APPENDIX A

#	Pincite (Tr.)	Question Asked	Objection	Answered Substantively?	Topic(s) and Explanation
1	9:19-10:7	"[W]hat is the definition of 'piracy' in your 30(b)(6) capacity here today? What is Meta's definition?"	"Object to form. Scope. I think it's outside the scope."	Yes: "So I believe the company would agree that pirating involves making a profit by redistributing the works that one did not actually -- that one does not have the rights to redistribute."	Topic 1 covers torrenting generally, including the distribution of copyrighted works via the same, which constitutes piracy. Topic 8 covers content in the AIDC, which includes express reference to "pirated sites."
2	11:5-12:20	"Is it Meta's position sitting here today that I could have gone on to LibGen and torrented and redistributed the works of secondary criticism, copyrighted books written within the last few decades of when I was working on my doctoral thesis, that it would have been okay for me to do that?"	"Object to form. Outside the scope." (repeated)	Yes: "Yes, I believe so… There's a material difference between acquiring the work via torrenting and intentionally redistributing it for profit. Certainly the latter, with an emphasis on for profit."	Topic 1 covers torrenting. The question was meant to elicit from Meta's corporate witness his position on whether torrenting is permissible so long as the work torrented is subsequently used for a "fair use." The example of the "fair use" in this hypothetical, was Plaintiffs' counsel's doctoral thesis.
3	13:6-15:10	"So it is Meta's position sitting here today that Meta can torrent from Library Genesis and redistribute copyrighted works, and so long as it does not charge for those works or make money on the redistribution of those works, it's okay?	"Object to form. Outside the scope."	No: "I couldn't answer."	Topic 1 covers Meta's torrenting. Topic 3 covers Meta's approval of the same.
4	63:6-20	"Was it – the approval [for Libgen] given by Joelle Pineau?"	"Object to form. Outside the scope."	Yes: "I have seen Joelle's name referenced."	Topic 1 covers Meta's torrenting, which is how LibGen was acquired in 2022 (the date of this discussion. Topic 3 covers Meta's approval of shadow libraries including Libgen.

5	70:13-17	"Was Meta concerned about the risk of distribution when Roller was torrenting from LibGen?"	"Object to form. Outside the scope."	No: "I don't know"	Topic 1 covers Meta's torrenting. Topic 3 covers Meta's approval of the same.
6	71:6-16	"Why didn't Roller directly download the content he torrented from Libgen?"	"Object to form. Calls for speculation. Outside the scope."	Yes: "I don't think it is he who said he was unable to download it directly. The rest of the record and my general knowledge allows me to say that there aren't easy ways to -- to directly download because these sites are too swamped, and that is why BitTorrent is required."	Topic 1 covers Meta's torrenting, and Topic 3 covers Meta's approval of the same, which necessarily involves the availability of alterative mechanisms of acquisition.
7	83:21:25	"Did Peter Albert ever express concerns about using LibGen as a data source?	"Object to form. Outside the scope."	No: "I don't know."	Topic 1 covers Meta's torrenting. Topic 3 covers Meta's approval of Libgen.
8	89:13-22	"So why would -- why did Meta have this engineer -- when it had had two prior, why does this engineer directly download datasets that Meta had actually already acquired?"	"Object to form. Outside the scope."	Yes: "It's very common for individuals in a company of a hundred thousand people to reinvent the wheel and do something again that somebody else has figured out how to do a better way before."	Topic 1 covers torrenting and this line of questioning is focused on why one of Meta's engineers chose *not* to torrent from LibGen and instead directly download the books there. Decisions to acquire data via alternatives to torrenting fall within the scope of Topic 1.
9	170:13-171:7	"So the testimony is Mr. Lample received approval to acquire data from LibGen, but Meta is not sure who provided that approval, however, the use of torrenting was not considered in that approval process?"	"Object to form. Outside the scope."	Yes: "Other -- two things in there I don't think I said.· So the first one is it isn't that we're not sure. I understand that the business made the decision. And the way decisions are typically made at Meta, they would be made by a small group of leaders, not an individual person. . . . As I understand it, this would be the FAIR leadership team."	Topic 1 covers Meta's torrenting. Topic 3 covers Meta's approval of Libgen.
10	64:2-10	"Is it possible that Mark Zuckerberg approved the use of Libgen?"	"Object to form. Outside the scope. Calls for speculation."	Yes: "I would not anticipate that he did, but I'm -- I wasn't involved. And so I am telling you what the company norms would be, which is that this doesn't strike me as the kind of decision he'd be involved in."	Topic 3 covers Meta's approval of Libgen.

