-client privilege; (b) the work-product doctrine; or (c)
9 any other constitutional, statutory, common law or regulatory protection, immunity, or proscription
10 from disclosure. In responding to the Interrogatories, Meta will not produce information or
11 documents protected by such privileges. Where Meta withholds information under claim of
12 privilege, Meta will adhere to the requirements under the ESI Order. Meta does not intend the
13 inadvertent production of any privileged or protected information to constitute a waiver of Meta's
14 rights to assert any applicable privilege or protection, and Meta further reserves the right to demand
15 that Plaintiff return, destroy, or sequester any privileged or protected documents inadvertently
16 produced and all copies thereof consistent with Fed. R. Civ. P. 26(b)(5)(B) and the terms and
17 conditions of the ESI Order and the Protective Order entered in this action.

18 **7.** Meta objects to the Interrogatories to the extent that they purport to require Meta to
19 provide information or documents in violation of a legal, contractual, or other obligation of
20 nondisclosure owed to a third party. Meta will not provide information or documents subject to
21 any such obligation unless and until it receives the necessary consent from any such third party or,
22 in the absence of such consent, unless and until it is ordered to do so by a court of competent
23 jurisdiction.

24 **8.** Meta objects to the Interrogatories to the extent they call for, or can be interpreted
25 as calling for, Meta to make legal conclusions, or to the extent they assume disputed issues or are
26 phrased in such a way as to be lacking foundation, argumentative, prejudicial, or otherwise
27 improper.
28

1 **9.** Meta objects to the Interrogatories to the extent that they seek confidential, trade
 2 secret, competitively sensitive, business, financial, or other proprietary information. To the extent
 3 documents or information containing any such information are nonprivileged, responsive to the
 4 Interrogatories, and are not otherwise objectionable, Meta will only produce such documents and
 5 information subject to the ESI Order and the Protective Order entered in this action. Meta further
 6 reserves the right to redact or otherwise withhold any confidential, proprietary, trade secret, or
 7 competitively sensitive information that is neither relevant nor proportional to the needs of this
 8 action.

9 **10.** Meta objects to the Interrogatories to the extent they seek documents that are not
 10 proportionate to the needs of the case and/or not relevant to any party's claims or defenses.

11 **11.** Meta objects to the Interrogatories that seek to impose an obligation to investigate
 12 or discover information or materials from third parties or sources that are equally accessible to
 13 Plaintiff as duplicative, overbroad, unduly burdensome, and oppressive.

14 **12.** Meta reserves the right to assert additional objections, or to supplement its
 15 Responses as appropriate, particularly if any additional information regarding the Interrogatories
 16 or the underlying claims at issue is provided.

17 **IV. OBJECTIONS AND RESPONSES TO INDIVIDUAL INTERROGATORIES**

18 **INTERROGATORY NO. 1:**

19 Which PLAINTIFFS' WORKS did YOU obtain from a SHADOW LIBRARY?

20 **RESPONSE TO INTERROGATORY NO. 1:**

21 Meta incorporates by reference its objections and definitions above, including to the terms
 22 PLAINTIFFS' WORKS, YOU, and SHADOW LIBRARY.

23 Subject to and without waiving the foregoing objections, Meta is undertaking a reasonable
 24 investigation to identify which PLAINTIFFS' WORKS, if any, Meta downloaded as part of Anna's
 25 Archive, Books3 and LibGen for use in training Meta's LARGE LANGUAGE MODELS. As
 26 previously communicated to counsel for PLAINTIFFS, Meta may need additional information,
 27 including the text of the PLAINTIFFS' WORKS, in order to provide a more complete response to
 28

1 this Interrogatory. Meta will supplement this response on a rolling basis as it identifies responsive
2 information.

3 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 1 (AUGUST 24, 2026):**

4 Subject to and without waiving the foregoing objections, and following a reasonable
5 inquiry, Meta further responds as follows:

6 Attached as **Exhibit A** hereto is a table identifying, based on Meta's current knowledge,
7 information, and belief based on a reasonable investigation, the authors and titles of PLAINTIFFS'
8 WORKS that Meta believes were downloaded from Libgen, Zlib, and Books3.

9 Based on Meta's current knowledge, information and belief based on a reasonable
10 investigation, to date Meta has not identified PLAINTIFFS' WORKS that were downloaded from
11 the Internet Archive Lending Library.

12 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
13 responsive information becomes available.

14 **INTERROGATORY NO. 2:**

15 Identify each PLAINTIFFS' WORK that YOU obtained through WEB SCRAPING.

16 **RESPONSE TO INTERROGATORY NO. 2:**

17 Meta incorporates by reference its objections and definitions above, including to the terms
18 PLAINTIFFS' WORK, YOU, and WEB SCRAPING.

19 Subject to and without waiving the foregoing objections, Meta is undertaking a reasonable
20 investigation to identify which PLAINTIFFS' WORKS, if any, Meta downloaded as part of any
21 web archiving activity conducted by Meta's SpiderMate web archiver for use in training Meta's
22 LARGE LANGUAGE MODELS. As previously communicated to counsel for PLAINTIFFS,
23 Meta may need additional information, including the text of the PLAINTIFFS' WORKS, in order
24 to provide a more complete response to this Interrogatory. Meta will supplement this response on
25 a rolling basis as it identifies responsive information.

26 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 2 (AUGUST 24, 2026):**

27 Subject to and without waiving the foregoing objections, and following a reasonable
28 inquiry, Meta further responds as follows:

1 Based on Meta's current knowledge, information, and belief based on reasonable
2 investigation, Meta has not identified any PLAINTIFFS' WORKS that Meta believes were
3 obtained from Common Crawl, an open repository of web crawl data made available by the
4 Common Crawl Foundation.

5 Meta also operates a web crawler "SpiderMate" that obtains publicly available web data
6 from the Internet. The SpiderMate data is organized and searchable based on web domains (e.g.
7 domain portion of Uniform Resource Locators (URLs)). Meta is willing to search SpiderMate data
8 for a reasonable number of web domains identified by PLAINTIFFS upon request.

9 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
10 responsive information becomes available.

