EXHIBIT 1

EXHIBIT D

In the Matter Of:

RICHARD KADREY

vs

META PLATFORMS, INC.

JAMES BELDOCK 30(B)(6) - HIGHLY CONFIDENTIAL - AEO

August 28, 2026

CERTIFIED COPY

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

RICHARD KADREY, et al.,

Plaintiffs,

vs.

META PLATFORMS, INC.,

Defendant.
_____/

No. 3:23-cv-03417-VC



VIDEO DEPOSITION OF

JAMES BELDOCK, 30(b)(6)

HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY

Date: August 28, 2026

Time: 10:26 a.m.

Location: Cooley LLP
3175 Hanover Street
Palo Alto, California 94304

Reported by: Jennifer J. Matteo, CSR, RPR, CRR
CSR No. 12139

A P P E A R A N C E S

FOR THE PLAINTIFF(S):

BOIES SCHILLER FLEXNER LLP
BY: MARGAUX POUEYMIROU, ESQ.
MAXWELL V. PRITT, ESQ.
ANNABEL WEINBACH, ESQ.
44 Montgomery Street, 41st Floor
San Francisco, California 94104
mpoueymirou@bsfllp.com
mpritt@bsfllp.com
aweinbach@bsfllp.com

FOR THE DEFENDANT(S):

COOLEY LLP
BY: MARK R. WEINSTEIN, ESQ.
3175 Hanover Street
Palo Alto, California 94304
mweinstein@cooley.com

ALSO PRESENT: TOMMY MADUENA, Videographer
NIKKI VO, In-house counsel for Meta

INDEX OF EXAMINATION

PAGE

JAMES BELDOCK, 30(b)(6)

Examination By Ms. Poueymirou: 7

Examination By Mr. Pritt: 256

INDEX OF EXHIBITS

NUMBER	DESCRIPTION	PAGE
Exhibit 1	Defendant Meta Platforms, Inc.'s Second Further Supplemental and Amended Responses and Objections to Plaintiffs' Second Set of Interrogatories	21
Exhibit 2	Notice of Deposition Pursuant to Fed.R.Civ.P. 30(b)(6) to Meta Platforms, Inc.	27
Exhibit 3	Binder	34
Exhibit 4	One-page document	54
Exhibit 5	LibGen dataset	77
Exhibit 6	Meta_Kadrey_00254952 - Meta_Kadrey_00254955	113
Exhibit 7	Meta_Kadrey_00254952 - Meta-Kadrey_00254955	126
Exhibit 8	Meta_Kadrey_00237553	144
Exhibit 9	Meta_Kadrey_00237731	159
Exhibit 10	Libtorrent documentation	172
Exhibit 11	Meta_Kadrey_00238129	185
Exhibit 12	Incident logs	201
Exhibit 13	Meta_Kadrey_00239021	201
Exhibit 14	Meta_Kadrey_00238347	225
Exhibit 15	Meta_Kadrey_00238408 - Meta_Kadrey_00238415	231
Exhibit 16	Meta_Kadrey_00255018	244
Exhibit 17	Meta_Kadrey_00211852	244
Exhibit 18	Defendant Meta Platforms, Inc.'s Objections and Responses to Rule 30(b)(6) Deposition Notice	256

INDEX OF EXHIBITS (Continued)

NUMBER	DESCRIPTION	PAGE
Exhibit 19	Two-page document	265
Exhibit 20	Meta_Kadrey_00238302	287
Exhibit 21	Meta_Kadrey_00244487 - Meta_Kadrey_00244497	300

AUGUST 28, 2026 10:27 A.M.

P R O C E E D I N G S

--o0o--

THE VIDEOGRAPHER: Good morning. Here begins Media Number 1 of the deposition of James Beldock, 30(b)(6), Volume 1, in the matter of Richard Kadrey, et al., versus Meta Platforms, Inc. This case is in the United States District Court, Northern District of California, San Francisco Division. And the case number is 3:23-cv-03417-VC. 10:26 10:26 10:26 10:27 10:27 10:27

Today's date is August 28th, 2026, and the time is 10:27 a.m. This deposition is taking place at 3175 Hanover Street, Palo Alto, California 94304. 10:27 10:27

The videographer is Tommy Maduena appearing on behalf of Array Legal Services. 10:28 10:28

Would counsel please identify yourselves and state whom you represent, beginning with the noticing attorney. 10:28 10:28

MS. POUEYMIROU: My name is Margaux Poueymirou. I'm with Boies Schiller Flexner, and I'm appearing on behalf of the Kadrey plaintiffs. 10:28 10:28 10:28

MS. WEINBACH: Annabel Weinbach and Maxwell Pritt, also from Boies Schiller Flexner on behalf of the Kadrey plaintiffs. 10:28 10:28 10:28

MR. WEINSTEIN: Mark Weinstein from the Cooley 10:28

LLP law firm for the defendant, Meta Platforms, Inc. 10:28
And with me is Nikki Vo, who is in-house legal with 10:28
Meta. 10:28
THE VIDEOGRAPHER: Would the reporter please 10:28
introduce themselves and then swear in the witness. 10:28
THE REPORTER: Good morning. My name is 10:28
Jennifer Matteo. I am a California-licensed 10:28
stenographer, No. 12139. 10:28
Please raise your right hand. 10:28
JAMES BELDOCK, 10:28
Being first duly sworn by the Certified 10:28
Shorthand Reporter to tell the truth, the 10:28
whole truth, and nothing but the truth, 10:28
testified as follows: 10:28
EXAMINATION 10:28
BY MS. POUEYMIROU: 10:29
Q. Good morning, Mr. Beldock. 10:29
A. Good morning. 10:29
Q. So we are on day three now of depositions. 10:29
The first day you were deposed in your 10:29
30(b)(6) capacity, and I believe yesterday in your 10:29
30(b)(1); is that correct? 10:29
A. That is correct. 10:29
Q. And how do you understand the differences 10:29
between the testimony with respect to 30(b)(6) and 10:29

30(b)(1)? 10:29

MR. WEINSTEIN: And you can give your general understanding. Just exclude from your answer any advice you may have received from counsel. 10:29

THE WITNESS: My layman's understanding of this is that as of 30(b)(6) witness, I am responsible for representing here the company's best understanding on certain facts, whether I know them directly or by research and diligence internally in preparation for this deposition. 10:29

And in the context -- excuse me. 10:30

And in the alternate case of my deposition as a 30(b)(1) witness, I understand that I was being deposed in a personal capacity representing to this, and therefore, testifying to what I personally know as an employee of the company, but not on behalf of the company. 10:30

BY MS. POUEYMIROU: 10:30

Q. Great. 10:30

Have you ever in your life pirated books? 10:30

MR. WEINSTEIN: Object to form. Scope. 10:30

THE WITNESS: I don't know what your definition of "pirated" is. 10:30

BY MS. POUEYMIROU: 10:30

Q. So let's -- we can answer that question 10:30

pursuant to your definition of "pirated."

So what is your definition of "pirated"?

MR. WEINSTEIN: Object to form. Scope.

THE WITNESS: The challenge is that this word is defined differently by many people. My definition is -- is -- involves both -- it would involve making a profit from the redistribution of specific works.

BY MS. POUEYMIROU:

Q. Is your definition today different than Meta's definition today?

MR. WEINSTEIN: Object to form. Scope. Calls for speculation.

THE WITNESS: Given that we don't -- oh, given that Meta doesn't believe it's pirated things, I'm struggling to figure that out on the fly.

BY MS. POUEYMIROU:

Q. You just testified "given that Meta doesn't believe it's pirated things"? Okay.

We'll come back to that, obviously, but what is the definition -- if Meta doesn't believe it's pirated things, what is the definition of "piracy" in your 30(b)(6) capacity here today? What is Meta's definition?

MR. WEINSTEIN: Object to form. Scope. I think it's outside the scope. I'm not going to stop

the witness from answering the question.

THE WITNESS: I'm making this up on the fly. It wasn't something -- I'm making this up on the fly.

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.

BY MS. POUEYMIROU:

Q. Okay. Using that as a context, I'm going to start with a personal anecdote.

Before I became a lawyer, I got my Ph.D., and my doctoral thesis was on late 19th Century British and French writers. I very quickly realized I could not make a living out of that, and I went on to law school. But I taught for about four years after that in New York. The primary works of my thesis, right, works of Oscar Wilde and Walter Pater and Baudelaire were all published in the late 19th Century, between 1850 and 1900. And my doctoral thesis, I think we can agree, was transformative. It was a work of criticism, and I was reading those works to train my mind to write that thesis.

My bibliography at the end of my doctoral thesis was divided between primary and secondary works. So I had the primary works of my 19th Century writers,

and I had all of the secondary works of the academics who write about that time period, the biographers, the cultural critics, many of whom published within the last four decades. I got my Ph.D. in 2008. 10:34

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? 10:35

MR. WEINSTEIN: Object to form. Outside the scope. 10:35

THE WITNESS: I don't think I understand what you mean by going on to LibGen in the context of that question. 10:35

BY MS. POUEYMIROU: 10:35

Q. So why don't I rephrase it, then. 10:35

Is it Meta's position sitting here today that I could have done what Meta did between 2000 -- 2022 and 2024 and used BitTorrent to acquire the works, copyrighted works, of the secondary works of criticism I relied on for my doctoral thesis? 10:35

MR. WEINSTEIN: Object to form. Outside the scope. 10:36

THE WITNESS: Yes, I believe so. 10:36

BY MS. POUEYMIROU: 10:36

Q. And is it Meta's position sitting here today that I could have gone on to LibGen through torrenting, I could have torrented from LibGen and also redistributed all of those works that I was going to read for my doctoral thesis? 10:36

MR. WEINSTEIN: Object to form. Outside the scope. 10:36

And when I make these objections, you can still answer the question. I'm just putting on the record that I think the question is outside the scope. You can answer. We would argue later as to whether that's an individual answer or a corporate answer. But unless I instruct you otherwise, you can still answer her question to the best of your ability. 10:36

THE WITNESS: Understood. 10:37

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. 10:37

BY MS. POUEYMIROU: 10:37

Q. Okay. 10:37

A. Would not be our position. 10:37

Q. So if I understand correctly, you're distinguishing between torrenting and intentionally 10:37

redistributing for profit? 10:37

MR. WEINSTEIN: Object to form. Outside the scope. 10:37

THE WITNESS: Yes. 10:37

BY MS. POUEYMIROU: 10:37

Q. 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? 10:37 10:38 10:38 10:38

MR. WEINSTEIN: Object to form -- object to form. Outside the scope. 10:38 10:38

You can answer. 10:38

THE WITNESS: I didn't say that. 10:38

BY MS. POUEYMIROU: 10:38

Q. Okay. So you tell me. Tell me -- tell me what it is. 10:38 10:38

A. Could we read it back? 10:38

Q. Sure. 10:38

The question that I had, which you said you didn't say that, is: Is it Meta's position sitting here today that Meta can torrent from Library Genesis and redistribute copyrighted works so long as it does not charge for those works or make a profit, and you said "I didn't say that." 10:38 10:38 10:38 10:38 10:38

