EXHIBIT E

Redacted

From: Jay Schuffenhauer <Jschuffenhauer@bsfllp.com>
Sent: Monday, December 16, 2024 6:24 PM
To: z/Meta-Kadrey; Dunning, Angela L.
Cc: Llama C-Counsel
Subject: Kadrey v. Meta - Discovery Letters for 12/20

[External]

Counsel,

In light of the Friday deadline for raising discovery disputes with the Court, we intend to brief the following topics unless the issues can be resolved via meet-and-confer. These topics all originated from recent depositions and/or Meta's 12/13 document productions and privilege logs. We propose conducting this M&C immediately after Mr. Zuckerberg's deposition concludes tomorrow.

1. **Meta's inadequate supplemental RFA/ROG responses.**

   12/13 RFA Responses

   Plaintiffs intend to submit a letter brief challenging the following RFA responses absent Meta's agreement to amend them:

   - Meta's "subject to and without waiving the foregoing" language: This language obscures whether any of Meta's objections qualify its responses.
   - Meta's "except as expressly admitted" denials: Tacking on unexplained partial denials to most of the RFA responses is improper.
   - RFAs 3-7, 17, 20, 23, 34, 43, 45-91, 94 & 96: Meta does not respond to the RFAs as written. The RFAs ask about copyrighted works and/or books. Meta is not allowed to answer in terms of "text" or "some text." If Meta needs to qualify its admission to account for the CMI that it stripped from these works, then it can say so after admitting the RFAs.
   - Meta's RFA 7 response's limitations to "one or more Datasets" and "one or more copyrighted books" are improper.
   - Meta's RFA 8 response's qualification of "claim in this action to have authored" is improper. If Meta has a basis to believe any particular plaintiff did not author any work at issue in the Datasets used to train Llama then it should say so and deny the RFA as to that plaintiff. But otherwise Meta should admit the RFA without this qualification.
   - Meta's RFA 9 response's qualifications of "purportedly authored" and "to the extent those Datasets included the content of books they purportedly authored" is improper. If Meta has a basis to believe any particular plaintiff did not author any work at issue and/or any Dataset does not include the work, then it can say so and deny the RFA as to that work and/or dataset but otherwise admit the RFA without those qualifications.
   - Meta's "portion of" and "some content" limitations in RFAs 16, 19, and 22 do not answer the question as asked. To the extent Meta denies it uses any portion of Books3, LibGen and/or The Pile then it can say what portions in a partial denial. We note that Meta deponents testified Meta used b3g, which includes Books3. Further, Meta should admit it used all the works in the datasets it produced.
   - Meta's response to RFA 26 does not answer the RFA as written.
   - Meta fails to support its relevance and proportionality objections in response to RFA 35.
   - Meta's response to RFA 39 does not answer the RFA as written.

Further, Meta fails to answer RFAs 38, 44, and 98 but offers to meet and confer. There is nothing ambiguous about these RFAs and Meta must answer them as written. Regarding RFA 38, to the extent you are unsure what "store" means, it means maintain a copy or copies of, however arranged in part or in whole. Regarding RFA 98, to the extent you are unsure what "sourced" means, you can omit that word; the RFA is the same with or without that word.

12/13 ROG Responses

- Meta's Responses to Plfs' ROGs, Set 1
  - These responses continue to improperly limit Meta's responses to Llamas 1-3 only.
  - The responses improperly limit the plain meaning of "agreements," including per Judge Hixon's orders in this case.
  - Meta improperly limits the definition of training data to Books3.
  - ROG 2: Meta failed to respond.
  - ROG 1: Meta's response does not include Llamas 4 or 5.
  - ROGs 3 & 7: It is unclear whether Meta's response includes Llamas 4 and 5. And, regardless, it fails to describe in any detail how the model was fine-tuned as to any Llama model, let alone all of them. The responses similarly is scant on details of the specifics of Meta's efforts around memorization with respect to the Llama models.
  - ROG 4: Meta's response fails to identify or explain all of the "risks" it references or the "mitigation" it supposedly employed.
  - ROG 8: Does not address Llamas 3 and thereafter.
  - ROG 10: Fails to provide the information requested for each copy Meta made of the datasets it uses in pre- and/or post-training.
    ROG 13: Does not address Llamas 4 or 5, and the documents identified do not provide all the information requested.

- Meta's Responses to Plfs' ROGs, Set 2
  - These responses continue to improperly limit Meta's responses to Llamas 1-3 only.
  - The responses improperly limit the plain meaning of "agreements," including per Judge Hixon's orders in this case.
  - Meta improperly limits the definition of training data to Books3.
  - ROG 16: The incorporation by reference without citation at 15:6-8 is improper.