11 **SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 2 (SEPTEMBER 4, 2026):**

12 Subject to and without waiving the foregoing objections, and following a reasonable
13 inquiry, Meta further responds as follows:

14 Meta states that it cannot feasibly search SpiderMate based on the text of the PLAINTIFFS'
15 WORKS because that platform is indexed and searchable by URL. Any search through the text of
16 SpiderMate for PLAINTIFFS' WORKS would be extraordinarily burdensome, not technically
17 feasible or proportional to the needs of the case. SpiderMate holds petabytes of data indexed by
18 URL, and any attempt to search for the text of PLAINTIFFS' WORKS would require significant
19 engineering work to develop and deploy a new search tool at significant expense and considerable
20 effort, disproportionate to the needs of the case. Moreover, PLAINTIFFS have not provided any
21 information indicating that PLAINTIFFS' WORKS are present in any material quantity on publicly
22 available internet web pages. The benefits of any such search are substantially outweighed by the
23 burden. Additionally, Plaintiffs' response to Meta Interrogatory No. 15 (served August 31, 2026)
24 states that Plaintiffs seek statutory damages under 17 U.S.C. § 504 and "is not seeking actual
25 damages, lost profits, lost sales, lost licensing revenue, or any other form of compensatory monetary
26 relief based upon Plaintiff's economic losses." (Plaintiffs' Responses and Objections to Meta's
27 Second Set of Interrogatories, August 31, 2026, at 2.) Under 17 U.S.C. § 504, a plaintiff can only
28 recover a single award of statutory damages "for all infringements involved in the action, with

respect to any one work.” 17 U.S.C. § 504(c)(1); *see also* H.R. Rep. No. 1476, 94th Cong., 2d Sess. at 162 (1976) (“A single infringer of a single work is liable for a single amount [of statutory damages], no matter how many acts of infringement are involved in the action and regardless of whether the acts were separate, isolated, or occurred in a related series.”); *Desire, LLC v. Manna Textiles, Inc.*, 986 F.3d 1253, 1271 (9th Cir. 2021) (noting that statutory damages awards “should not provide copyright owners a windfall. Congress specifically limited the amount of statutory damages available per award and limited a plaintiff to a single award no matter how many infringements an individual defendant... committed.”). Accordingly, Plaintiffs can only recover a single award of statutory damages for each PLAINTIFFS’ WORK, regardless of the number of times that work was downloaded by Meta or how many Meta datasets may have incorporated such work. Accordingly, requiring Meta to attempt to develop a new search tool for purposes of searching SpiderMate would be disproportionate to the needs of the case under Rule 26(b)(1).

INTERROGATORY NO. 3:

For each PLAINTIFFS’ WORK that YOU obtained from a SHADOW LIBRARY or through WEB SCRAPING, when and how did you obtain it? Identify and describe each instance in detail, including the source or sources, date or dates of acquisition, and means or method of acquisition.

RESPONSE TO INTERROGATORY NO. 3:

Meta incorporates by reference its objections and definitions above, including to the terms PLAINTIFFS’ WORK, YOU, SHADOW LIBRARY, and WEB SCRAPING.

Meta objects to this request to the extent it is duplicative of information sought by Interrogatory Nos. 1 and 2. Meta further objects to this interrogatory as vague, ambiguous, overbroad, unduly burdensome, and disproportionate to the needs of the case to the extent the phrase “[i]dentify and describe each instance in detail, including” seeks unspecified information beyond the recited request for “the source or sources, date or dates of acquisition, and means or method of acquisition” of data obtained from a SHADOW LIBRARY or WEB SCRAPING.

Subject to and without waiving the foregoing objections, as stated for Interrogatory Nos. 1 and 2, Meta is undertaking a reasonable investigation to identify which PLAINTIFFS’ WORKS, if any (1) Meta downloaded as part of Anna’s Archive, Books3, and/or LibGen for use in training

Meta's LARGE LANGUAGE MODELS and (2) Meta downloaded as part of any web archiving activity conducted by Meta's SpiderMate web archiver for use in training Meta's LARGE LANGUAGE MODELS. Meta will provide reasonably available information regarding the source, timing, and method of download. As previously communicated to counsel for PLAINTIFFS, Meta may need additional information, including the text of the PLAINTIFFS' WORKS, in order to provide a more complete response to this Interrogatory. Meta will supplement this response on a rolling basis as it identifies responsive information.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 3 (AUGUST 24, 2026):

Subject to and without waiving the foregoing objections, and following a reasonable inquiry, Meta further responds as follows:

Attached as **Exhibit A** hereto is a table identifying, based on Meta's knowledge, information, and belief based on reasonable investigation, the authors and titles of PLAINTIFFS' WORKS that Meta believes were downloaded from the Z-Lib, Libgen Fiction, Libgen Scitech and Books3 collections. Based on Meta's current knowledge, information, and belief based on reasonable investigation, Meta has not identified any PLAINTIFFS' WORKS that Meta believes were obtained from Common Crawl, an open repository of web crawl data made available by the Common Crawl Foundation. With respect to each of these data sources, Meta identifies, to the best of its knowledge, information and belief after reasonable investigation, how each of these data sources was obtained:

Anna's Archive (which includes Z-Lib): Meta explained, in its response to Interrogatory No. 9 (served August 8, 2026), the way in which Meta obtained data from the Z-Lib collection of Anna's Archive via torrenting, and that response is incorporated herein by reference.

Libgen Fiction: Meta explained, in its response to Interrogatory No. 9 (served August 8, 2026), the way in which data from Libgen Fiction was obtained via torrenting, and that response is incorporated herein by reference. Additionally, beginning in or about Spring 2023, data from the Fiction branch of Libgen was downloaded, and this download was conducted primarily by Nikolay Bashlykov. The download was conducted using a direct download technique that did not involve use of torrents. Meta has produced a software script (download_libgen_direct.py, META-

1 KADREY-SC-000197), a version of which Meta believes was used in the direct download process
2 for Libgen Fiction.

3 **Libgen Scitech:** Meta explained, in its response to Interrogatory No. 9 (served August 8,
4 2026), the way in which data from Libgen Scitech was obtained via torrenting, and that response is
5 incorporated herein by reference. Additionally, beginning in or about Spring 2023, data from the
6 Scitech branch of Libgen was downloaded, and the download was conducted primarily by Nikolay
7 Bashlykov. The download was conducted using a direct download technique that did not involve
8 use of torrents. Meta has produced a software script (download_libgen_direct.py, META-
9 KADREY-SC-000197), a version of which Meta believes was used in the direct download process
10 for Libgen Scitech.

11 **Common Crawl:** Meta believes that Common Crawl data was first obtained by Meta prior
12 to 2023. Common Crawl was available for download over the Internet from a number of different
13 websites and online databases. Based on Meta's knowledge and information after a reasonable
14 investigation, Meta has not been able to definitively determine where it originally obtained
15 Common Crawl data, but believes it may have been obtained from
16 <<https://www.tensorflow.org/datasets/catalog/c4>>, <<https://github.com/speedingghzl/CCNet/>>, or
17 <[https://github.com/tensorflow/models/tree/archive/research/lm_commonsense#1-download-](https://github.com/tensorflow/models/tree/archive/research/lm_commonsense#1-download-datafiles)
18 [datafiles](https://github.com/tensorflow/models/tree/archive/research/lm_commonsense#1-download-datafiles)>.