A. Right. 10:38

Q. So what is it that I did not understand about your statement regarding the intentional redistribution for profit? 10:38

A. Sorry. I was asking you to read back what I said, since that's what I actually said. 10:38

Q. Oh, I'm sorry. 10:39

So you said, "There's a material difference between" -- I think you said "acquiring the works." It says "inquiring" here, but she's writing in real time. 10:39

A. Of course. 10:39

Q. "Acquiring the work via torrenting and intentionally redistributing it for profit." 10:39

A. Right. 10:39

Q. "Certainly the latter with an emphasis on for profit." 10:39

So I just want you to break down for me the distinction between acquiring the works via torrenting and intentionally redistributing them for profit. 10:39

What is okay and what is not okay? 10:39

A. You just parsed it correctly. So acquiring the works via torrenting is very different from intentionally redistributing those same works for a profit. 10:39

Q. What if you intentionally redistribute the 10:39

same works but you don't make a profit? 10:39

MR. WEINSTEIN: Object to form. Outside the scope. 10:39

THE WITNESS: I couldn't answer. 10:40

BY MS. POUEYMIROU: 10:40

Q. Meta's position today is it's not sure whether that would be okay or consistent with its policies? 10:40

MR. WEINSTEIN: The witness's position is that -- again, I'm going to object. Outside the scope. 10:40

THE WITNESS: I couldn't answer. 10:40

BY MS. POUEYMIROU: 10:40

Q. Okay. What do you mean by "for profit"? 10:40

A. I draw a distinction between a verbatim copy of something that is sold for the purpose of transmitting that verbatim copy, and that would be what I would define as selling something -- selling the good for a profit. And I distinguished that from other uses. 10:40

I'm narrowly defining profit, is what I just described, or selling for profit is what I just described. 10:41

Q. So your -- "for profit" doesn't mean the money that Meta has made from its Llama models or the incorporation of its AI technology into its platform. You're literally talking about the money that one could have made or Meta could have made were it to go and sell 10:41

works it was not permitted to sell and the money it would have made from the sale of those works? 10:41

MR. WEINSTEIN: Object to form. Outside the scope. 10:41

THE WITNESS: Corporations and people make money all the time from the transformative use of material that they have read or otherwise consumed. That has helped them build a profitable business. I don't -- I certainly don't think that's a problem. 10:41

BY MS. POUEYMIROU: 10:41

Q. Okay. And to go back to my Ph.D. example, then, just so that I'm clear, it's Meta's position today that it would be consistent with corporate policy that I could go and torrent from LibGen and redistribute all the copyrighted works that I torrented so long as I was using them for a transformative later use, which was my doctoral thesis? 10:41 10:42 10:42 10:42 10:42

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 10:42

THE WITNESS: I didn't say that. 10:42

BY MS. POUEYMIROU: 10:42

Q. Okay. So tell me -- tell me how I -- tell me what is wrong about that statement. 10:42

I'm not selling it. I'm just redistributing it. There's no profit. 10:42

A. For starters, Meta is not a Ph.D. student. Meta did not write a thesis. So none of your false analogy adheres. 10:42

Q. Okay. What about if I was using torrenting from LibGen to acquire copyrighted books that I intended to make a parody of? 10:42 10:43

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 10:43

BY MS. POUEYMIROU: 10:43

Q. It sounds like you don't agree that a doctoral thesis is necessarily a transformative use the way building a large language model is. 10:43

Is that where we're not meeting? 10:43

MR. WEINSTEIN: Object to form. Calls for speculation. Outside the scope. 10:43

THE WITNESS: Your entire analogy is inappropriate. 10:43

BY MS. POUEYMIROU: 10:43

Q. Okay. So what makes it inappropriate? 10:43

A. It's just fundamentally flawed. Meta doesn't write Ph.D. theses. It's just not an accurate characterization of what the company does. 10:43

Q. This is true. 10:43

But is it fair to say that you consider the creation of a large language model a transformative use 10:43

of data?

MR. WEINSTEIN: Object to form. Outside the scope.

THE WITNESS: Would you repeat the question, please?

BY MS. POUEYMIROU:

Q. Yeah.

Earlier today, you testified -- and let me find this.

Earlier today, I asked you: Is it Meta's position sitting here today that I could have done what Meta did between 2022 and 2024 and used BitTorrent to acquire the works, the copyrighted works, of criticism I relied on for identify doctoral thesis.

And you said, "Yes, I believe so."

I'm asking you whether, if I take that one step further because we know that Meta did not just download works but also redistributed them, is it Meta's position sitting here today that I could have done what Meta did between 2022 and 2024 and used BitTorrent to download and redistribute copyrighted works, and that would be consistent with company policy today?

MR. WEINSTEIN: Object to the form. Outside the scope. Lacks foundation.

THE WITNESS: Okay. First of all, you -- your

question has changed. 10:45

BY MS. POUEYMIROU: 10:45

Q. Okay. How has it changed? 10:45

A. You just inserted three or four new predicates. So let's go one by one. 10:45

Could you repeat the first part of your question? 10:45

Q. So my question is: You previously testified it would be okay for me to torrent secondary works of criticism that were copyrighted from LibGen to use for my doctoral thesis? That would be fine? 10:45

A. As I understand it, yes. 10:45

Q. Okay. If you can just state at what point it becomes not fine. 10:45

MR. WEINSTEIN: Objection, Counsel. 10:46

THE WITNESS: I've answered your question on this three times, or at least twice that I recall. 10:46

MR. WEINSTEIN: Let's just pause for a second for the objection. 10:46

Objection. Outside the scope. Asked and answered. Lacks foundation. 10:46

BY MS. POUEYMIROU: 10:46

Q. And that is simply it only becomes not fine not when there's redistribution but only when there is redistribution for profit. 10:46

Is that your testimony -- is that Meta's testimony today? 10:46

A. I didn't say that. 10:46

Q. Okay. So what is it? When does it become not fine? 10:46

MR. WEINSTEIN: Same objections. 10:46

THE WITNESS: Well, it is certainly not fine when there is intentional redistribution for profit of the verbatim text. 10:46

BY MS. POUEYMIROU: 10:46

Q. Okay. What makes something intentional? 10:46

MR. WEINSTEIN: Object to form. Outside the scope. 10:46

THE WITNESS: It's a word commonly used in the English language. 10:47

BY MS. POUEYMIROU: 10:47

Q. It is, and you've put a lot of emphasis on it in this answer. 10:47

So I would like to know what it means for redistribution to be intentional versus unintentional so I can understand your point. 10:47

MR. WEINSTEIN: Object to form. Outside the scope. 10:47

BY MS. POUEYMIROU: 10:47

Q. Do you have to know that you're redistributing 10:47

for it to be intentional? 10:47

A. I think all common interpretations of the word "intentional" involve intent. I'm unaware of circumstances in which one can unknowingly intend something. So, yes, I think a person would have to know they're doing it. 10:47

Q. Okay. Can you pull out tab 39? 10:47

MS. POUEYMIROU: We're going to mark this as Exhibit 1. 10:47

(Exhibit 1 was marked for identification.) 10:47

MS. POUEYMIROU: And this is Meta's Second Further Supplemental and Amended Responses and Objections to Plaintiffs' Second Set of Interrogatories. 10:48

It's a long -- it's a long document. 10:48

Mr. Beldock, I'm only going to be asking you about page 16. 10:48

BY MS. POUEYMIROU: 10:48

Q. Did you review this document in preparation for today? 10:48

Feel free to take some time to look at it. 10:48

(Witness reviews document.) 10:49

THE WITNESS: I did review this document. And I have seen the -- the -- I have not actually seen -- I have seen the page. I have now reviewed the rest of the page. 10:50

A. We don't -- I don't see evidence that it is -- that it does. 11:52

And the -- yeah, sorry. I don't see evidence that it does. 11:52

Q. No problem. 11:52

What happened to it? 11:52

MR. WEINSTEIN: Object to form. Calls for speculation. 11:52

THE WITNESS: I don't know. 11:52

BY MS. POUEYMIROU: 11:52

Q. Do you know what the fairspark data backup is? 11:53

A. No. 11:53

Q. So sitting here today, you don't know who at Meta gave Mr. Roller approval to torrent from LibGen? 11:53

MR. WEINSTEIN: Object to form. 11:53

THE WITNESS: I don't. I think it is quite likely that it was not an individual but a decision made by the business, which typically involves more than one person. 11:53

BY MS. POUEYMIROU: 11:53

Q. Okay. So Meta does not know exactly how Roller came to obtain approval for using BitTorrent to -- 11:53

A. Oh, I see. That's a misconception. One did not require approval to use BitTorrent at this time. 11:53

And, in fact, there are published -- before I came to Meta, which is over a decade ago, Meta was very well known for actually using the BitTorrent protocol inside its own facilities to distribute its own code inside its own data centers. 11:53 11:54 11:54 11:54 11:54

So it was not -- there was no red line around the use of BitTorrent because there were -- are plenty of uses of BitTorrent that are very efficient, and nobody would have thought, "That's a thing I should get approval for." 11:54 11:54 11:54 11:54 11:54

Q. Okay. Were there any -- are there any written policies at Meta regarding the use of BitTorrent, then, since it was pervasively used at the company? 11:54 11:54 11:54

A. There are -- there is now such a policy, which I came to know in the process of my preparation for this. 11:54 11:54 11:54

Q. And is that policy applicable to the work of the individuals in Meta's generative AI program? 11:54 11:54

A. Organizationally, it would be applicable, but temporally, it did not exist at the time in question. 11:54 11:55

Q. When did it come into existence? 11:55

A. My understanding is that it came into existence in 2025. 11:55 11:55

Q. So your -- so Meta has a policy now on torrenting. 11:55 11:55

What does it cover? 11:55

A. It -- we could produce it. I didn't bring it with me today. My recollection from reading it is simply that it requires manager approval. Today. 11:55

Q. What was the impetus for creating this policy? 11:55

A. That, I couldn't tell you. 11:55

Q. So in 2022, Meta had no policy regarding the use of BitTorrent? 11:55

MR. WEINSTEIN: Object to form. 11:55

THE WITNESS: On the contrary. It would have been the only way to access large datasets that would actually work. 11:55

BY MS. POUEYMIROU: 11:56

Q. I don't understand what a statement about access to large datasets has to do with whether Meta, as a company, had a policy regarding torrenting. 11:56

Can you connect those dots for me? 11:56

A. I suspect you do. 11:56

Q. I truly don't. 11:56

First of all, I don't agree with your statements. 11:56

But my question is: You said that a torrenting policy now exists that applies to the generative AI group. 11:56

What is the generative AI group called? 11:56

I know at one point it was FAIR, and then I believe it changes names? 11:56

A. Just to be clear, what I said is it would have applied organizationally because it applies to the company. 11:56

Q. Okay. So the whole company? 11:56

A. It is not a -- 11:56

Q. Okay. 11:56

A. -- generative AI specific thing. 11:56

Q. Okay. And prior to the memorialization of that policy in writing, your testimony is that, as a company, folks at Meta knew how to use BitTorrent? 11:56

A. I didn't say that. 11:57

Q. Okay. 11:57

A. I said that we had used BitTorrent and still do, and some people would, therefore, know how to use it. I also said that there was not a policy that -- that required one to check one's -- with one's manager prior to 2025. 11:57

Q. Okay. So in 2022, you testified that Mr. -- Meta has testified today that Mr. Roller was torrenting from LibGen at the same time as another team.