- Meta's Responses to Plfs' ROGs, Set 3
  - These responses continue to improperly limit Meta's responses to Llamas 1-3 only.
  - Meta improperly restricts the definition of shadow datasets to the extent it excludes any dataset it used for any Llama model.
  - Meta improperly limits the definition of training data to Books3.
  - ROG 19: Meta fails to set forth the revenue its anticipated receiving through [redacted]
  - ROG 22: This ROG does not ask Meta for its opinion about what it needed to do or didn't need to do. If Meta didn't do anything then it should just say so and that's the end of it. *See* Dkt. 315 at 5.
  - ROG 23: Meta's response fails to identify what datasets it obtained from Anna's Archive. It's also unclear from Meta's response whether it obtained datasets from any other sources for any of the Llama models or if it's only the three listed.

2. Meta's assertion of attorney-client privilege over business documents.

a. Meta's 12/13 privilege logs appear to reflect numerous redacted and withheld documents whose primary purpose is business-related.

b. Plaintiffs will circulate a full list of challenged documents along with their letter brief.

3. Meta's assertion of attorney-client privilege over documents subject to the crime-fraud exception.

a. Meta's 12/13 privilege logs appear to reflect numerous redacted and withheld documents where the attorney advice, if any, was used in furtherance of a crime or illegal scheme: in this case, the knowing torrenting of copyrighted works from pirated datasets like LibGen and subsequent use of those pirated works in connection with training the Llama Models.
b. Plaintiffs will circulate a full list of challenged documents along with their letter brief.

4. Requests for additional deposition time.

a. Plaintiffs are still reviewing Meta's 12/13 document production, but it already is clear that the production contains some of the most important documents of the case. Based on the file path names, many of these documents were collected by Meta as early as June 2024 but were withheld until the very last night of fact discovery.
b. As of now, Plaintiffs believe that at least the following fact witnesses were deposed at a time where Meta had already collected (but had not produced) critical documents pertaining to their knowledge:
    i. Melanie Kambadur; Sergey Edunov; Nikolay Bashlykov; Joelle Pineau; Eleonora Presani
c. As discussed via earlier emails, the belated production, including of Mark Zuckerberg's documents, may also prompt a request for additional deposition time of Mr. Zuckerberg.

5. Other RFP Refusals.

a. Llama 4 and 5 datasets
    i. Plaintiffs believe these are covered by RFP 81, particularly in light of Judge Hixson's orders regarding Llama 4-5 relevance.
b. Torrenting Materials
    i. In addition to RFP 85, the items Plaintiffs requested via email are also covered by RFP 119 (processing of copyrighted material). Regardless of whether governed by RFP 85 or 119, Plaintiffs believe Meta's bit torrent client, application logs, and peer lists are relevant and should be produced. It also appears at least Mr. Bashlykov may have used personal devices to download/access datasets. See, e.g., Meta_Kadrey_00204223. Please confirm Mr. Bashlykov's personal computer(s) were preserved and searched. Please also confirm whether any other Meta employees used their personal computers to download/access datasets.
c. Supervised Fine Tuning Data
    i. This data is relevant to Plaintiffs' RFPs 118 and 119, which encompass efforts to prevent Llama from emitting copyrighted material as well as Meta's processing of copyrighted material. The supervised fine tuning data directly serves this purpose.

In light of the large quantity of issues, Plaintiffs will submit their briefs to Meta on a rolling basis beginning tomorrow afternoon. All briefs will be with Meta no later than Wednesday at 3 PM PT. Plaintiffs are also open to discussing an omnibus letter brief if Meta would prefer.

Best,
Jay

Jay Schuffenhauer
Associate

BOIES SCHILLER FLEXNER LLP
1401 New York Avenue, N.W.

Washington, DC 20005
(m) +1 952 288 5729
jschuffenhauer@bsfllp.com
www.bsfllp.com

The information contained in this electronic message is confidential information intended only for the use of the named recipient(s) and may contain information that, among other protections, is the subject of attorney-client privilege, attorney work product or exempt from disclosure under applicable law. If the reader of this electronic message is not the named recipient, or the employee or agent responsible to deliver it to the named recipient, you are hereby notified that any dissemination, distribution, copying or other use of this communication is strictly prohibited and no privilege is waived. If you have received this communication in error, please immediately notify the sender by replying to this electronic message and then deleting this electronic message from your computer. [v.1 08201831BSF]