19 **Books3:** Meta believes that Books3 data was first obtained prior to 2023. At the time it
20 was obtained by Meta, Books3 was part of a larger dataset known as "The Pile" published by
21 EleutherAI, which was available for download over the Internet from a number of different
22 websites and online databases. Meta has not been able to definitively determine where it originally
23 obtained Books3 data, but believes Books3 may have been obtained from a website known as The
24 Eye (<<https://theeye.eu>>), which at the time hosted a number of publicly-available datasets
25 including Books3 and The Pile. Meta has produced a spreadsheet (Meta_Kadrey_00000250)
26 listing names of files in a Books3 dataset obtained by Meta.

27 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
28 responsive information becomes available.

1 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 3 (SEPTEMBER 4, 2026):**

2 Subject to and without waiving the foregoing objections, and following a reasonable
3 inquiry, Meta further responds as follows:

4 Meta incorporates by reference its previous responses to this Interrogatory. **Exhibit A**
5 provided with the August 24, 2026 supplemental response contains, for each PLAINTIFFS' WORK
6 that was downloaded by Meta, a "Downloaded Dataset(s)" field that identifies, for each work, the
7 online repository or repositories that were downloaded by Meta that contained such work. For
8 example, if a work lists "Z-Lib," in this column, that means it was downloaded as part of the torrent
9 download from Anna's Archive described in Meta's response to Interrogatory No. 9 (served August
10 8, 2026). If a work lists "Books3" in this column, it was downloaded as part of the pre-2023
11 download of Books3 data as described in Meta's response to Interrogatory No. 9 (served August 8,
12 2026).

13 If a work lists "Libgen Scitech," that dataset was downloaded during the multiple
14 downloads of this dataset, specifically (a) the torrent downloads of Libgen Scitech described in
15 Meta's response to Interrogatory No. 9 (served August 8, 2026), and (b) the direct download by of
16 Libgen Scitech Mr. Bashlykov described in Meta's August 24, 2026 supplemental response to
17 Interrogatory No. 3. To the best of Meta's knowledge, information and belief after reasonable
18 investigation, the downloads of Libgen Scitech by Bashlykov and Wang are believed to have been
19 substantially complete as of the date of their respective downloads, whereas the downloads by
20 former employees Lample and Roller were partial. Pursuant to Rule 33(d), Meta has produced
21 listings of downloaded files at META_KADREY_00237775 and META_KADREY_00243793
22 corresponding to the partial downloads of Libgen Scitech by Lample and Roller, respectively,
23 which provide a more precise identification of the works believed to have been downloaded by
24 these individuals. To the best of Meta's knowledge, information, and belief, the portions of Scitech
25 believed to have been downloaded by Mr. Lample did not contain Plaintiff's works. With respect
26 to Mr. Roller, attached as **Exhibit E** is a chart showing, to the best of Meta's knowledge,
27 information, and belief, the particular works of Plaintiff that were included in the portions of
28 Scitech believed to have been downloaded by him.

1 If a work lists “Libgen Fiction,” it was downloaded during the multiple downloads of this
 2 dataset, specifically (a) the torrent download of Libgen Fiction by Mr. Lample described in Meta’s
 3 response to Interrogatory No. 9 (served August 8, 2026), and (b) the direct download by Mr.
 4 Bashlykov described in Meta’s August 24, 2026 supplemental response to Interrogatory No. 3. To
 5 the best of Meta’s knowledge, information and belief after reasonable investigation, the downloads
 6 of Libgen Fiction by Mr. Lample and Mr. Bashlykov are believed to have been substantially
 7 complete as of the date of their respective downloads.

8 If multiple datasets are listed in this column for a particular work, that means the work was
 9 separately downloaded during each of these dataset downloads. Details about each of these
 10 instances of dataset downloading, including the date or dates of acquisition, and the method of
 11 acquisition, are set forth (a) for torrent downloads, in Meta’s response to Interrogatory No. 9 (served
 12 August 8, 2026), and (b) for Mr. Bashlykov’s direct downloads of Libgen Fiction and Libgen
 13 Scitech, and for Books3, in Meta’s August 24, 2026 supplemental response to Interrogatory No. 3.
 14 With respect to **Exhibit B** provided with the August 24, 2026 supplemental response, Meta has not
 15 been able to reliably determine precise dates when downloads of Common Crawl occurred, but
 16 states that Common Crawl data was likely downloaded multiple times as new crawls were made
 17 available by the Common Crawl Foundation.

18 **INTERROGATORY NO. 4:**

19 Which PLAINTIFFS’ WORKS did YOU use, in whole or in part, to train any of YOUR LLM
 20 MODELS?

21 **RESPONSE TO INTERROGATORY NO. 4:**

22 Meta incorporates by reference its objections and definitions above, including to the terms
 23 PLAINTIFFS’ WORKS, YOU, and YOUR LLM MODELS.

24 Meta further objects to this interrogatory to the term “train” as vague and ambiguous to the
 25 extent it is undefined. Meta will interpret the term “train” as the act of pre-training and/or fine-
 26 tuning a large language model.

27 Subject to and without waiving the foregoing objections, Meta is undertaking a reasonable
 28 investigation to identify which PLAINTIFFS’ WORKS, if any, were used to pre-train Meta’s LLM

1 MODELS. As previously communicated to counsel for PLAINTIFFS, Meta may need additional
 2 information, including the text of the PLAINTIFFS' WORKS, in order to provide a more complete
 3 response to this Interrogatory. Meta will supplement this response on a rolling basis as it identifies
 4 responsive information.

5 **SUPPLEMENTAL RESPONSE TO INTERROGATORY No. 4 (AUGUST 24, 2026):**

6 Subject to and without waiving the foregoing objections, and following a reasonable
 7 inquiry, Meta further responds as follows:

8 Attached as **Exhibit C** is a list of PLAINTIFFS' WORKS that based on Meta's knowledge,
 9 information, and belief based on reasonable investigation were used in training one or more of the
 10 Llama models. **Exhibit C** is broken into two parts. **Part 1** provides a listing that identifies the
 11 work files as produced by PLAINTIFFS (with corresponding title), in which Meta believes the
 12 work (or a substantial portion thereof) was used to train at least one of its Llama models. **Part 2**
 13 provides a listing that shows, for each work file as produced by PLAINTIFFS, an identification of
 14 the training dataset and specific training dataset file, that Meta believes includes the work (or a
 15 substantial portion thereof). To the best of Meta's knowledge, information and belief, these
 16 training files were used in the training of Meta models and thus provide an identification of specific
 17 instances in which the identified PLAINTIFFS' WORKS were used in training.