What was that team? 11:57

A. There were two groups, one in Paris and one in -- I actually don't know where Roller is. I think it 11:57

was New York. 11:58

Q. Okay. 11:58

A. Certainly not in Paris. 11:58

And the records shows that these two events took place roughly around the same time. They -- one of them stopped fairly quickly, which would be the Roller case. The other one is -- hang on. I'll get the name right. 11:58

Q. Guillaume Lample? 11:58

A. I believe it's Guillaume Lample. 11:58

Yes, it is. 11:58

Q. Uh-huh. 11:58

So to go back, while Roller did not obtain approval to torrent because there was no policy to Meta in 2022, what did -- 11:58

A. I -- I don't think I agree with your phrasing there. 11:58

Q. Okay. Give me the right phrasing. 11:58

A. He didn't need to get approval. 11:58

Q. He didn't need to get approval. 11:58

A. So it isn't that he didn't get approval, as in he did not receive -- he asked for and didn't receive it. 11:58

Q. Okay. 11:59

A. It is that there was no approval required. 11:59

Q. And your testimony today on behalf of Meta is Meta doesn't know whether he sought approval for torrenting? 11:59

A. That's why I'm making the distinction. 11:59

Q. Okay. 11:59

A. It would not have -- the torrenting itself, which is a known mechanism for getting large data files, would not have required an approval that nobody had been instructed to seek. 11:59

Q. Okay. And what is Meta's testimony today on who, even if not an individual, if it was a team if it was a group of people, on who gave Roller that approval in 2022? 11:59

MR. WEINSTEIN: Object to form. Lack of foundation. 11:59

THE WITNESS: Sorry. I literally just said there was no approval. 11:59

BY MS. POUEYMIROU: 11:59

Q. Not of the torrenting, but of the -- of using LibGen as a data source. 11:59

MR. WEINSTEIN: Object to form. Outside the scope. 11:59

THE WITNESS: Ah. My recollection is that that work was performed by FAIR, which is -- used to be the Facebook AI Research Lab, which was founded in Paris 12:00

but had branches in the U.S. as well. And I can't tell you whether FAIR now stands for foundational or fundamental AI research. It is one of those 2Fs. Either way, the approval was from that organization. 12:00

BY MS. POUEYMIROU: 12:00

Q. Was it -- the approval given by Joelle Pineau. 12:00

MR. WEINSTEIN: Object to form. Outside the scope. 12:00

THE WITNESS: I have seen Joelle's name referenced. And so that -- unless I sit here and tell you on a person-by-person basis how many times Joelle's name appears in the same section of the binder, I can't tell you exactly which person referenced Joelle. 12:00

It's clear, given that they were both working in this area, that some decision was made by the organization to use these data, and the way you access the dataset this large was to BitTorrent it. It's also clear that they took different technical approaches. You can infer from that that somebody in the organization made the call. 12:01

BY MS. POUEYMIROU: 12:01

Q. But Meta doesn't know who? 12:01

A. Again, we don't tend to have individuals make important technical decisions in -- in a vacuum. We tend to have organizational leadership teams take 12:01

decisions.

Q. Is it possible that Mark Zuckerberg approved the use of LibGen?

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation.

THE WITNESS: 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.

BY MS. POUEYMIROU:

Q. Okay. But just so that we have clean testimony today, in your 30(b)(6) capacity, Meta does not know sitting here today who -- which individuals gave Roller approval to use LibGen as a data source?

MR. WEINSTEIN: Object to form. Outside the scope.

THE WITNESS: I believe that's correct.

MR. WEINSTEIN: Lacks foundation.

BY MS. POUEYMIROU:

Q. Okay. Which models was Roller's data used for training?

A. That -- that we do not know.

Q. Is it possible that they were not used -- that the data was not used for any production model?

MR. WEINSTEIN: Object to form. Outside the scope. 12:02

THE WITNESS: Is it possible, yes. 12:02

BY MS. POUEYMIROU: 12:02

Q. Okay. Why was Roller torrenting from LibGen contemporaneously with Lample in Paris? 12:02

MR. WEINSTEIN: Object to form. Outside the scope. 12:03

THE WITNESS: Both teams were involved in LLM research. It was a hot area of research at the time. 12:03

BY MS. POUEYMIROU: 12:03

Q. Would you say that teams internally at Meta were competitive with each other? 12:03

MR. WEINSTEIN: Object to form. Outside the scope. 12:03

THE WITNESS: Are you asking as a general matter? 12:03

BY MS. POUEYMIROU: 12:03

Q. I'm trying to -- I'm wondering why you had two separate teams that were targeting the same data source from LibGen at the same time rather than just sharing with each other. 12:03

And my question is predicated on that, whether it's because they are competitive with each other, wanting to build models that beat other models 12:03

internally at Meta?

A. These are -- first of all, it's very common to access date sources in parallel. Therefore, at least part of the foundation of your -- of -- the foundation of your question is -- is -- doesn't sit right. That downloading is -- excuse me -- a first step in getting data. When you're building an AI model, people get data first.

Second -- sorry, go back to whether or not I'm supposed to be generalizing or speaking about the specific case -- instance, excuse me.

Q. I would -- my question is whether at the time that Roller and Lample are torrenting from LibGen, I'm curious whether they were doing that at the same time and not sharing data because teams were competing with each other to build better models?

A. Oh. So you are asking me not for the general case but the specific case?

Q. I'm asking specifically in this context.

A. I don't think the company knows exactly why these were done in parallel. And I could say that we -- it doesn't surprise me that work would be done in parallel. This was a hot area of research. It still is.

Q. Meta -- Meta does not know how or who gave

Roller approval to acquire data from LibGen. Does Meta know today who gave Lample approval to acquire data from LibGen? 12:05

MR. WEINSTEIN: Object to form. Outside the scope. 12:05

THE WITNESS: Okay. 12:05

MS. POUEYMIROU: The topic -- Topic 3 is on approvals, Mark. 12:05

MR. WEINSTEIN: Right, but I think there's been approval of using a dataset versus approval for torrenting. I just want to make that distinction clear. 12:05

MS. POUEYMIROU: I mean, they were torrenting the datasets, though, at the time; right? So approval to -- 12:05

MR. WEINSTEIN: I understand what you're saying. There's approval related to whether there was an approval on torrenting as a method of acquisition versus was there approval related to the use of dataset more generally. That's the distinction I'm calling. I would agree that one of those two is within the scope of your topics. The other one I think arguably is not. But I'm not going to stop you from asking the witness the question to the extent he knows. 12:06

THE WITNESS: I'm afraid I'm now lost, so you will have to ask the question again. 12:06

BY MS. POUEYMIROU: 12:06

Q. So my question is -- Meta has testified that it does not know who gave Roller approval to acquire data from LibGen And my question is: Roller was doing this contemporaneous to Lample -- 12:06

A. Right. 12:06

Q. Does Meta know who gave Lample approval? 12:06

A. Okay. So to your -- the predicate to your question, again, I want to reinforce that it wouldn't have been an individual. It would have been a business decision made by the leadership team of the business or some team within the business. 12:06

Secondly, to your second question, I'm going to answer with that predicate, which is that it would be the business that made the decision, not an individual, so for as I know. 12:06

And then I can try to answer your question. 12:07

No. What we have entered as Bates Meta Kadrey 000238836 underscore CT has Lample simply saying, "We have a greenlight for Lib Genesis." 12:07

Q. Is there anyone at the company today that would know how or who gave him approval? 12:08

MR. WEINSTEIN: Object to form. Calls for speculation. Outside the scope. 12:08

THE WITNESS: I don't know. 12:08

BY MS. POUEYMIROU:

Q. Do you know how much data Roller distributed when he was torrenting from LibGen?

MR. WEINSTEIN: Object to form. Lacks foundation.

You can answer.

THE WITNESS: Back to Roller?

BY MS. POUEYMIROU:

Q. Uh-huh.

A. Okay. The evidence relating to the amounts of uploaded information can -- would require a calculation using the Amazon Web Services, AWS, cost and usage information, which is produced at Kadrey 00238273. There are 43,026 rows in that dataset, and that -- which is for the period in which this downloading, we believe, occurred. There are also the EWR logs.

As I understand it --

Q. The EMR?

A. Did I say W? I did. It is EMR, which is Elastic MapReduce.

Q. Uh-huh?

A. As I understand it, what one would have to do is to statistically analyze what was available to the server at the time -- at the moment in question of the billing log, and then look at the upload and download

rate, and then infer from that some percentage of what could have been as an absolute maximum uploaded. That's what our expert, I believe, has been asked to do. 12:10

I think that's the answer to your question. 12:10

Q. So that is the answer to how one would reach the uploading amount that corresponds to Roller's torrenting. 12:10

So Meta sitting here today does not know how much data Roller torrented yet? You're saying that that is going to be something that the -- Meta's expert will handle? 12:10

A. Yes. 12:10

Q. Was Meta concerned about the risk of distribution when Roller was torrenting from LibGen? 12:10

MR. WEINSTEIN: Object to form. Outside the scope. 12:11

THE WITNESS: I don't know. 12:11

BY MS. POUEYMIROU: 12:11

Q. Why do you not know? 12:11

MR. WEINSTEIN: Object to form. 12:11

THE WITNESS: Because I wasn't present. 12:11

BY MS. POUEYMIROU: 12:11

Q. But sitting here today in your 30(b)(6) capacity, as the corporate deponent, did you review anything or speak to anyone that flagged a concern that 12:11

the company had about the risk of distributing copyrighted data to third parties while torrenting? 12:11

MR. WEINSTEIN: Object to form. 12:11

THE WITNESS: No. 12:11

BY MS. POUEYMIROU: 12:11

Q. Okay. Why didn't Roller directly download the content he torrented from LibGen? 12:11

MR. WEINSTEIN: Object to form. Calls for speculation. Outside the scope. 12:11

THE WITNESS: 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. 12:11 12:12 12:12

