

# EXHIBIT B



Lauren Martin  
COUNSEL

STRIS & MAHER LLP  
17785 Center Court Dr N, Ste 600  
Cerritos, CA 90703  
stris.com

lmartin@stris.com

June 30, 2026

Privileged & Confidential

VIA E-MAIL

Karen L. Dunn  
Jessica E. Phillips  
Dunn Isaacson Rhee LLP  
401 9th Street, NW  
Washington, DC 20004

Re: **Re: Discovery *Cognella, Inc. v. Meta Platforms, Inc.*, No. 3:26-cv-04053-VC, *Chicken Soup for the Soul, LLC v. Meta Platforms, Inc.*, No. 3:26-cv-02333-VC; *Cambronne, et al. v. Meta Platforms, Inc.*, No. 3:26-cv-03725-VC**

Dear Counsel:

We write on behalf of Plaintiffs Cognella, Inc. (“Cognella”), Chicken Soup for the Soul, LLC (“Chicken Soup”), and each Plaintiff in the *Cambronne* matter (collectively, “Plaintiffs”) regarding Meta’s deficient responses to the written discovery in the above captioned matters.

On May 21, 2026, Cognella and Chicken Soup served Meta with a very limited number of initial written discovery requests: eight (8) Interrogatories, three (3) Requests for Admission, and one (1) Request for Production. The *Cambronne* Plaintiffs served Meta with the same discovery on May 27, 2026. Plaintiffs’ requests were designed to elicit key information central to this case, including which of Plaintiffs’ copyrighted Works Meta obtained, the source from which each work was obtained (including any shadow libraries), and the manner in which those Works were acquired and used. Despite representing in open court on May 22<sup>1</sup> that Meta would be forthcoming with this information, Meta’s has not responded in substance to any of Plaintiffs’ requests.

Meta’s responses also repeatedly depart from the plain language of the requests by unilaterally redefining key terms to inappropriately narrow the scope of each inquiry and avoid even partial disclosure of the requested information about Plaintiffs’ works. Under Rule 37(a)(4), an evasive or incomplete disclosure, answer, or response is treated as a failure to disclose, answer, or respond.

Plaintiffs’ counsel would like to confer this Thursday or by Monday, July 6, 2026 regarding Meta’s failure to substantively respond, including the specific deficiencies outlined below. Absent complete and forthcoming supplemental responses from Meta by **July 10, 2026**, Plaintiff will have no choice but to seek appropriate relief from the Court. Please provide your availability to discuss.

---

<sup>1</sup> Case Management Conference Hearing Transcript, May 22, 2026.

Jessica E. Phillips  
June 30, 2026  
Page 2

Privileged & Confidential

### **Burden Shifting**

Despite Meta’s repeated contention that it may “need” additional information, including the text of Plaintiffs’ copyrighted works to conduct a search, the burden of identifying which works obtained from shadow libraries and what it did with those works rests squarely with Meta. Plaintiffs dispute that Meta needs anything from Plaintiffs to disclose what it obtained and used from shadow libraries, but in any event, Meta has had “additional information” from Plaintiffs for some time. To be helpful and try to avoid a dispute, Plaintiffs already provided the requested ISBNs known to them for their works on June 2, 2026. It has also provided text for the Chicken Soup and Cambronne works on June 22, 2026 and June 24, 2026, respectively. Text for the Cognella works is forthcoming. Delay in responding to Plaintiffs’ request based on the need for further information is unwarranted.

### **Definitional Objections**

To start, Plaintiffs challenge Meta’s definitional objections and unilateral interpretations of Plaintiffs’ definitions as described in Set One. The definitions are unambiguous and must stand as defined because they are foundational to the nature of each corresponding discovery request. For example:

#### *COPYRIGHT MANAGEMENT INFORMATION*

Meta’s contention that it “is not in a position to know whether any particular information qualifies as legally protected Copyright Management Information” and that such information is outside of its possession, custody or control. Meta’s objection is inconsistent with its prior discovery responses and litigation positions in *Kadrey*, where it has demonstrated familiarity with the concept of copyright management information. Meta has admitted to its use of “pirate libraries” and understands what CMI is as defined in 17 U.S.C. § 1202(c). Plaintiffs have not asked Meta to reach legal conclusions regarding the statutory status of particular information; Plaintiffs defined the term for purposes of discovery. If Meta believed clarification was necessary, it was obligated to answer the attendant requests using the definition provided or explain any genuine ambiguity—not rewrite the definition to avoid responding.