11	65:5-10	"Why was Roller torrenting from LibGen contemporaneously with Lample in Paris?	"Object to form. Outside the scope."	Yes: "Both teams were involved in LLM research. It was a hot area of research at the time."	Topic 1 covers Meta's torrenting.
12	68:21-25	"Is there anyone at the company today who would know how or who gave [Lample] approval [for Libgen]?"	"Object to form. Calls for speculation. Outside the scope."	No: "I don't know."	Topic 3 covers Meta's approval of Libgen.
13	77:18-19	"Does Meta know how many works Bashlykov directly downloaded in 2023 from LibGen [fiction] and nonfiction?"	"Object to form. Outside the scope. Calls for speculation."	No: "If we do, I do not." (limited to personal knowledge)	Because Bashlykov chose to directly download after flagging the legal risks of torrenting, testimony regarding the data he was able to acquire via direct downloading and his motivations for choosing an alternative method to torrenting fall within Topic 1 (torrenting) and Topic 3 (approvals).
14	81:14-20	"So Meta knew in 2023 it was possible to directly download ~6 million works from LibGen [fiction] and nonfiction?"	"Object to form. Outside the scope."	Yes: "No. An employee named Nikolay knew that. It's a company of 100,000 people." (limited to individual employee's knowledge, distinguished from corporate knowledge)	Same as above.
15	104:23-105:5	"Is it Meta's position today that Bashlykov did not know in 2023 how to modify Libtorrent to prevent uploading when he was torrenting?"	"Object to form. Scope."	Yes/No: "He was using code that had some of those modifications on it. Whether he knew, I couldn't tell you."	Topic 1 covers Meta's torrenting.
16	136:4-10	"Did Bashlykov distribute any data when he was directly downloading it from LibGen?"	"Object to form. Scope."	Yes: "When he was directly downloading, my understanding is the only data that he consumed was downstream. He did technically distribute URLs back upstream in order to get them."	Because Bashlykov chose to directly download after flagging the legal risks of torrenting, testimony regarding whether direct downloading carried comparable legal risks to torrenting falls within Topic 1.

17	150:14-25	"At this point in time, July – sorry June 27, 2024, did Meta have a policy regarding torrenting data that self-identified as pirated?"	"Object to form. Outside the scope. Calls for speculation."	Yes: "We -- we -- we had a policy around reviewing what data could be used in training a model. We also knew that various archives contained a mixture of public domain, copyrighted, and things that were ambiguously identified as pirated. And so we would very often review which was which after it had been downloaded."	Topic 1 covers Meta's torrenting and Topic 3 covers Meta's approval of the same.
18	158:15-25	"Was Meta aware that Anna's Archive, LibGen Z-Library, Internet Archive, any of these sources that Meta torrented data from, were either under government enforcement action or the subject of civil lawsuits at any point in time that Meta was acquiring data from these sites?"	"Object to form. Scope."	No: "I feel like that's not the first time I've heard that, but I didn't – I don't recall in reviewing materials for today whether I saw anything to that effect."	One of the exemplar AIDC pages cited in the deposition notice includes an "IP Policy Analysis" that flags federal government watchlists and refers to such sites as "sites known to engage in IP infringement."
19	159:2-10	"Was there a policy at Meta to not source data from sites that the Federal Government was targeting in actions?"	"Object to form. Outside the scope. Lacks foundation."	No: "If there was one, I'm not aware of it."	Whether Meta had a policy not to source data from pirated websites targeted by the federal government falls within Topic 3 (approvals). However, as noted above, one of the exemplar AIDC pages cited in the deposition notice includes an "IP Policy Analysis" that flags federal government watchlists and refers to such sites as "sites known to engage in IP infringement."
20	Tr. 193:7-199:7.	"Did Meta have a policy -- since it was going to deduplicate duplicate copies of books, did Meta have a policy to minimize the acquisition of duplicate copies when torrenting from Anna's Archive or LibGen?"	"Object to form. Scope."	Yes: "You can't do it by looking at the metadata names. They're notoriously wrong. And you can't do it streaming. You need the thing -- the whole thing to do it."	Topic 1 covers Meta's torrenting and Topic 3 covers Meta's approval of the same. This includes whether Meta had a policy not to torrent the same content repeatedly.
21	221:7-14	"Do you know whether Anthropic torrented and distributed copyrighted data from LibGen or Anna's Archive to build these Large Language Models?"	"Object to form. Massively out of scope. Calls	No: "I have seen evidence that these datasets were used by at least two or three competitors. I don't know how they acquired them"	The witness suggested that Meta's reliance on torrenting to acquire AI training data was industry practice. Questions regarding Meta's views on its competitors' torrenting practices fall under Topic 1, therefore. This question is also in scope of Topic 8