BY MS. POUEYMIROU: 12:12

Q. Let's pull up the supplemental interrogatory response that we started with again. 12:12

A. Okay. 12:12

Q. That was Exhibit 2. 12:12

MR. WEINSTEIN: One was the interrogatory responses. 12:12

MS. POUEYMIROU: Number 1, Exhibit 1, yeah. Thank you. 12:12

THE WITNESS: Do you want me to look at 12:12

interrogatory number 2? 12:12

BY MS. POUEYMIROU: 12:12

Q. We're going to to go back to page 16. 12:12

So you just testified, "My general knowledge" -- "the rest of the record and my general knowledge allows me to say that there aren't easy ways to directly download because these sites are too swamped, and that is why BitTorrent is required." 12:12 12:12 12:13

What do you mean by "these sites are too swamped"? 12:13 12:13

A. First of all, I should say I don't know whether there is a direct download available. That's the first predicate here. And I should have said that earlier. But let's start with that. 12:13 12:13 12:13 12:13

Q. So. Did you -- are you familiar with an employee at Meta called Nikolay Bashlykov? 12:13 12:13

A. I am. 12:13

Q. Did you speak with him in preparation for today? 12:13 12:13

A. I did, and he referenced direct download not working. I don't know whether all of it was available for direct download. 12:13 12:13 12:13

Q. Okay. So Mr. Bashlykov directly downloaded all of LibGen and nonfiction. 12:13 12:13

A. Did he? 12:13

Q. Yes, he did. 12:13

A. Oh. I hear you saying that. Please allow me to -- 12:13 12:14

Q. In fact, we can -- let's pull up Exhibit 6. 12:14

A. Okay. So, counselor, what -- your assertion is inconsistent with what I have in my binder. 12:14 12:14

Q. Let's look at your binder. Let's look at your binder. 12:14 12:14

Which -- which document are you looking at right now? 12:14 12:14

A. So I'm in the tab that begins 1.3, Bashlykov. 12:14

Q. Okay. 12:14

A. Sorry, it doesn't begin. That is what is on the tab. 12:14 12:14

If you go to the first page, there's a summary of Meta dataset downloads. This was provided by counsel. 12:14 12:14 12:14

And the first block reads "Nikolay Bashlykov's download (torrent)," and then there's another block underneath it which reads "non torrenting downloads of dataset data, Nikolay Bashlykov direct download." 12:14 12:14 12:14 12:14

Do you -- do you see that? 12:14

Q. Uh-huh? 12:14

A. And that says data "from" LibGen fiction and LibGen SciTech. I don't know whether that means they 12:14 12:14

were complete. 12:14

Q. Oh. So, yes, this page shows that Bashlykov torrented SciMag, but directly download LibGen and nonfiction? 12:14 12:15 12:15

A. Data from. It doesn't say the whole thing. 12:15

Q. Okay. 12:15

A. That you said entire. 12:15

Q. So I will represent to you and we will look at other documents that it was the entire library of fiction and nonfiction. 12:15 12:15 12:15

And so my question to you is: On what basis are you saying in your answer -- give me a moment. 12:15 12:15

So Meta -- Meta did know that direct download was available, right, because it's engineer directly downloaded data in 2023? 12:15 12:15 12:15

A. Yes. 12:15

MR. WEINSTEIN: Object to form. 12:15

BY MS. POUEYMIROU: 12:15

Q. Okay. And Meta -- 12:15

A. I just said I didn't know. 12:15

Q. Okay. Yes. 12:16

So I've been a little sloppy with the "yous" and the "Metas," but I'm always -- I'm always focused here on -- on Meta's knowledge. So Meta did know that. 12:16 12:16 12:16

And Meta knew that in 2022, the option to 12:16

directly download or torrent existed with respect to Library Genesis; right? 12:16

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 12:16

THE WITNESS: Okay. So, first of all, no. We'll do these in your order. 12:16

Would you reask your question, please? 12:16

BY MS. POUEYMIROU: 12:16

Q. We know that Meta knew in 2023 that it could directly download LibGen fiction and nonfiction because Mr. Bashlykov did that, and -- 12:16

A. The record makes it clear that that was quite difficult. And it's unclear whether that actually worked completely. 12:17

Q. Where are you seeing it -- what document are you looking at? 12:17

A. So let's take a look at Bates No. 00231275. And in the middle of this message between Nikolay and Aurelien he says, "The con is that it is quite slow, and I am regularly throttled by the server." 12:17

Q. I'm still looking to find this document, sorry. 12:17

MR. WEINSTEIN: It's at the end of the Bashlykov tab, the next to last blue sheet. 12:17

MS. POUEYMIROU: Okay. Hold on one second. 12:17

BY MS. POUEYMIROU: 12:18

Q. Okay. Can you actually read to me that passage from Mr. Bashlykov? 12:18

A. Uh-huh. "I'm loading the library file by file using a download link. This way I'm not seeding, sharing any of the files outside, only loading it to the FAIR cluster. The cons is" -- he means "the con is" -- "that it is quite slow, and I am regularly throttled by the server." 12:18

So it was quite difficult, and BitTorrent was clearly the preferred mechanism. 12:18

BY MS. POUEYMIROU: 12:18

Q. Can you finish reading that? 12:18

A. Yeah. "The alternative is to download the files through torrent, but this way we will seed the files from our FAIR cluster servers." 12:18

Q. Okay. Can you continue to the end of this, before Rodriguez writes back in the chat? 12:18

A. "But this would be much faster. Do we have any guidelines on that?" 12:18

Sorry, "I think this would be much" -- "but this would be much faster. I think you unloaded it this way previously. Do we have any guidelines on that?" 12:19

Q. Okay. Now, previously you testified that Mr. Xu -- was it Mr. Xu that asked Roller to stop 12:19

torrenting because it was -- what was happening on the cluster? 12:19

A. This is what leads me to suspect that it's hard both ways, and that BitTorrenting was more efficient and that's why they did it. 12:19

Q. Well, Roller was torrenting at that point so -- 12:19

A. That was -- 12:19

Q. What were the -- 12:19

A. It still sucks. It just sucks less. 12:19

Q. It still sucks. It sucks less. 12:19

Do you know how long it took Bashlykov to directly download all of LibGen fiction and nonfiction -- does Meta know how long? 12:19

A. I still don't know that he did. I know that he did some direct downloading because he said he did. You're telling me he did the whole thing. 12:19

Q. Does Meta know how many works Bashlykov directly downloaded in 2023 from LibGen and nonfiction? 12:19

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 12:20

THE WITNESS: If we do, I do not. 12:20

MS. POUEYMIROU: Can you give me the LibGen notebook? Thanks. We'll enter this into Exhibit 5. 12:20

(Exhibit 5 was marked for identification.) 12:20

THE WITNESS: Okay.

BY MS. POUEYMIROU:

Q. Okay. Did you review this in preparation for today? And this is Exhibit 5, Meta_Kadrey_00089791.

A. Okay. Well, let's -- it's -- what is this? It's 60-some odd pages.

Q. Yes. And I'm just going to direct you to a few, but feel free to familiarize yourself.

A. Did I read every -- every line of this? No, I did not. I did review the overall file.

Q. Okay. Does this refresh your recollection that Mr. Bashlykov directly downloaded fiction and SciTech, which is also called nonfiction?

A. I did not review this document for the purposes of distinguishing direct download from -- from the preferred BitTorrent download.

Q. What did you review this document for?

A. It was mostly to understand elements of deduplication, and I knew that -- this document is in my binder. And specifically also about filtering.

Q. Okay. How many -- how many individual works did Meta directly download from LibGen in 2023, as reflected in this notebook?

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation.

THE WITNESS: I don't know if these were downloaded via BitTorrent or direct download. 12:22

BY MS. POUEYMIROU: 12:22

Q. So in your binder, you have a document that distinguishes between SciTech and fiction and SciMag. And SciMag is the torrented data of Bashlykov's, and SciTech and fiction is the directly downloaded data. 12:22 12:22 12:23 12:23

A. I see. Okay. 12:23

Q. And we know that -- if you look on page ending in 792, he has a chart, where he has total document numbers noted, and it is -- 12:23 12:23 12:23

A. I see. 12:23

Q. It ends with tokenized counts; right? 12:23

A. Okay. 12:23

Q. Is it fair to say -- is it accurate to say that these are the works -- that this tracks the number of works he acquired from LibGen which Meta later tokenized? 12:23 12:23 12:23 12:23

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 12:23 12:23

THE WITNESS: Would you repeat your question, please? 12:23 12:23

BY MS. POUEYMIROU: 12:23

Q. Yes. 12:23

Can you tell me, if this graph tracks the 12:24

number of works that Mr. Bashlykov acquired from LibGen, how many works did he acquire through direct download from fiction and SciTech? 12:24

MR. WEINSTEIN: Same objections. 12:24

THE WITNESS: I'm sorry. I don't see a graph. 12:24

BY MS. POUEYMIROU: 12:24

Q. On page 792? 12:24

A. No. 12:24

Q. So 89792? 12:24

A. This page? 12:24

Q. Yes. I'm calling that a graph. I'm calling that a chart. Sorry. 12:24

MR. WEINSTEIN: You mean a table? 12:24

MS. POUEYMIROU: A table. 12:24

THE WITNESS: It's a table. 12:24

BY MS. POUEYMIROU: 12:24

Q. Excuse me. Okay. 12:24

A. Okay. So reading this table, I now understand this table. It looks like the unique fiction count is 1.6 million. The SciTech count is 3.27 million, 3.3 million. And the SciMag, which was the larger file that he BitTorrented, was 72.6 million. 12:24 12:25 12:25

Q. Okay. In this table, the first row says "total document number." You were reading from the unique document number. 12:25

A. Right?

Q. Did Meta directly download the total documents --

MR. WEINSTEIN: Object to form.

BY MS. POUEYMIROU:

Q. -- or the unique documents, because the numbers are different?

MR. WEINSTEIN: Object to form. Outside the scope.

THE WITNESS: It would have been the total documents.

BY MS. POUEYMIROU:

Q. Total documents. Okay.

So Meta knew in 2023 that it was possible to directly download about 6 million works from LibGen and nonfiction?

MR. WEINSTEIN: Object to form. Outside the scope.

THE WITNESS: No. An employee named Nikolay knew that. It's a company of 100,000 people.

BY MS. POUEYMIROU:

Q. How many people were in this group at this time? And by "this group," I mean, the generative AI group.

How many employees?