#### *CORPORA and DATASET(S)*

Plaintiffs defined these terms to encompass the collections of works and training data likely at issue in this action. Rather than responding under Plaintiffs definitions, Meta has unilaterally narrowed those definitions by limiting them to a self-selected list of datasets. Once again, Meta may not rewrite Plaintiff’s definitions to exclude responsive sources or otherwise narrow the scope of discovery.

#### *LARGE LANGUAGE MODELS and LLM MODELS*

Meta may not unilaterally narrow the definition of “LARGE LANGUAGE MODEL(S)” and “LLM MODEL(S)” to mean the publicly released Llama 1, Llama 2, Llama 3, Llama 4, and Muse

Jessica E. Phillips  
June 30, 2026  
Page 3

Privileged & Confidential

Spark models. This is a transparent effort to limit Meta's searches to its commercially released models. But Meta has developed numerous internal and experimental models that have never been released to the public but for which Plaintiffs' works were likely copied, processed and tested. These are all relevant to the claims of infringement and Meta's searches and responses must comport with Plaintiffs' definition, not Meta's self-serving one.

#### *WEB SCRAPING*

Plaintiffs defined "web scraping" broadly to include the automated collection of online content. Meta attempts instead to redefine the term to mean only "the automated collection of text data from the public internet for use in training the Meta Language Models" and in some instances, limit it to Meta's SpiderMate web archiver. This revised definition improperly inserts limitations that do not appear in Plaintiffs' requests—including restricting the inquiry to text data, public internet sources, and a particular purpose. Meta cannot avoid responding by substituting its own narrower definition for the ones Plaintiffs propounded.

#### *SHADOW LIBRARY*

Plaintiffs defined "shadow library" broadly to encompass all repositories containing unauthorized copyrighted works. Meta instead limited the definition to only a handful of specifically identified repositories. Nothing in Plaintiffs' definition restricts the term in that manner. By replacing Plaintiffs' definition with its own, Meta materially narrows the scope of the requests and withholds information concerning any other repositories that fall within Plaintiffs' definition. Meta must perform its searches and prepare its responses in accord with Plaintiffs' definitions, not its own.

#### **Interrogatory Responses**

As explained to Meta's counsel by U.S. Magistrate Judge Thomas S. Hixon in the pending *Kadrey* litigation, "Plaintiffs are allowed to take discovery into the full scope of their copyright claim, not just part or most of their claim." See *Kadrey*, Dkt. 399, p. 3. A party may not evade discovery by providing partial, qualified, or intentionally generalized responses. See Fed. R. Civ. P. 33(b)(3); Rule 37(a)(4) (providing that an evasive or incomplete disclosure, answer, or response must be treated as a failure to disclose, answer, or respond).

Meta has served evasive and incomplete responses to all of Plaintiffs' ROGs, failing to answer them as propounded. Plaintiffs agreed to a compressed discovery schedule ending in September in reliance on Meta's commitment to provide these foundational facts promptly, and Plaintiffs' trial preparation depends upon receiving that information without further delay. Meta cannot shift its burden onto Plaintiffs by requiring Plaintiff to provide its copyrighted works before conducting a reasonable search.