18 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
 19 responsive information becomes available.

20 **INTERROGATORY No. 5:**

21 For each PLAINTIFFS' WORK that YOU used, in whole or in part, to train any of YOUR
 22 LLM MODELS, when and how did you do so? IDENTIFY and describe each instance in detail.

23 **RESPONSE TO INTERROGATORY No. 5:**

24 Meta incorporates by reference its objections and definitions above, including to the terms
 25 PLAINTIFFS' WORKS, YOU, and YOUR LLM MODELS.

26 Meta further objects to this interrogatory to the term "train" as vague and ambiguous to the
 27 extent it is undefined. Meta will interpret the term "train" as the act of pre-training and/or fine-
 28 tuning a large language model. Meta further objects to the phrase "how did you do so?" as it does

1 not specify which actions the interrogatory seeks information about. Meta will interpret the phrase
2 “how did you do so?” as seeking a high-level general description of the way a work like a
3 PLAINTIFFS’ WORK would be used in the pre-training process of Meta’s LLM MODELS. Meta
4 further objects to this interrogatory as vague, ambiguous, overbroad, unduly burdensome, and
5 disproportionate to the needs of the case to the extent the phrase “[i]dentify and describe each
6 instance in detail” seeks unspecified information about the use of PLAINTIFFS’ WORKS in pre-
7 training. Meta will interpret this phrase to seek a high-level general description of the way a work
8 like a PLAINTIFFS’ WORK would be used in the pre-training process of Meta’s LLM MODELS.

9 Subject to and without waiving the foregoing objections, as stated for Interrogatory No. 4,
10 Meta is undertaking a reasonable investigation to identify which PLAINTIFFS’ WORKS, if any,
11 that were used to pre-train Meta’s LLM MODELS and Meta will provide a high-level general
12 description of the way a work like a PLAINTIFFS’ WORK would be used in the pre-training
13 process of Meta’s LLM MODELS. As previously communicated to counsel for PLAINTIFFS,
14 Meta may need additional information, including the text of the PLAINTIFFS’ WORKS, in order
15 to provide a more complete response to this Interrogatory. Meta will supplement this response on
16 a rolling basis as it identifies responsive information.

17 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 5 (AUGUST 24, 2026):**

18 Subject to and without waiving the foregoing objections, and following a reasonable
19 inquiry, Meta further responds as follows:

20 Meta has provided, in response to Interrogatory 7, *infra*, a description of how textual
21 training data (which would include the PLAINTIFFS’ WORKS identified in response to
22 Interrogatory No. 4) was processed and used in the training process for Meta’s models. As
23 described in response to Interrogatory No. 4, *supra*, **Exhibit C, Part 2**, provides a listing that
24 shows, for each work file as produced by PLAINTIFFS, an identification of the training dataset
25 and training dataset file, where Meta believes the work (or a substantial portion thereof), is
26 contained. To the best of Meta’s knowledge, information and belief, these training files were used
27 in the training of Meta models and thus provides an identification of specific instances in which the
28 identified PLAINTIFFS’ WORKS were used in training.

Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional responsive information becomes available.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 5 (SEPTEMBER 4, 2026):

Subject to and without waiving the foregoing objections, and following a reasonable inquiry, Meta further responds as follows:

Meta incorporates by reference its previous responses to this Interrogatory. **Exhibit C** provided with the August 24, 2026 supplemental response identifies, for each work, codenames of the dataset(s) in which each identified work was located, and each codename corresponds to b3g, Libgen Fiction, Libgen Scitech, Z-Lib, or Common Crawl (cc). Meta states that the b3g dataset was used to train Llama 1 (released February 2023), Llama 2 (released July 2023), Llama 3 (released April 2024), and Llama 4 (released April 2025). As such, any work in **Exhibit C** that lists b3g was used in training each of these models. The Libgen Scitech dataset was used to train Llama 3 and 4, and as such, any work in **Exhibit C** that lists a Libgen Scitech dataset (*i.e.*, scitech_20231120, scitech_20231126, scitech_en, scitech_non_en) was used in training each of those models. The Libgen Fiction dataset was used to train Llama 3 and 4, and as such, any work in **Exhibit C** that lists a Libgen Fiction dataset (*e.g.*, libgen_fiction_20231120, libgen_fiction_rest_safe, llama_fiction_split) was used in training each of those models. The Z-Lib dataset was part of Anna's Archive and first used to train Llama 4, and as such, any work in **Exhibit C** that lists a Z-Lib dataset (*i.e.*, zlib, llama_bespoke_zlib) was used in training Llama 4. Pursuant to Rule 33(d), Meta refers Plaintiff to the "b3g" dataset produced at META_KADREY_DATA_001, the "Libgen" and "Anna's Archive" datasets produced as part of META_ENTREPRENEUR_DATA_001. Common Crawl datasets were used to train Llama 1 through Llama 4. Pursuant to Rule 33(b), Meta identifies the Common Crawl datasets produced by Meta identified by codenames in **Exhibit C**, which were produced at META_KADREY_DATA_002 ("c4" and "edouard_cc"), META_KADREY_DATA_002A ("ccnet"), and META_KADREY_DATA_003 ("ccnew_new"). Additionally, **Exhibit D** identifies the works that were used to train the Muse Spark family of models, the first of which was released in April 2026.

1 Additionally, the training run for each of these models starts from scratch with the model
 2 parameters randomly initialized; as a result, to the extent a PLAINTIFF'S WORK was used in
 3 training a Meta model, that training run represents a separate instance in which text from the work
 4 was used. Additionally, prior to the public release of any Meta Llama or Muse model and consistent
 5 with standard LLM development practices, a number of internal models are trained for
 6 experimental and evaluation purposes including ablation studies, performance benchmarking, and
 7 other research to refine the training datamix and hyperparameters for the training run resulting in
 8 the model intended for public release. Meta is unable to reliably and reasonably ascertain the
 9 number of research and internal models that were created for these purposes. Additionally, Meta
 10 has also developed and utilized a number of purpose-built variants of its LLMs that serve discrete
 11 internal purposes such as data processing and classification, evaluation of LLM prompts and
 12 responses for safety purposes, and content analysis and filtering.