A. Zero at this time. 12:26

Q. Okay. So what existed at this time? Was it FAIR? 12:26

A. Correct. 12:26

Q. How many people were part of FAIR at this time? 12:26

A. I don't know. 12:26

Q. It wasn't 100,000, though; is that correct? 12:26

MR. WEINSTEIN: Object to form. Outside the scope. 12:26

THE WITNESS: It was definitely not 100,000, but my point remains. 12:26

BY MS. POUEYMIROU: 12:26

Q. How many people are in this document? 12:26

A. Which document? 12:26

Q. The document you're looking at, the LibGen notebook. 12:26

You reviewed this document in preparation for today? 12:27

A. I did. 12:27

Q. Is Bashlykov the only individual inside of this document speaking? 12:27

A. I haven't reviewed it exhaustively. I can if you want me to. It's a lot of work. It looks like it was mostly Nikolay. If there are other names here. 12:27

Q. If you turn to the back, you'll see comments. 12:27

A. I see. 12:27

Q. Are any of those employees still at Meta today? 12:27

MR. WEINSTEIN: Outside the scope. 12:27

THE WITNESS: You mean the comments from -- first of all, a third of these are from Nikolay himself. 12:27

BY MS. POUEYMIROU: 12:27

Q. Okay. Who else? 12:27

A. Then you've got Melanie Kambadur and Lukas. 12:27

Q. Okay. And then -- 12:28

A. Unless I'm misreading it, those are the three people, one of whom is the author. 12:28

Q. If you go up to the page ending in 00018. 12:28

A. Okay? 12:28

Q. You see the names Nikolay and Peter. 12:28

Do you know who that Peter is? 12:28

A. I do. 12:28

Q. Who is it? 12:28

A. His name is Peter Albert. 12:28

Q. Okay. Did Peter Albert ever express concerns about using LibGen as a data source? 12:28

MR. WEINSTEIN: Object to form. Outside the scope. 12:28

THE WITNESS: I don't know. 12:28

BY MS. POUEYMIROU: 12:28

Q. You also see the name Todor. 12:28

Do you see that? 12:28

A. You'll have to show me where. 12:28

Q. Under the red row saying "still to answer," you have -- sorry, above it, you say -- you see Todor, and you see Peter. 12:28

MR. WEINSTEIN: Council, which page are you on? 12:28

MS. POUEYMIROU: On the same page, 00018. 12:29

THE WITNESS: Oh. 00018. 12:29

MR. WEINSTEIN: Oh, it's in the back section. Okay. 12:29

THE WITNESS: Sorry. Where are we here? On the back of the page? 12:29

MR. WEINSTEIN: Well, there's a bunch of numbering. 12:29

THE WITNESS: No, that 0019. 12:29

MS. POUEYMIROU: So it's the second-to-last page, and you see a bunch of names. It's -- 12:29

THE WITNESS: Oh, I see. 12:29

BY MS. POUEYMIROU: 12:29

Q. Yeah. 12:29

A. So you're on 00018, and there is a Todor in brackets. 12:29

Q. Okay. So Todor, Peter. 12:29

Do you see Mel? 12:29

A. Yeah. 12:29

Q. Do you think that's Melanie Kambadur? 12:29

MR. WEINSTEIN: Object to form. Outside the scope. 12:29

THE WITNESS: It -- it appears to be her -- the shorter version of her name, yes. 12:29

BY MS. POUEYMIROU: 12:29

Q. Okay. And we've also got Lukas Blecher. So we've got at least five people in this document. 12:29

Are you -- sitting here today, are you saying that Meta was unaware that you could directly download from LibGen but these five individuals were? 12:30

A. I didn't say that. 12:30

MR. WEINSTEIN: Object to form. 12:30

BY MS. POUEYMIROU: 12:30

Q. Okay. So can you clarify what you meant by Bashlykov was aware, not Meta. 12:30

MR. WEINSTEIN: Object to form. Outside the scope. 12:30

THE WITNESS: That statement speaks for itself. 12:30

BY MS. POUEYMIROU: 12:30

Q. What is the distinction you're drawing between 12:30

Bashlykov and Meta right now?

A. The difference between an individual human's knowledge and the company's overall distributed knowledge.

Q. I see.

So your position today is that Meta was unaware of what these engineers were doing?

MR. WEINSTEIN: Object to form.

THE WITNESS: I didn't say that, counselor, and you know I didn't say that.

BY MS. POUEYMIROU:

Q. Okay. So tell me what is your position today on Meta's knowledge regarding the works of these engineers?

MR. WEINSTEIN: Object to form. Outside the scope.

THE WITNESS: Generally, regarding the works of these engineers?

BY MS. POUEYMIROU:

Q. I'm trying to understand the distinction you're drawing between what the company knew and what these five people working on data knew?

A. You're the one who asserted that the company knew, and I disagreed with you.

Q. So how does a company know anything?

MR. WEINSTEIN: Object to form. Outside the scope. 12:31

THE WITNESS: A company has general knowledge of facts that are shared, and there are plenty of things that individual employees at a company know that not everybody at the company knows. You asserted the company knew something. I'm disagreeing with your assertion. 12:31

BY MS. POUEYMIROU: 12:31

Q. Okay. So is it Meta's position today that all of these employees were engaging in these efforts to acquire data from LibGen covertly? 12:31

MR. WEINSTEIN: Object to form. Outside the scope. 12:31

THE WITNESS: That is a fabrication of what I did not say. 12:31

BY MS. POUEYMIROU: 12:31

Q. Okay. Meta had meetings all the time. All of these -- Bashlykov has testified in this case multiple times, Kambadur has testified in this case. This is an exhibit that has appeared in multiple depositions. 12:31

Is it Meta's position today that it was unaware, even though you can see it online, that you could directly download from LibGen in 2023? 12:32

MR. WEINSTEIN: Object to form. Outside the 12:32

scope. 12:32

THE WITNESS: I'm telling you that I know this employee knew that. Probably some number of people -- numbered on one hand -- knew that. 12:32

BY MS. POUEYMIROU: 12:32

Q. That direct downloading was possible? 12:32

MR. WEINSTEIN: Same objection. Outside the scope. 12:32

THE WITNESS: Not only did they know that direct downloading was possible, they knew it was painful. They said that on the record we pointed to. And that the more efficient way to do it was to do what everyone else does, which was to BitTorrent. 12:32

BY MS. POUEYMIROU: 12:32

Q. So why didn't he BitTorrent fiction and nonfiction? 12:32

MR. WEINSTEIN: Object to form. Outside the scope. 12:32

THE WITNESS: It's a learning -- first of all, I'm speculating. But it's a learning process. And when one does something that doesn't work, one replaces the thing with something that does work. 12:33

BY MS. POUEYMIROU: 12:33

Q. It didn't have anything to do with the risk of seeding data? 12:33

MR. WEINSTEIN: Object to form. Outside the scope. 12:33

THE WITNESS: The order of -- you've got the order of operations wrong here. He did a thing that others subsequently did via BitTorrent because it was more efficient. 12:33

BY MS. POUEYMIROU: 12:33

Q. Actually, the order of operation is that Roller and Lample in 2022 used BitTorrent, and then Bashlykov decides to directly download? 12:33

A. Much to his pain, and he learned what they already knew. 12:33

Q. 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? 12:33

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 12:34

THE WITNESS: 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. 12:34

BY MS. POUEYMIROU: 12:34

Q. Mr. Beldock, you keep referring to this company of a hundred thousand people, which I understand 12:34

did not consult the libtorrent documentation before using libtorrent? 12:50

MR. WEINSTEIN: Object to form. 12:50

THE WITNESS: I didn't say that. 12:50

BY MS. POUEYMIROU: 12:50

Q. The libtorrent documentation which we're going to look at, makes clear that you can prevent uploading through three simple lines of code, so I'm trying to understand. You've said they didn't know. I spoke with them. Nobody knew. 12:50

A. Hold on a second. I have spoken with them. 12:50

Q. Who have you spoken with? 12:50

A. I spoke with Xiao Lin. I spoke with Nikolay. I spoke with Jacob, who really didn't know about Roller's downloads. I know he knew that he had to stop it in order to get access to the cluster because it was too busy. That's all he knew. 12:51

Q. Okay. And all of these folks that you spoke with said they were unaware that you could stop uploading during the peering phase? 12:51

A. I don't know that they said it. My understanding from speaking to them is that at the time Xiao Lin didn't know and didn't understand the details at all of the protocol and was simply running a script she had been given. 12:51

And that when I spoke to Bashlykov, it was quite clear that at the time he didn't understand the nuances either. There was a paste running around, which is a mechanism of sharing code, that one often uses with details that one doesn't fully understand. 12:51 12:51 12:52 12:52 12:52

And all of this is consistent with my understanding that research engineers, and even those working on data acquisition because fundamentally what they do is work on data preparation -- wouldn't have understood at -- at this level of detail that we're getting to. 12:52 12:52 12:52 12:52 12:52 12:52

Now that they've been deposed as many times as they have, they understand the situation, at least in the case of Bashlykov, better. 12:52 12:52 12:52

Q. It's Meta's testimony today that Bashlykov did not know how to configure libtorrent to prevent uploading during the peering phase -- 12:52 12:52 12:52

MR. WEINSTEIN: Object to form. 12:52

BY MS. POUEYMIROU: 12:52

Q. -- at the time that he was doing it? 12:52

A. I don't know if he did. I'm not surprised if he didn't. 12:52

Q. I'm asking you not as an I. I'm asking you: Is it Meta's position today that Bashlykov did not know in 2023 how to modify libtorrent to prevent uploading 12:53 12:53 12:53

when he was torrenting?

MR. WEINSTEIN: Object to form. Scope.

THE WITNESS: He was using code that had some of those modifications in it. Whether he knew, I couldn't tell you.

BY MS. POUEYMIROU:

Q. How could those modifications exist if he hadn't written them?

A. Oh, people use code snippets all the time. One doesn't always read every line in that code and understand it.

Q. Did you review anything today that showed that Bashlykov actually did configure the upload settings at any point with libtorrent?

A. I did see one where he did.

Q. Okay. What about with Lample?

A. I'll have to check. Okay. So for Lample -- Lample, ah, here's the script.

MS. POUEYMIROU: And we can take a break after this. I know I've been going very long. I'm sorry.

MR. WEINSTEIN: You read our minds actually. Thank you.

MS. POUEYMIROU: We should just break for lunch, right?

THE WITNESS: Okay. So Lample's script is --

THE VIDEOGRAPHER: Time is 2:04 p.m., and we're back on the record. 02:04

BY MS. POUEYMIROU: 02:04

Q. So before we went off the record, Mr. Beldock, you said that -- when I asked you: What did you review today that indicated that torrenting was the most reliable and efficient method of obtaining data when Roller was torrenting? 02:04

You said, "in fact, Bashlykov's testimony -- testimony is the wrong word. The work chat from Bashlykov that backs up my assertion and general knowledge that peer-to-peer transferring is highly reliable versus clogged single-source downloading"; right? 02:04

A. I believe that's what the transcript says. 02:04

Q. Okay. But we did review a document today that showed that Mr. Bashlykov directly downloaded approximately 7 million works from fiction and nonfiction LibGen; is that correct? 02:04

MR. WEINSTEIN: Object to form. Scope. 02:05

THE WITNESS: Yes. 02:05

BY MS. POUEYMIROU: 02:05

Q. Okay. Is it fair to call you a software engineer? 02:05

A. I've been one in my career. 02:05

Q. Okay. How -- I assume that Mr. Bashlykov was not sitting at his computer and tapping each URL.

How did Meta directly download 7 million works from LibGen during that timeframe that Bashlykov was doing it?