#### *ROG No. 1*

Plaintiffs maintain the position that Meta must identify those of Plaintiffs' copyrighted works it obtained from the web. The onus of identifying those works, their associated CMI, and the sources

Jessica E. Phillips

Privileged &amp; Confidential

June 30, 2026

Page 4

from which they were derived, rests with Meta, and Meta alone. Even so, Plaintiffs have already provided Meta with ISBNs and is providing text for many of the works for this purpose. Notwithstanding, Meta has not, to date, identified *any* of Plaintiffs' works obtained from a specified Shadow Library

Meta's response states that it will only be searching its files for Plaintiffs' works that Meta downloaded from Anna's Archive, Books3 and LibGen. But Plaintiffs asked Meta to identify all works obtained from any SHADOW LIBRARIES, which include the aforementioned repositories, but are not limited to those alone. Meta needs to search which works it obtained from any online repository or electronic collection or source of digital copies of copyrighted books, articles, media and similar materials, **including but not limited to** the datasets commonly known as Anna's Archive, Bibliotik, Books3, C4, Library Genesis (or LibGen), Sci-Hub, The Pile, Z-Library, Bespoke, CORE, UCSF, NDLI, and m2c2, SpiderMate GitHub, CCNews, Deeper Web Crawl, and CommonCrawl.<sup>2</sup>

#### *ROG No. 2*

Similarly, Meta may not limit its search of Plaintiffs' works it obtained through WEB SCRAPING to a search of Meta's SpiderMate web archiver. WEB SCRAPING, as we defined in our requests, means the automated, semi-automated, programmatic, or systematic extraction, collection, copying, downloading, crawling, indexing, harvesting, or retrieval of data, text, books, articles, images, metadata, or other content from websites, online repositories, databases, shadow libraries, torrent indexes, APIs, or internet-accessible sources for purposes including, without limitation, aggregation, storage, processing, Dataset creation, Corpus creation, or training, fine-tuning, evaluating, benchmarking, validating, or developing machine learning systems or Large Language Models.

In addition to its own SpiderMate, Meta has previously disclosed additional sources of web scraping (*see e.g.*, Meta\_Kadrey 000880065 and 00161479) and must search for works obtained through **any** Web Scraping efforts, including **but not limited to** Bespoke, CORE, UCSF, NDLI, and m2c2, GitHub, CCNews, Deeper Web Crawl, and CommonCrawl.

#### *ROG No. 3*

Meta's unilateral limitations outlined in response to ROG Nos. 1 and 2 are likewise unwarranted for any response to ROG 3. Meta must identify when and how it obtained each work from a SHADOW LIBRARY or WEB SCRAPING, and identify the source or sources, date or dates of acquisition, and means or method of acquisition.

---

<sup>2</sup> Meta's interrogatory response is narrower even than its own already narrowed definition. Meta objects to the term SHADOW LIBRARY saying in its definitional objections that it will construe the term to include Anna's Archive, Bibliotik, Books3, Library Genesis (LibGen), Sci-Hub, The Pile, and Z-Library but then says in its interrogatory response that Meta will only conduct a search of Anna's Archive, Books3, and LibGen.

Jessica E. Phillips  
June 30, 2026  
Page 5

Privileged & Confidential

*ROG No. 4*

Meta has objected to this Interrogatory arguing that the term “train” is vague and ambiguous and that it will interpret the term “train” and “the act of pre-training and/or fine-tuning a large language model.” Plaintiffs do not accept Meta’s unilateral narrowing of this interrogatory. Deferring a substantive response to unspecified future date is not compliant with Rule 33, which does not permit a party to avoid answering an interrogatory by promising future supplementation. Meta necessarily possesses information concerning the datasets used to train its models and has represented that it is investigating which Plaintiffs’ works are contained within those datasets. The interrogatory seeks the results of that investigation as they presently exist—not only those results Meta chooses to disclose after its investigation is complete.

Plaintiffs are entitled to know which of its works were used, in whole or in part, to train any of Meta’s LLM models. In this sense, Meta should disclose which of Plaintiffs’ works were used to pre-train, train, fine-tune, or post-train any of Meta’s LLM models. To be clear, the term “train” includes all processes taken after Meta’s acquisition of the Work (or any part thereof), *including but not limited to* the process of teaching any model (commercial or otherwise) to understand and digest the data. This necessarily includes ablations of the copyrighted works.