13 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
 14 responsive information becomes available.

15 **INTERROGATORY NO. 6:**

16 IDENTIFY each PLAINTIFFS' WORK from which YOU removed or altered any COPYRIGHT
 17 MANAGEMENT INFORMATION.

18 **RESPONSE TO INTERROGATORY NO. 6:**

19 Meta incorporates by reference its objections and definitions above, including to the terms
 20 PLAINTIFFS' WORKS, YOU, and COPYRIGHT MANAGEMENT INFORMATION.

21 Meta further objects to this interrogatory as overbroad, unduly burdensome, and
 22 disproportionate to the needs of the case to the extent it seeks information about PLAINTIFFS'
 23 WORKS that were not used as part of Meta's pre-training of large language models.

24 Subject to and without waiving the foregoing objections, Meta is undertaking a reasonable
 25 investigation to identify for which PLAINTIFFS' WORKS identified in its responses to
 26 Interrogatory Nos. 1 and 2 Meta is able to reasonably determine that COPYRIGHT
 27 MANAGEMENT INFORMATION was removed or altered by Meta. As previously
 28 communicated to counsel for PLAINTIFFS, Meta may need additional information, including the

1 text of the PLAINTIFFS' WORKS, in order to provide a more complete response to this
 2 Interrogatory. Meta will supplement this response on a rolling basis as it identifies responsive
 3 information.

4 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 6 (AUGUST 24, 2026):**

5 Subject to and without waiving the foregoing objections, and following a reasonable
 6 inquiry, Meta further responds as follows:

7 As explained in response to Interrogatory No. 7, *infra*, Meta did employ processes in which
 8 repetitive text from books was not included in the training data used for training certain Meta Llama
 9 models, including data that commonly appear in works such as "ISBN," "Copyright," the copyright
 10 symbol ©, "All Rights Reserved," and "DOI." Based on Meta's current understanding and
 11 investigation, Meta believes this processing step may have been performed in the processing of at
 12 least some PLAINTIFFS' WORKS identified in response to Interrogatory No. 4, *supra*. Meta is
 13 undertaking a reasonable search of the produced training data to determine if it can more precisely
 14 identify particular PLAINTIFFS' WORKS to which this processing step was applied. As further
 15 explained in response to Interrogatory No. 7, *infra*, Meta did not employ this processing step for
 16 scanned and digitized books it used as training data for its Muse models.

17 Meta's investigation is ongoing and Meta will supplement this Interrogatory as additional
 18 responsive information becomes available.

19 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 6 (SEPTEMBER 4, 2026):**

20 Subject to and without waiving the foregoing objections, and following a reasonable
 21 inquiry, Meta further responds as follows:

22 Meta believes that for a substantial majority of the CHICKEN SOUP FOR THE SOUL
 23 WORKS set forth in Exhibit C, repetitive text was not included in the training data used for training
 24 certain Meta Llama models, including data that commonly appear in works such as "ISBN,"
 25 "Copyright," the copyright symbol ©, "All Rights Reserved," and "DOI." Meta's investigation is
 26 ongoing and Meta will supplement if appropriate.

27 **INTERROGATORY NO. 7:**

28 Explain what you did with each PLAINTIFFS' WORK that YOU obtained from a SHADOW

1 LIBRARY or through WEB SCRAPING or that YOU used, in whole or in part, to train any of
 2 YOUR LLM MODELS. For each PLAINTIFFS' WORK, your explanation should state whether,
 3 when, and how it was:

4 (a) retained, stored, or preserved;

5 (b) ingested;

6 (c) processed;

7 (d) tokenized;

8 (e) duplicated or reproduced;

9 (f) de-duplicated;

10 (g) stripped of COPYRIGHT MANAGEMENT INFORMATION;

11 (h) modified or altered in any other way;

12 (i) distributed;

13 (j) torrented;

14 (k) incorporated, in whole or in part, into any of YOUR LLM MODELS or model

15 parameters; or

16 (l) deleted.

17 **RESPONSE TO INTERROGATORY NO. 7:**

18 Meta incorporates by reference its objections and definitions above, including to the terms
 19 PLAINTIFFS' WORKS, SHADOW LIBRARY, WEB SCRAPING, COPYRIGHT
 20 MANAGEMENT INFORMATION, YOU, and YOUR LLM MODELS.

21 Meta further objects to this interrogatory as compound because it comprises at least twelve
 22 subparts and would potentially exceed the allowable number of interrogatories under the Federal
 23 Rules of Civil Procedure. Meta further objects to the terms "[e]xplain what you did", "processed",
 24 "tokenized", "stripped", "modified or altered in any other way", "incorporated, in whole or in part,
 25 into any of YOUR LLM MODELS or model parameters" as vague and ambiguous. Meta further
 26 objects to this interrogatory to the extent it seeks information sought by Interrogatory Nos. 1-6.
 27 Meta further objects to this interrogatory to the extent it requires Meta to create or re-create
 28 information or data that is not in Meta's possession, custody, or control.

1 Subject to and without waiving the foregoing objections, as stated for Interrogatory Nos. 1-
2 6, Meta is undertaking a reasonable investigation to identify which PLAINTIFFS' WORKS, if any,
3 were downloaded and used to pre-train Meta's LLM MODELS. Meta will provide a high-level
4 general description of the way a work like a PLAINTIFFS' WORK would be utilized by Meta in
5 conjunction with Meta's LLM MODELS. As previously communicated to counsel for
6 PLAINTIFFS, Meta may need additional information, including the text of the PLAINTIFFS'
7 WORKS, in order to provide a more complete response to this Interrogatory. Meta will supplement
8 this response on a rolling basis as it identifies responsive information.