MR. WEINSTEIN: Object to form. Beyond the scope.

BY MS. POUEYMIROU:

Q. Did they -- did Meta write a script?

A. Oh. So we don't know. The --

Q. This is with Mr. Bashlykov, Mr. Beldock, who was in the company?

A. I'm sorry.

Q. It's okay.

A. Too many names.

So we're talking about Bashlykov here?

Q. Yes.

A. Then we do know. All right.

So we -- oh, I see. We have only the torrent code here.

Q. And where are you looking in Exhibit 3?

A. This is Bates Meta Kadrey SC.

Q. In Bashlykov's folder?

A. Sorry. Right. So tab 1.3 in this binder, Bashlykov, the first page is the summary prepared by

counsel, then there's a blue page, then you've got Meta hyphen Kadrey hyphen SC hyphen 000201. This is a Python script, which -- excuse me, is clearly -- not clearly. It is the torrenting one. 02:06 02:06 02:07 02:07

Q. Uh-huh? 02:07

A. The next page is the code that gives you the list of torrent files itself. And then the next page is a small print version of the Exhibit 5 that you handed me at the correct -- well, I think at larger size. It's the same content, as far as I can tell, printed at different magnifications and aspect ratios. Or I guess orientation of the page, is the right phrase. 02:07 02:07 02:07 02:07 02:07 02:07 02:07

So that means we -- oh, I'm sorry. It says here -- I just don't have it in front of me, and perhaps you have it -- that we produced the download script download dot PY, which is Python, at SC000197, which I don't have in front of me but does exist. 02:08 02:08 02:08 02:08 02:08

Q. Okay. Do you have -- does Meta have an understanding of how that script works? 02:08 02:08

MR. WEINSTEIN: Outside the scope. 02:08

THE WITNESS: So I haven't prepared -- sorry, I haven't seen prepared this piece of paper. I'm happy to look at it, if you want me to. 02:08 02:08 02:08

BY MS. POUEYMIROU: 02:08

Q. It's okay. I'm just -- I was just curious if 02:08

you had.

MR. WEINSTEIN: Can I just quickly -- I think our position in preparing was the direct downloads, it is not torrenting. It wasn't within the scope. I'm not going to stop you from asking the questions. That's why he's not prepared for it.

MS. POUEYMIROU: Right. And I think our position is that because Meta has taken the position, both today and in its interrogatory response, that torrenting was the most efficient and reliable means of acquiring this data, the way that it acquired through direct download and both actually acquired and the time it took is within the scope.

MR. WEINSTEIN: Look I -- you know my view. I don't like to get in arguments about which topic do you think it falls within the scope of. We can just put it on the record, then we can move on.

MS. POUEYMIROU: Okay.

BY MS. POUEYMIROU:

Q. You said that in your review of Bashlykov's testimony or work chat.

A. I used the wrong word.

Q. Work chat.

A. It was a chat.

Q. That he indicated that torrenting was more

THE WITNESS: This is now entirely hypothetical. 02:31

BY MS. POUEYMIROU: 02:31

Q. Okay. Did Bashlykov distribute any data when he was directly downloading it from LibGen? 02:31

MR. WEINSTEIN: Object to form. Scope. 02:31

THE WITNESS: 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. 02:31

BY MS. POUEYMIROU: 02:31

Q. Okay. If you turn to this document, it's the last one in the Bashlykov folder or tab in your binder. 02:31

A. Just a second, please. 02:31

Q. I would take a moment to review that, if you'd like. 02:31

(Witness reviews document.) 02:32

THE WITNESS: Hmm. Okay. 02:32

BY MS. POUEYMIROU: 02:32

Q. Why did you say, "hmm." 02:32

A. The CAPTCHA solving is an important wrinkle here. 02:32

Q. Okay. And what is that? 02:32

A. CAPTCHA is -- I don't remember what it stands for, but it's -- it's a mechanism, we've all done them, 02:32

THE WITNESS: I can tell you what the document says, which is the description indicates that the books are pirated. Again, we don't actually know what the word "pirated" means in this context or that all the books in there were.

BY MS. POUEYMIROU:

Q. Okay. In 2024, you were the head of data acquisition, weren't you?

A. No. I was responsible for the product management functions around data strategy. There -- I -- there was a, later in the year, a strategy around which data to acquire, but in -- at no point was I uniquely responsible for acquisition of data.

Q. Okay. At this point in time, July -- sorry June 27, 2024, did Meta have a policy regarding torrenting data that self-identified as pirated?

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation.

THE WITNESS: 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.

BY MS. POUEYMIROU: 02:52

Q. And how did that policy distinguish between what you've referred to as archives that contained copyrighted works and pirated works? 02:52 02:53 02:53

MR. WEINSTEIN: Object to form. Outside the scope. 02:53 02:53

THE WITNESS: That would have been -- that would have been a matter that we reviewed with legal. 02:53 02:53

BY MS. POUEYMIROU: 02:53

Q. Okay. But it was a -- it was a policy; right? You said it was a policy. 02:53 02:53

A. We -- 02:53

MR. WEINSTEIN: Object to form. Scope. 02:53

THE WITNESS: We had a policy that there was a review. 02:53 02:53

BY MS. POUEYMIROU: 02:53

Q. I see. 02:53

So this is not a policy that was conveyed to engineers or research scientists who were working in data acquisition groups? 02:53 02:53 02:53

MR. WEINSTEIN: Answering her question, obviously, exclude any advice you may have received from counsel or anyone else may have received from counsel. 02:53 02:53 02:53

THE WITNESS: Then I can't answer the question. 02:53 02:53

BY MS. POUEYMIROU: 02:53

Q. I'm not asking for legal advice. I'm asking about an actual policy that would have been shared with folks at Meta who were tasked with acquiring data. 02:53 02:54

MR. WEINSTEIN: Again, if the policy was animated by legal advice -- which I can't determine, only you can -- I would ask you to exclude that from your answer. 02:54 02:54 02:54 02:54

THE WITNESS: There was a review. And whether the -- yeah, I'm sorry. That's all I can tell you. 02:54 02:54

MS. POUEYMIROU: I mean, aren't all policies animated by legal advice -- or legal advice, or many of them, but the fact of the policy is not privileged, Mark. 02:54 02:54 02:54 02:54

MR. WEINSTEIN: It's the content of the policy that's the issue. And your questions are getting into that, that the witness has already testified implicate legal advice. 02:54 02:54 02:54 02:54

BY MS. POUEYMIROU: 02:54

Q. So was there a policy that was written and shared with engineers or with employees at Meta who were tasked with acquiring data that guided them on the kind of data they could acquire? 02:54 02:54 02:54 02:54

A. There was a review process. That review process involved lawyers. 02:54 02:55

Q. Okay.

MS. POUEYMIROU: Can you give me the other exhibit? 03:02

BY MS. POUEYMIROU: 03:02

Q. In your preparation for today, Mr. Beldock, did you review any document that -- did you review any document where Meta knew that some of the websites it was trying to get data from were banned? 03:02

MR. WEINSTEIN: Object to form. Outside the scope. Lacks foundation. 03:02

THE WITNESS: I think I would need to know what you mean by "banned" and who was doing the banning in order to give you a precise answer. 03:02

BY MS. POUEYMIROU: 03:02

Q. 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? 03:02

MR. WEINSTEIN: Object to form. Scope. 03:03

THE WITNESS: 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. 03:03

BY MS. POUEYMIROU: 03:03

Q. Was there a policy at Meta to not source data from sites that the Federal Government was targeting in actions? 03:03

MR. WEINSTEIN: Object to form. Outside the scope. Lacks foundation. 03:03

THE WITNESS: The U.S. Federal Government? 03:03

BY MS. POUEYMIROU: 03:03

Q. Yes. 03:03

A. If there was one, I'm not aware of it. 03:03

MS. POUEYMIROU: We'll put Exhibit 9, Beldock Meta_Kadrey_00237731. 03:04

(Exhibit 9 was marked for identification.) 03:04

MS. POUEYMIROU: We can break after this one. I know we've been going a long time. 03:04

THE WITNESS: Okay. Yes. I have not seen this before, but I'm looking at it now. 03:05

BY MS. POUEYMIROU: 03:05

Q. Okay. So in your preparation for today, you reviewed evidence with respect to Ms. Wang's torrenting from Anna's Archive and LibGen; correct? 03:05

A. And I spoke to her. 03:05

Q. And you spoke to her. 03:05

What did you learn from that discussion? 03:05

MR. WEINSTEIN: Object to form. 03:05

BY MS. POUEYMIROU: 03:38

Q. Who approved Lample's torrenting from LibGen? 03:39

MR. WEINSTEIN: Object to form. Lacks foundation. 03:39

(Witness reviews document.) 03:40

THE WITNESS: It isn't clear exactly who. It is clear that they were approved to download the data. But, again, I don't think anybody approved the use of the libtorrent -- sorry, not libtorrent. Let's do this right. The -- torrenting is the preferred access mechanism. 03:40

BY MS. POUEYMIROU: 03:40

Q. 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? 03:40

MR. WEINSTEIN: Object to form. Outside the scope, in part. Calls for speculation. 03:41

THE WITNESS: 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. 03:41

/ / / 03:41

BY MS. POUEYMIROU: 03:41

Q. And at this time, who would have been among those small group of leaders? 03:41

MR. WEINSTEIN: Object to form. Outside the scope. 03:41

THE WITNESS: As I understand it, this would be the FAIR leadership team. 03:41

BY MS. POUEYMIROU: 03:41

Q. Who was that, in addition to Joelle Pineau I think? 03:41

A. I don't know. We can find out. 03:41

Q. Okay. 03:41

MS. POUEYMIROU: That would be great, Mark, if we can follow up on that. 03:41

MR. WEINSTEIN: You probably have asked that question on multiple depositions. You probably already have that information. But I don't think it's in the scope. 03:41

MS. POUEYMIROU: Who approved? 03:41

MR. WEINSTEIN: No. Who the leadership team of the FAIR group is. 03:41

MS. POUEYMIROU: But Mr. Beldock just said they are the ones who approved, and Topic 3 is approval. 03:42

MR. WEINSTEIN: Let's take this offline. I -- are you asking me -- 03:42

My understanding is that the pound sign followed by human readable code would just be a comment? 04:06

A. That's right. So 129 is comment. 04:06

Q. Yeah? 04:06

And the comment is saying only include and exclude; is that correct? 04:06

A. Well, what it says is extensions to include slash exclude. 04:06

Q. Uh-huh. Do you know what it means by an extension there? 04:06

A. I believe I do. 04:06

Q. What is that? 04:06

A. An extension is a -- typically the last few letters of a filename after a period. 04:06

Q. So like dot ePub? 04:06

A. Yes, exactly. 04:06

Q. Okay. And if you go to the next page, 229 to 241. 04:06

Is it accurate to say that that code is saying to check what pieces -- what piece numbers in torrents are particular extension? 04:07