*ROG No. 5*

Meta’s response to this ROG suffers the same deficiencies as those outlined above and inappropriately narrows the obligation to respond to a specific inquiry with a “high-level general description” instead of a straightforward answer. Plaintiffs request a complete response to identify, on a work-by-work basis, each instance when a work was used to train an LLM model (as defined by Plaintiffs, not Meta), including when and how.

*ROG No. 6*

This is a straightforward ROG calling for a straightforward answer asking Meta to identify each work from which it removed or altered any Copyright Management Information. This information is solely within Meta’s possession and control and it cannot delay further response by suggesting that it may need additional information from Plaintiffs to identify responsive information.

Whether Meta removed or altered Copyright Management Information, directly or through third parties acting on its behalf, is uniquely within Meta’s possession, custody, or control. Meta therefore must identify responsive works based on the information presently available.

*ROG No. 7*

Plaintiffs are willing to meet and confer regarding this scope of this request and to better understand Meta’s concerns.

Jessica E. Phillips  
June 30, 2026  
Page 6

Privileged & Confidential

ROG No. 8

Meta's objection that this ROG is improperly compound lacks merit. The interrogatory seeks a single category of information—the factual basis for each response to Plaintiffs' Requests for Admission that is not an unqualified admission. The requested facts, knowledgeable persons, supporting documents, and corresponding RFA numbers are logically subsumed within that single inquiry and therefore do not constitute separate interrogatories under Rule 33. Moreover, Local Rules 33-2 and 36-2 expressly contemplate discovery concerning the factual basis for denials of requests for admission. Meta's statement that it will respond "to the extent required" by those Local Rules is not a response. Defendant must identify, for each Request for Admission that was denied or only partially admitted, the supporting facts, persons with knowledge, and responsive documents presently known after a reasonable inquiry, and supplement its response if additional information is later discovered.

As invited by Meta's response to this ROG, counsel welcomes further discussion about the scope of Meta's intended supplement.

### **Requests for Admissions**

*RFA No. 1*

Meta has admitted only a different proposition than the one Plaintiffs requested. Meta's qualified admission improperly limits the defined term "shadow library" to a self-selected subset of repositories. Plaintiffs' definition is broader and encompasses any repository falling within the defined criteria, whether or not expressly named. By substituting its own narrower definition, Meta must now supplement its response by applying Plaintiffs' definition of "shadow library" and admitting or denying the RFA accordingly.

*RFA No. 2*

Meta's response to this RFA suffers from the same deficiency as its response to RFA No. 1. By substituting its own narrowed definition of "shadow library," and "web scraping," Meta materially alters the Request and admits only a different, narrower proposition than the ones Plaintiffs served.

*RFA No. 3*

Meta again responds based on its own narrowed definitions of expressly defined terms and fails to answer the Request as propounded. The RFA asked Meta to admit whether it used any of Plaintiffs' Works "in whole or in part." Meta instead limited its response to use of the entirety of each Work. That qualification materially alters the Request and leaves unanswered whether Meta used portions of Plaintiffs' Works.

Jessica E. Phillips  
 June 30, 2026  
 Page 7

Privileged & Confidential

### **Requests for Production**

Plaintiffs have made but one simple request for the production of documents, asking for unredacted copies of all dispositive motion briefings filed in *Kadrey* and *Entrepreneur Media*. To date, Plaintiffs have not received those documents and Meta has objected to providing them in the absence of an ESI Order. This is an unnecessary delay of Meta's own making. Meta has had Plaintiffs' draft of the ESI Order since June 11, 2026 and has not responded to Plaintiffs' repeated requests to finalize the Order. Please provide Plaintiffs' counsel with the status of third-parties notices and a response to the draft of the ESI Order no later than July 6, 2026 to facilitate a production of these documents no later than July 10, 2026.

+++++

Please provide your availability to confer on Thursday July 2 or Monday, July 6 to address these issues. Please provide complete supplemental responses to Plaintiffs' First Set of Written Discovery by July 10, 2026.

Sincerely,

/s/ Lauren E. Martin

Lauren E. Martin