9 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7 (AUGUST 24, 2026):**

10 Subject to and without waiving the foregoing objections, and following a reasonable
11 inquiry, Meta further responds as follows:

12 **This response is designated as Highly Confidential - Attorneys' Eyes Only under the**
13 **Protective Order.**

14 As explained in response to Interrogatory No. 3, Meta used publicly-available works that
15 were downloaded from online repositories (in some cases through torrent downloads) and web data,
16 in connection with the research, development, and training of certain AI models. An initial step is
17 to download the files, which are (in the case of books) typically in Portable Document Format
18 (PDF) or Electronic Publication (EPUB) format. Regardless of the format, downloaded files must
19 then be processed into a form suitable for model training. In many instances, the downloaded book
20 files contained a series of scanned pages, which were run through optical character recognition
21 (OCR) to extract textual information from the files. Textual information from the downloaded files
22 was further filtered and processed for quality and de-duplication. In some cases, the processing of
23 the downloaded files resulted in certain text from those files not being included in the data
24 ultimately used for model training. For example, Meta created software scripts to filter out text
25 comprising certain personally identifiable information (PII) (such as email addresses) and repetitive
26 text in book files from certain downloaded data. This repetitive text included (as explained at
27 Meta_Kadrey_00065249), lines that contain fewer than 8% unique words, and repetitive words and
28 phrases that commonly appear in works such as "ISBN," "Copyright," the copyright symbol ©,

1 “All Rights Reserved,” and “DOI.” *See also* Meta’s Response to Interrogatory No. 6, for more
2 specificity. Meta notes that, to the extent other data from the original works was not included in
3 the training data, that fact will be reflected and can be ascertained from the training data produced
4 to Plaintiffs.

5 After the downloading, de-duplication, processing and cleaning processes, the resulting data
6 may then be used as input text for training a large language model (LLM). The LLM training
7 process is a complex one involving numerous steps, including steps of tokenization and
8 embedding. Tokenization generally breaks the input text into smaller units (tokens), which are
9 assigned unique numbers. The embedding process then identifies multi-dimensional vectors
10 (embeddings) for each token in the input sequence, which attempt to capture the meaning of the
11 words in the input text. The training process continues by feeding these word embeddings into the
12 transformer layers of the LLM, to obtain next token probabilities from the neural network, which
13 are compared and used to update the model weights (also commonly referred to as
14 parameters). These training steps are computationally intensive and are repeated billions or even
15 trillions of times (depending on the training dataset size) until the LLM has made predictions about
16 all of the input text in the training dataset. Billions of LLM weights are adjusted throughout this
17 process to improve the model’s next-token prediction accuracy. *See* Meta’s Response to
18 Interrogatory No. 4, for an identification of the particular PLAINTIFFS’ WORKS that Meta
19 believes were used (in whole or part) for model training.

20 Once an LLM has been trained and is available to be used, it can be provided to users in a
21 number of ways. For Llama 1 through Llama 4, Meta made these models available for distribution
22 to the open source community for use by researchers, developers, and other third parties, pursuant
23 to applicable open source license agreements. Although the weights in an open source model are
24 distributed and available to the open source community, the underlying text data used to train the
25 LLM was not provided with these models or their model weights. Beginning in or about April
26 2024, Meta also made functionality from Llama broadly available to the general public through
27 Meta AI and incorporated into certain other Meta products.

1 Beginning in or around mid-2025, Meta began acquiring physical copies of books for use
2 in training later LLMs. Many of the processes described above for models trained on downloaded
3 works also apply to these models, with some exceptions. The acquired physical copies were first
4 scanned into digital form and then run through OCR to extract their textual data from the scanned
5 images as discussed above, and then de-duplicated, processed and cleaned, but copyright-related
6 information from these purchased books was not filtered out before being used for model
7 training. In April 2026, Meta introduced the initial version of its Muse Spark model, which is
8 available through certain Meta products and services, and later versions were made available to
9 developers through an application programming interface (API). In August 2026, Meta also made
10 its Muse Glimmer model available for distribution to the open source community for use by
11 researchers, developers, and other third parties, pursuant to an open source license agreement.

12 **SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7 (SEPTEMBER 4, 2026):**

13 Subject to and without waiving the foregoing objections, and following a reasonable
14 inquiry, Meta further responds as follows:

15 Meta incorporates by reference its previous responses to this Interrogatory. Meta further
16 states that to the extent any work of PLAINTIFFS was used in training a Llama or Muse model,
17 each of the processing, filtering and training steps above would apply to each of those works (with
18 the exception of copyright-related information processing with respect to purchased books for the
19 Muse models). A number of transitory and persistent copies of each of these works would have
20 been created in the conversion and transformation of the original text into the processed form
21 suitable for input into Meta's LLM training pipelines. Additional copies are also made when the
22 processed data is used in a training run, including transitory and intermediate copies made when
23 the processed data (which includes text from the original works) is transformed into tokenized data
24 representations used in the embedding and next-token prediction processes described above. To
25 the extent a work of PLAINTIFFS is used in training a Llama or Muse model, Meta has no feasible
26 way to reliably ascertain the number of additional persistent or transitory copies of text from such
27 work that are made in the course of these processing and training processes. Meta further states
28 that, to the extent Plaintiff contends this interrogatory requires a cataloguing of every conceivable

1 copy of its works that may have been made internally at Meta in connection with its LLM
 2 development efforts, this information would be irrelevant and not proportional to the needs of the
 3 case. *See Kadrey v. Meta Platforms, Inc.*, 788 F.Supp. 3d 1026, 1048 (N.D. Cal. 2025) (“In any
 4 event, even if Meta did download some copies that weren't ultimately used for training, fair use
 5 doesn't require that the secondary user make the lowest number of copies possible.”). Plaintiff has
 6 informed Meta that it seeks only statutory damages in this action and no other form of compensatory
 7 monetary relief, and as such, plaintiff can only recover a single award of statutory damages “for all
 8 infringements involved in the action, with respect to any one work.” 17 U.S.C. § 504(c)(1); *see*
 9 *also* H.R. Rep. No. 1476, 94th Cong., 2d Sess. at 162 (1976) (“A single infringer of a single work
 10 is liable for a single amount [of statutory damages], no matter how many acts of infringement are
 11 involved in the action and regardless of whether the acts were separate, isolated, or occurred in a
 12 related series.”); *Desire, LLC v. Manna Textiles, Inc.*, 986 F.3d 1253, 1271 (9th Cir. 2021). Meta
 13 further states that to the extent it used any work of PLAINTIFFS in training its LLMs, Meta is
 14 unaware of any uses of such work that were not in furtherance of the research, development,
 15 evaluation and training of its LLMs. Pursuant to Rule 33(d), Meta further refers to the deposition
 16 testimony of James Beldock and Shaun Lindsay given in this action, which provides additional
 17 information about how datasets are acquired and used in the research, development, evaluation and
 18 training of Meta’s LLMs.