MR. WEINSTEIN: Object to form. 04:07

THE WITNESS: This is the first time I've read the code at this level of detail, and you should -- the company's answer should come from a Python and 04:08

networking expert. My initial read is that this portion of the code filters which extensions on which files will be kept. 04:08

BY MS. POUEYMIROU: 04:08

Q. And when you say "filters," you mean that -- as I stated earlier, and I don't think it's in dispute, the torrent files, for the most part, that were torrented at this time from LibGen contained 1,000 works inside of each. 04:08

Are you saying that you read this code to be in those 1,000 works in that torrent file, you could include or exclude torrenting ePubs or dot MOBIs or something like that? 04:08

MR. WEINSTEIN: Object to form. Scope. Calls for expert testimony. 04:09

BY MS. POUEYMIROU: 04:09

Q. You can also look at lines 257 to 260, set zero priority no download to pieces that were not selected. 04:09

A. Yeah. So I'll reiterate my statement that I -- at least I don't think I am required to be able to read code live as a company witness. We do have experts whom we've hired to do that, and you are free to ask them. 04:09

Q. Okay. Fair enough. 04:09

A. It's clearly filtering, and it appears to be filtering by extension. 04:09

I want to stipulate that that's my personal read, and the company's opinion should come from the expert who can confirm that. 04:09

Q. And at any point between 2022 and 2024, did Meta have a policy to not torrent duplicative data? 04:09

MR. WEINSTEIN: Object to form. Scope. 04:10

BY MS. POUEYMIROU: 04:10

Q. In other words -- 04:10

A. Sorry. Please explain what you mean by that. 04:10

Q. So I know that Mark is going to object to this being outside of the scope. 04:10

But as a context, why does Meta deduplicate data when training? 04:10

MR. WEINSTEIN: Object to form. Scope. 04:10

THE WITNESS: Okay. Let me reset here. 04:10

Why does Meta -- training a model on duplicative text overfits the model to that particular type of duplicate content. 04:10

BY MS. POUEYMIROU: 04:10

Q. Is it accurate to say that deduplication is a copyright mitigation at Meta? 04:10

MR. WEINSTEIN: Object to form. Outside the scope. 04:11

THE WITNESS: I'm sorry. Would you say that again? 04:11

BY MS. POUEYMIROU: 04:11

Q. Is it accurate to say that deduplication is one of Meta's copyright mitigations? 04:11

MR. WEINSTEIN: Outside the scope. Substantially out of scope. 04:11

THE WITNESS: I -- I've never thought of it that way. Deduplication is important to avoid overfitting the model. It's important to avoid skewing the dataset that trains the model. If there's a copyright ramification, I'm not aware of it. 04:11

BY MS. POUEYMIROU: 04:11

Q. So my understanding and training is that you're using the term "overfit" -- one other way to look at it is that the more a model trains on the same data, the more it tends to memorize and regurgitate that data. 04:11

But I think we can at least agree on the fact that Meta does -- deduplication is something that Meta does as part of training; is that fair? 04:12

MR. WEINSTEIN: Object to form. Outside the scope. 04:12

THE WITNESS: Oh, I see. All right. 04:12

I disagree with your first assertion because at the scales of pretraining, the -- the repetition that 04:12

you're referring to would not train the model to regurgitate. It would, instead, simply train the model to an inaccurate statistical view of the corpus. And inaccurate statistical view of the corpus would result in poor model performance. 04:12 04:12 04:12 04:12 04:13

Again, we're talking about one -- multiple copies of a single element of a corpus that is as large as 60 trillion tokens, where a book is on the order of maybe 100,000 plus. 04:13 04:13 04:13 04:13

So, no, your first assertion part is just not right. That's not -- the overfitting of the model to skewed data is the problem. 04:13 04:13 04:13

BY MS. POUEYMIROU: 04:13

Q. Okay. 04:13

A. Now, you'll have to forgive me, but I was concentrating on that and didn't get the second part of the question. 04:13 04:13 04:13

Q. So as a factual matter, deduplication is what it says it is; right? It's removing duplicate copies of data from training data or raw data or any data, but it is the removal of a duplicate; is that correct? 04:13 04:13 04:13 04:13

MR. WEINSTEIN: Object to form. Outside the scope. 04:13 04:13

THE WITNESS: I see this changes sides of the page. Let me find it. 04:14 04:14

Here. Yeah. So I agree with your assertion that deduplication removes duplicates, by definition. And a good example here would be the removal of excessive whitespace, which I think happened, you know, on the order of hundreds of millions of times. 04:14

BY MS. POUEYMIROU: 04:14

Q. Uh-huh. Okay. 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? 04:14

MR. WEINSTEIN: Object to form. Scope. 04:15

You can answer. 04:15

THE WITNESS: Those -- that's a false equivalence because the downloading is done from a source with unknown content. And the Meta data within these libraries is notoriously bad. 04:15

And so the way you do the deduplication is not at the download stage because you don't know if it actually is a duplicate. You don't, in some cases, even have the metadata to tell you what it is. It is done after you have the entire corpus download -- sorry, corpus is the wrong word in this context. It is done after the entire archive is downloaded, and then you use a process called MinHash or an equivalent process which 04:15 04:16 04:16 04:16

measures the distance at the text level between the various elements of the content and deduplicates at that level. 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. 04:16

BY MS. POUEYMIROU: 04:16

Q. So it's Meta's testimony today that -- let me back up. Strike that. 04:16

MR. WEINSTEIN: Yeah. This is not Meta's testimony. You're so far outside the scope. I've kind of lost where we are. 04:16

MS. POUEYMIROU: It's the torrenting of data and the approval of torrenting duplicate copies. So it's a -- it's a Topic 1, 3. 04:17

MR. WEINSTEIN: I respectfully disagree with you, but you can proceed. 04:17

BY MS. POUEYMIROU: 04:17

Q. Doesn't Meta use the metadata that it acquires from libraries to identify works -- 04:17

MR. WEINSTEIN: Object to form. 04:17

MS. POUEYMIROU: -- that it has? 04:17

MR. WEINSTEIN: I'm sorry. Object to form. Outside the scope. Calls for speculation. 04:17

THE WITNESS: My understanding is that that 04:17

metadata -- those metadata are notoriously inaccurate and highly unreliable for the purpose that you are describing. The duplication assessment is performed after download in order to do it without being dependent on noisy metadata. 04:17 04:17 04:17 04:17 04:18

BY MS. POUEYMIROU: 04:18

Q. Is Meta's position today that it had a -- the question that I asked you was whether Meta had a policy to minimize the torrenting of duplicate data between 2022 and 2024? 04:18 04:18 04:18 04:18

A. You keep using the word "policy." 04:18

MR. WEINSTEIN: I'll interpose. Outside the scope. 04:18 04:18

You can answer. 04:18

THE WITNESS: These were the early research days of an entirely new field. Using the word "policy" is already giving this more weight than it deserves. These were research efforts to figure out what data would produce a better model in a world where we didn't even know how to produce that model. 04:18 04:18 04:19 04:19 04:19

The idea that there would be formal policies of the nature that you've just described, which, first of all, I've explained wouldn't have worked technically, itself is flawed because this was early days of research where the primary objective was to get the data and 04:19 04:19 04:19 04:19 04:19

it -- and learn how to build a model that does these things. 04:19

The -- we have spent more time talking just today about the details of this process than the research engineers did, whose entire purpose was to get the data and use it. There would not have been a policy established because nobody had ever thought of this question at the time to do the thing that you're describing. 04:19 04:19 04:19 04:20 04:20

And I still tell you that the thing you're describing would not have been technically the right way to do it, based on my understanding of what we now know four years later, is the best way to measure duplication in the first place. This was early research. It's messy. People do things more than once. They learn from it. 04:20 04:20 04:20 04:20 04:20 04:20

BY MS. POUEYMIROU: 04:20

Q. So every file -- or many files in these libraries have MD5 hashes associated with them. 04:20 04:20

Do you know what an MD5 hash is? 04:20

A. I do. 04:20

Q. Is it your understanding that those uniquely identify the works that they are attached to? 04:20 04:20

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 04:20 04:20

THE WITNESS: They mostly uniquely identify. 04:20
MD5 Hash collisions do happen. 04:20
BY MS. POUEYMIROU: 04:21
Q. How often do they happen? 04:21
A. It's a birthday paradox phenomenon. When 04:21
you're dealing with millions of files, they are actually 04:21
not that uncommon. 04:21
Q. So Meta's own expert in this case has 04:21
testified that it is exceedingly -- I mean, it is -- 04:21
I think what she said was as common as twice every two 04:21
years or some -- does that sound possible to you? 04:21
MR. WEINSTEIN: Object to form. Outside the 04:21
scope. 04:21
THE WITNESS: First of all, if we're going to 04:21
get into a mathematics debate on MD5 Hashes, I defer to 04:21
our expert. 04:21
BY MS. POUEYMIROU: 04:21
Q. It would have been possible for Meta to only 04:21
torrent files that were uniquely identified by hashes 04:21
that didn't match hashes Meta had already acquired. 04:21
Is that fair to say? 04:21
MR. WEINSTEIN: Object to form. Outside the 04:21
scope. 04:21
THE WITNESS: It's ludicrously unfair to say. 04:21
/ / /

BY MS. POUEYMIROU: 04:21

Q. Why? 04:21

A. This is a small group of researchers who have been asked to go get data. What you're describing is a high-end hash-based deduplication system that nobody would have invested in in the early stages, when nobody really knew if any of these data were going to be useful in the first place. And torrenting them was the most expeditious way to get all the things that needed to be gotten. 04:21 04:22 04:22 04:22 04:22 04:22 04:22

And trying to assert that there were optimizations that could have been made, which you're doing, is just not true to the nature of the process at the time. 04:22 04:22 04:22 04:22

Q. I'm asking you if it's technically possible to write source code where using a hash function, Meta could cross-reference hashes that it had already in a database of raw data it had collected and do exactly what Guillaume Lample's code does here? It says set piece priority to zero if you come across a hash that matches that we already have. 04:22 04:22 04:22 04:23 04:23 04:23

I'm just asking if it's possible. 04:23

MR. WEINSTEIN: Object to form. Calls for speculation. Calls for expert testimony. 04:23 04:23

THE WITNESS: There was no reason to do that. 04:23

BY MS. POUEYMIROU: 04:23

Q. That's not my question, though, about whether there was a reason. 04:23

I'm asking you if it is technically possible. 04:23

A. It's technically possible to write code that wires money to the wrong sources. We wouldn't do that either. 04:23

Q. Okay. 04:23

MR. WEINSTEIN: So, counsel, if we could take a break, I might be able to get you some clarity on an issue we talked about earlier. 04:23

MS. POUEYMIROU: Sure. 04:23

How much time do we have on the record? 04:23

THE VIDEOGRAPHER: We are at 4 hours and 27 minutes. 04:23

MS. POUEYMIROU: Okay. Great. Thank you. 04:23

THE VIDEOGRAPHER: Stand by. The time is 4:43 p.m., and we're going off the record. 04:23

(Recess taken.) 04:23

THE VIDEOGRAPHER: The time is 4:37 p.m., and we are back on the record. 04:36

BY MS. POUEYMIROU: 04:37

Q. Good afternoon, Mr. Beldock. 04:37

A. Hello again. 04:37

Q. How much total data did Guillaume Lample 04:37

versus unsafe version is entirely your characterization. And I couldn't disagree more that there should be a restrictive set of rearview mirror legalistic policies in a creative research organization. You just misunderstand the nature of early stage research.