			for speculation."		because it covers content within the AIDC, and one of the exemplar AIDC pages cited in the deposition notice includes a discussion of competitors, namely, OpenAI, avoiding crawling sites on the USTR Notorious Markets List.
22	Tr. 250:24-253:7	"Do you think that redistributing copyrighted books to third parties and piracy swarms harms this ordinary market for plaintiffs' works?" "Is it Meta's position today that where there is proof of actual dissemination of plaintiffs' works to third parties this causes market substitution and, therefore, harms plaintiffs in that ordinary market for books?"	"Object to form. Outside the scope."	Yes: "No.. . I don't understand that would be a substitution."	Meta's views on whether the distribution of pirated data via torrenting harms the ordinary market for Plaintiffs' books falls within Topic 1 (torrenting). Examining counsel was also citing Meta's Fair Use Interrogatory response, which discusses market harms from torrenting.
23	270:2-5	"Was the FAIR dataset catalog, was the information on the datasets in that catalog transferred or onboarded into AIDC?"	No objection.	No: "I don't know"	Topic 7 covers the "origin and timing of updates, changes, and content in Meta_Kadrey_00238411 and all other AIDC dashboards related to Library Genesis, Anna's Archive, Z-Library, Internet Archive, or Duxiu."
24	289:21-290:3	"And what are the different purposes that a dataset can be approved for?	No objection.	No: "I don't know the rest of that set of potential approved uses there are."	Topic 6 covers "All "approved usages" or "usages" related to Library Genesis, Anna's Archive, Z Library, Internet Archive or Duxiu."
25	290:4-9	"But sitting here today you are not prepared to tell me what all the different potential approved usages are for datasets in the AIDC?"	"Object to form. Scope."	No: "I am not prepared to tell you that sitting here today."	Topic 6 covers "All "approved usages" or "usages" related to Library Genesis, Anna's Archive, Z Library, Internet Archive or Duxiu." Topic 3 covers Meta's approval pirated datasets.

26	297:16-	"Is [Anna's Archive], in fact, still approved for exploration and training today?	"Object to form. Outside the scope."	Yes: "We have replaced this dataset so your hypothetical -- you're effectively asking me a hypothetical because we have different data with a different quality mix that we now use. . . . We wouldn't use it for reasons I've explained in my prior testimony."	Topic 6 covers "All "approved usages" or "usages" related to Library Genesis, Anna's Archive, Z Library, Internet Archive or Duxiu." Topic 8 covers "The terms, fields, and content of the AIDC dashboards for any and all AIDC dashboards related to Library Genesis, Anna's Archive, Z-Library, Internet Archive, and Duxiu." Topic 3 covers Meta's approval of Anna's Archive.
27	300:22-301:3	"Does -- is this referring to re-downloading the dataset or just getting a copy of the dataset that already exists?"	No objection.	No: "That's a good question. I interpret --I don't know for sure. I can tell you what this says. And I can interpret it to mean that it was, in fact, a request to download in the first place."	Topic 8 covers "The terms, fields, and content of the AIDC dashboards for any and all AIDC dashboards related to Library Genesis, Anna's Archive, Z-Library, Internet Archive, and Duxiu."