19 Meta made open source distributions of Llama 1-4 and Muse Glimmer (in August 2026),
 20 but these distributions include model weights but not the data (including any work of PLAINTIFFS)
 21 used to train the model. Additionally, Meta does not currently have information with respect to
 22 whether any work of PLAINTIFF (or any portion thereof) was uploaded or distributed to any third
 23 party during downloads via torrent, as the information available to Meta does not identify the
 24 content of data payloads, if any, that may have been uploaded during such downloads. Ascertaining
 25 whether a particular work of PLAINTIFF (or any portion thereof) was distributed during torrenting
 26 is a matter for an analysis as part of expert discovery. As explained in Meta’s response to
 27 Interrogatory No. 9 (served August 8, 2026), Meta has produced extensive file lists, cost and usage
 28 data, log files, and other data for each torrent download, to facilitate such analysis. Pursuant to

1 Rule 33(d), Meta further refers to the Rule 30(b)(6) deposition testimony of James Beldock
2 provided on August 28, 2026 in *Kadrey v. Meta Platforms, Inc.*, No. 3:23-cv-03417-VC (N.D.
3 Cal.), which provides additional information relating to the alleged distribution of copyrighted
4 works during torrent downloads.

5 **INTERROGATORY NO. 8:**

6 For each response to every Request for Admission served on YOU by Plaintiff in this case
7 that is not an unqualified admission, IDENTIFY the following information:

- 8 (a) The number of the Request for Admission;
9 (b) All facts on which YOU base your response;
10 (c) All PERSONS who have knowledge of those facts; and
11 (d) All DOCUMENTS that support YOUR response along with the custodian of those
12 DOCUMENTS.

13 **RESPONSE TO INTERROGATORY NO. 8:**

14 Meta incorporates by reference its objections and definitions above, including to the terms
15 YOU and DOCUMENTS.

16 Meta further objects to this interrogatory as compound because it comprises multiple
17 subparts and would potentially exceed the allowable number of interrogatories under the Federal
18 Rules of Civil Procedure. Meta will respond to this interrogatory to the extent required by Local
19 Rule 33-2 and 36-2.

20 Meta is willing to meet and confer regarding the appropriate scope, if any, of a response to
21 this interrogatory.

1 Dated: September 4, 2026

COOLEY LLP

2
3 By: /s/ Phillip E. Morton

4 Phillip E. Morton

5 Karen L. Dunn (*pro hac vice*)
6 Jessica E. Phillips (*pro hac vice*)
7 Kyle N. Smith (*pro hac vice*)
8 DUNN ISAACSON RHEE LLP
9 401 9th Street NW
10 Washington, DC 20004
11 Email: kdunn@dirllp.com
12 jphillips@dirllp.com
13 ksmith@dirllp.com

14 Mark R. Weinstein
15 Elizabeth Lee Stameshkin
16 COOLEY LLP
17 3175 Hanover Street
18 Palo Alto, CA 94304
19 Email: mweinstein@cooley.com
20 lstameshkin@cooley.com

21 Bobby A. Ghajar
22 COOLEY LLP
23 1333 2nd Street, Suite 400
24 Santa Monica, CA 90401
25 Email: bghajar@cooley.com

26 Phillip E. Morton (*pro hac vice pending*)
27 COOLEY LLP
28 1299 Pennsylvania Avenue NW, Suite 700
Washington, DC 20004
Telephone: (202) 842-7800
Email: pmorton@cooley.com

Angela L. Dunning
CLEARY GOTTLIEB STEEN & HAMILTON
LLP
151 University Avenue
Palo Alto, CA 94301-1632
Telephone: (650) 815-4131
Email: adunning@cgsh.com

Counsel for Defendant Meta Platforms, Inc.

Author	Book	Downloaded Dataset(s)
John Carreyrou	BAD BLOOD: Secrets and Lies in a Silicon Valley Startup	Libgen Scitech, Z-Lib, Books3
Philip Shishkin	Restless Valley: Revolution, Murder, and Intrigue in the Heart of Central Asia	Libgen Scitech
Matthew Sacks	Pro Website Development and Operations	Libgen Scitech
Michael S. Kochin	Five Chapters on Rhetoric: Character Action Things Nothing and Art	Libgen Scitech
Michael S. Kochin, Michael Taylor	An Independent Empire: Diplomacy & War in the Making of the United States	Libgen Scitech
Jane Adams	How to Sell What You Write	Libgen Fiction
Jane Adams	Boundary Issues: Using Boundary Intelligence to get the Intimacy you want and the Independence you need in life, love and work	Libgen Scitech

Produced Book ID	Title
Cambronne_Meta_0000037.txt	Bad Blood searchable (without 2022 afterword)
Cambronne_Meta_0000994.txt	Pro Website Development and Operations
Cambronne_Meta_0001111.txt	Kochin_Revised - review copy - Independent Empire
Cambronne_Meta_0001432.txt	Michael S. Kochin - Five Chapters on Rhetoric: Character, Action, Things, Nothing, and Art-Penn State University Press (2015) (1)

Produced Book ID	Training Dataset	Training Data File	Title
Cambronne_Meta_0000037.txt	llbgen_fiction_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/fiction-20231120/fiction_en_processed_20231120.chunk.31.jsonl	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	scilech_en	META_ENTREPRENEUR_DATA_001/scilech_en_union_train_bucket_91.csv	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	scilech_en	META_ENTREPRENEUR_DATA_001/scilech_en_union_train_bucket_31.csv	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	scilech_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.08.jsonl	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	b3g	META_KADREY_DATA_001tb3g/b3g.chunk.25.jsonl	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	llama3_fiction_split	META_ENTREPRENEUR_DATA_001/llama3_fiction_split.csv	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	llbgen_fiction_rest_safe	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/fiction/20231210/safe/data.chunk.029.jsonl	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000037.txt	scilech_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.30.jsonl	Bad Blood searchable (without: 2022 afterword)
Cambronne_Meta_0000094.txt	scilech_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.02.jsonl	Pro Website Development and Operations
Cambronne_Meta_0000094.txt	scilech_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.15.jsonl	Pro Website Development and Operations
Cambronne_Meta_0000094.txt	scilech_en	META_ENTREPRENEUR_DATA_001/scilech_en_union_train_bucket_61.csv	Pro Website Development and Operations
Cambronne_Meta_0000094.txt	scilech_en	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.12.jsonl	Kochin_ Revised - review copy - Independent Empire
Cambronne_Meta_0001111.txt	scilech_en	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.15.jsonl	Kochin_ Revised - review copy - Independent Empire
Cambronne_Meta_0001432.txt	scilech_20231120	META_KADREY_DATA_005leag-wsffair_jlm_v3/data/minhashdeduped/llb/scilech-20231120/scilech_en_processed_20231120.chunk.15.jsonl	Michael S. Kochin - Five Chapters on Rhetoric_ Character, Action, Things, Nothing, and Art-Penn State University Press (2015). (1)
Cambronne_Meta_0001432.txt	scilech_en	META_ENTREPRENEUR_DATA_001/scilech_en_union_train_bucket_61.csv	Michael S. Kochin - Five Chapters on Rhetoric_ Character, Action, Things, Nothing, and Art-Penn State University Press (2015). (1)
Cambronne_Meta_0001432.txt	scilech_en	META_ENTREPRENEUR_DATA_001/scilech_en_union_train_bucket_11.csv	Michael S. Kochin - Five Chapters on Rhetoric_ Character, Action, Things, Nothing, and Art-Penn State University Press (2015). (1)

ISBN	TITLE	AUTHOR	PUBLISHER	YEAR
9780271034560	Five Chapters on Rhetoric: Character, Action, Things, Nothing, and Art	Michael S. Kochin	Penn State University Press	2012
9780300184365	Restless Valley: Revolution, Murder, and Intrigue in the Heart of Central Asia	Philip Shishkin	Yale University Press	2013
9780300205916	Restless Valley: Revolution, Murder, and Intrigue in the Heart of Central Asia	Philip Shishkin	Yale University Press	2014
9780385306942	I'm Still Your Mother How to Get Along with Your Grown Up Children for the Rest of Your Life	Jane Adams	Delacorte Press	1994
9780399129827	How to sell what you write	Jane Adams	Putnam	1984
9780471660453	Boundary Issues: Using Boundary Intelligence to Get the Intimacy You Want and the Independence You Need in Life, Love, and Work	Jane Adams	Trade Paper Press	2005
9780472054404	An Independent Empire: Diplomacy & War in the Making of the United States	Michael S. Kochin, Michael Taylor	University of Michigan Press	2020
9780525431992	Bad Blood: Secrets and Lies in a Silicon Valley Startup	John Carreyrou	Knopf Doubleday Publishing Group	2020
9780595185641	Seattle Green	Jane Adams	iUniverse	2001
9780688087548	Wake Up, Sleeping Beauty: How to Live Happily Ever After-Starting Right Now	Jane Adams	William Morrow & Co	1990
9780738718507	The Street-Smart Psychic's Guide to Getting a Good Reading	Lisa Barretta	Llewellyn Worldwide	2009
9780877959113	Seattle Green: A Novel	Jane Adams	Arbor House Pub Co	1987
9781430239697	Pro Website Development and Operations: Streamlining DevOps for large-scale websites (Expert's Voice in Web Development)	Matthew Sacks	Apress	2012
9781509868070	Bad Blood: Secrets and Lies in a Silicon Valley Startup [Paperback] John Carreyrou	John Carreyrou	Picador	2018
9781509868087	Billion Dollar Whale By Tom Wright, Bradley Hope & Bad Blood Secrets and Lies in a Silicon Valley Startup By John Carreyrou 2 Books Collection Set	John Carreyrou	Scribe UK/Picador	2019
9781524731656	Bad Blood: Secrets and Lies in a Silicon Valley Startup	John Carreyrou	Knopf	2018
9781524731663	Bad Blood: Secrets and Lies in a Silicon Valley Startup	John Carreyrou	Vintage	2018

CONFIDENTIAL

2026-09-04 Cambronne Ex D (Muse Spark Works)

9781601632173	The Book of Transformation: Open Yourself to Psychic Evolution, the Rebirth of the World, and the Empowering Shift Pioneered by the Indigos	Lisa Barretta	Weiser	2012
9781632651143	Conscious Ink: The Hidden Meaning of Tattoos: Mystical, Magical, and Transformative Art You Dare to Wear	Lisa Barretta	Weiser	2017
97833328105909	Bad Blood: Die wahre Geschichte des größten Betrugs im Silicon Valley - Ein SPIEGEL-Buch		Kiligry	
9783421048233	Bad Blood: Die wahre Geschichte des größten Betrugs im Silicon Valley - Ein SPIEGEL-Buch	John Carreyrou	Deutsche Verlags-Anstalt	2019
9789867778383	Bad Blood (Chinese Edition)	John Carreyrou	Shang Ye Zhou Kan/Tsai Fong Books	2018

CONFIDENTIAL

Author	Title
John Carreyrou	BAD BLOOD: Secrets and Lies in a Silicon Valley Startup
Philip Shishkin	Restless Valley: Revolution, Murder, and Intrigue in the Heart of Central Asia
Matthew Sacks	Pro Website Development and Operations
Michael S. Kochin, Michael Taylor	An Independent Empire: Diplomacy & War in the Making of the United States
Jane Adams	How to Sell What You Write
Jane Adams	Boundary Issues; Using Boundary Intelligence to get the Intimacy you want and the Independence you need in life, love and work

CERTIFICATE OF SERVICE

I am a citizen of the United States and a resident of the California. I am employed in Santa Clara County, State of California. I am over the age of eighteen years, and not a party to the within action. My business address is Cooley LLP, 3175 Hanover Street, Palo Alto, CA 94304. On the date set forth below I served the documents described below in the manner described below:

**DEFENDANT META PLATFORMS, INC.'S SECOND SUPPLEMENTAL RESPONSES AND
OBJECTIONS TO PLAINTIFF'S FIRST SET OF INTERROGATORIES**

☒ (BY ELECTRONIC MAIL) I am personally and readily familiar with the business practice of Dunn Isaacson Rhee LLP for the preparation and processing of documents in portable document format (PDF) for e-mailing, and I caused said documents to be prepared in PDF and then served by electronic mail to the parties listed below.

on the following part(ies) in this action:

FREEDMAN NORMAND FRIEDLAND LLP	STRIS & MAHER LLP
Kyle Roche (<i>pro hac vice</i>)	Elizabeth Brannen (# 226234)
Devin (Velvel) Freedman (<i>pro hac vice</i>)	John Stokes (# 310847)
Alex Potter (<i>pro hac vice</i>)	Lauren Martin (# 294367)
155 E. 44th Street, Suite 915	17785 Center Court Dr N, Ste 600
New York NY 10017	Cerritos, CA 90703
Email: vel@fnf.law	Email: ebrannen@stris.com
Email: kroche@fnf.law	Email: jstokes@stris.com
Email: apotter@fnf.law	Email: Jason.Morris@ndlf.com

Christopher M. Rigali (*pro hac vice*)
 Jacqueline Sahlberg (*pro hac vice*)
 1717 K St NW Suite 900
 Washington, DC 20006
 Email: crigali@stris.com
 Email: jsahlberg@stris.com

Executed on September 4, 2026, at Palo Alto, California.

/s/ Jocelyn McIntosh
 Jocelyn McIntosh