BY MS. POUEYMIROU:

Q. Do you know whether Anthropic torrented and distributed copyrighted data from LibGen or Anna's Archive to build these large language models?

MR. WEINSTEIN: Object to form. Massively out of scope. Calls for speculation.

THE WITNESS: I have seen evidence that these datasets were used by at least two or three competitors. I don't know how they acquired them.

BY MS. POUEYMIROU:

Q. Are those two or three competitors Anthropic, OpenAI and someone else?

MR. WEINSTEIN: Same objections. Outside the scope. Calls for speculation.

THE WITNESS: I can't remember whether it was Anthropic specifically, but the set that you just enumerated plus Google and -- I think that's actually it, is roughly the same set.

BY MS. POUEYMIROU:

Q. Did you review a document today where Meta was

recognizing that OpenAI had put -- was refusing to use data appearing on the Notorious Markets List for training? 05:10

MR. WEINSTEIN: Object to form. Lacks foundation. Outside the scope. 05:10

THE WITNESS: There may have been a time at which they chose to make that announcement. And I do believe -- 05:10

BY MS. POUEYMIROU: 05:10

Q. Who is "they"? 05:10

A. -- they did -- OpenAI. 05:10

Q. And what was -- 05:10

A. And before that, all bets are off. 05:10

Q. Did you see that -- before that, all bets are off. 05:10

You don't actually know, though, whether OpenAI ever torrented and distributed copyrighted works from LibGen? 05:10

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 05:10

THE WITNESS: Our understanding is that they use these datasets. 05:10

BY MS. POUEYMIROU: 05:11

Q. I'm asking you whether you know -- whether your understanding is that they torrented and 05:11

distributed them like Meta? 05:11

MR. WEINSTEIN: Same objections. 05:11

THE WITNESS: Well, the only practical way to get them is to BitTorrent them. So by logical inference, I wouldn't be surprised if they did. 05:11

BY MS. POUEYMIROU: 05:11

Q. And so if they didn't, would that suggest to you that the sets were acquirable through direct downloading just as Bashlykov did? 05:11

MR. WEINSTEIN: Object to form. Lacks foundation. Calls for speculation. Outside the scope. 05:11

THE WITNESS: I don't know. 05:11

BY MS. POUEYMIROU: 05:11

Q. Is it accurate to say that if Meta had had a policy that required its engineers to track the data that Lample did in 2022 -- I understand that you think that's just absurd. But I'm asking you to hear my hypothetical and not debate with it. It is Meta has a policy and that policy says you track this data, plaintiffs in this case would have a lot more information about the upload rates of the torrent files their works were in? 05:11

MR. WEINSTEIN: Object to form. Calls for speculation. Outside the scope. 05:12

THE WITNESS: I just can't reconcile the 05:12

a little too long.

Can you do it, Annabel?

BY MS. POUEYMIROU:

Q. Okay. So I asked -- so you said, "There's a material difference between acquiring the works and acquiring the work via torrenting and intentionally redistributing them for profit."

"Right." You said, "Right." And then you said, "Certainly the latter with an emphasis on for profit."

A. Okay.

Q. In the supplemental rog response Meta stated -- on page 16 Meta states -- actually, let's turn to page 11.

A. Okay.

Q. The bottom of page 11, the final paragraph starts with "The ordinary market for plaintiffs' works is the market for people to purchase and read the books and possibly to create derivative works from those books such as for audio books and film and television adaptations."

Do you see that?

A. I do.

Q. Do you think that redistributing copyrighted books to third parties and piracy swarms harms this

ordinary market for plaintiffs' works? 06:20

MR. WEINSTEIN: Object to form. Outside the scope. 06:20

THE WITNESS: I disagree with your assertion of piracy swarms. 06:20

BY MS. POUEYMIROU: 06:20

Q. What do you mean by that? 06:20

A. You use the words "piracy swarms." And I disagree with your characterization. 06:20

Q. So how would you characterize a piracy swarm? 06:20

MR. WEINSTEIN: Outside the scope. 06:21

THE WITNESS: I've already defined piracy for you. 06:21

BY MS. POUEYMIROU: 06:21

Q. Okay. 06:21

A. A swarm is multiple. Piracy events or pirates. 06:21

Q. Okay. So when Meta states here in its verified interrogatory response that the ordinary market for plaintiffs' works is the market for people to purchase and read the books, my question to you is, when Meta distributed plaintiffs' works to third parties via torrenting, did it harm this ordinary market identified in this interrogatory response? 06:21

MR. WEINSTEIN: Object to form. Outside the 06:21

scope. Calls for a legal conclusion. 06:21

You can answer to the extent you can. 06:21

THE WITNESS: To the extent that there was any distribution, it was not intentional, and it certainly was not for profit. So, therefore, it does not qualify as piracy to my definition. 06:21 06:22 06:22

BY MS. POUEYMIROU: 06:22

Q. I'm asking for Meta's today. 06:22

MR. WEINSTEIN: And I'm going to object outside the scope. I don't think you are -- 06:22 06:22

BY MS. POUEYMIROU: 06:22

Q. So we can look at page 16, at the end of page 16, the end of that interrogatory response we've been looking at, Meta states, "Fair use also applies to this theory because there can be no harm to the market or potential market or value of plaintiffs' works because the near act of making their works available to third parties through BitTorrent without actual dissemination of their works does not result in any third party obtaining a copy of any of plaintiffs' works and, thus, can cause no market substitution or any other potential harm." 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22 06:22

Do you see that, Mr. Beldock? 06:22

A. I -- I do see that. 06:23

Q. Is it Meta's position today that where there 06:23

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? 06:23

MR. WEINSTEIN: Object to form. Outside the scope. Calls for legal conclusion. 06:23

THE WITNESS: No. 06:23

BY MS. POUEYMIROU: 06:23

Q. Why? 06:23

A. I'm not a lawyer. But my understanding of this document is that there's no substitution. 06:23

Q. Your understanding is that there's no substitution even if plaintiffs show that their works were distributed to third parties? 06:23

MR. WEINSTEIN: The same objections. Outside the scope. 06:24

THE WITNESS: I don't understand that that would be substitution. 06:24

BY MS. POUEYMIROU: 06:24

Q. Okay. So to date -- and, of course, this will unfold in the context of expert work. But based on our analysis, no one in the history of the United States, no single entity has distributed more copyrighted books than Meta did from 2022 to 2024. The amount of data distributed is comparable to the amount of data that 06:24

estimate it, and I am reasonably certain based on the fact that there are automated data transition notes in here from the summer of 2024. 06:49 06:50 06:50

Q. And how did the purpose change for the AIDC at that time? 06:50 06:50

A. Originally AIDC was a catalog of internal data sources and maintained metadata -- well, actually, no, indeed it was metadata about what kinds of information was contained and the origin of that information with respect to how it was sourced inside the company in order for the company to maintain its commitments or to be able to achieve its commitments on limiting data use of certain types and combining those data with other data of certain types. Which are the subject of regulations, either data use regulations or privacy regulations in many jurisdictions. 06:50 06:50 06:50 06:50 06:50 06:50 06:50 06:51 06:51 06:51 06:51

Sometime during 2024 as I understand -- as I recall, when we were scaling up our systematic approach to data lineage tracking for Llama model training, we transitioned other dataset review processes for external data into an expanded version of AIDC. Because AIDC had not exactly the right granularity, but could be modified as an enterprise system that the company relied on to include this additional functionality which allowed us to record metadata more 06:51 06:51 06:51 06:51 06:51 06:52 06:52 06:52 06:52

appropriate to external data sources.

And so the transition was from tracking internal datasets for compliance review to tracking all -- not all -- to tracking third-party datasets as well, is the way to put it, for, as I understand it, more than one purpose.

Q. Is there a difference between third-party and external data sources, or are you referring to that interchangeably here?

A. I -- I did use them interchangeably in my answer. There -- you've reminded me that -- that there is a difference between a dataset that we have acquired from the web or by download or by license. And I'm not certain the status of licensed datasets prior to this transition.

Q. Okay. And prior to the transition, was there a name for or a different catalog other than the AIDC for datasets acquired from external data sources?

A. There -- there wasn't a single system. I know of the existence of a system that was used by FAIR, which is the foundational or fundamental AI research lab. F stands for one of those two words.

And I also know of the existence of a spreadsheet that tracked a number of dataset approvals. I could not tell you whether those sets of datasets --

collections of datasets, just to clarify the wording, overlapped or intersected or were disjoint. 06:54

Q. Are you referring to the spreadsheet, as opposed to the dataset catalog that FAIR owned? 06:54

A. There was a dataset catalog that FAIR maintained. I don't know that they owned it. It was used for their purposes. 06:55

Q. Okay. 06:55

A. And then I am referring to the collection of datasets that FAIR was tracking in whatever their mechanism was which I don't know the identity of or the name, and the spreadsheet that had its own collection of datasets. And what I'm saying is, I don't know if those sets were disjoint, intersected or one was a subset of the other. 06:55

Q. When you say "I don't know," are you referring to you personally or Meta as the company? 06:55

MR. WEINSTEIN: Object to form. 06:55

THE WITNESS: I'm speaking of me. The company could answer that question. 06:55

BY MR. PRITT: 06:55

Q. Did you do anything to prepare to answer that question? 06:55

MR. WEINSTEIN: Object to outside the scope. 06:55

THE WITNESS: It didn't appear necessary. 06:55

BY MR. PRITT: 06:55

Q. Okay. Was the FAIR dataset catalog, was the information on the datasets in that catalog transferred or onboarded into AIDC? 06:55 06:56 06:56

A. I don't know. 06:56

Q. Okay. 06:56

A. I do know that the spreadsheet was. 06:56

Q. Okay. Do you know who used the spreadsheet? Which teams? 06:56 06:56

A. I -- I do. 06:56

Q. Okay. And which teams were they? 06:56

A. It was the dataset review XFN team. By XFN we mean cross-functional team. And that team consisted of legal, risk review or privacy review, and I believe other functions. 06:56 06:56 06:56 06:57

Q. Is that the only source in which Library Genesis was recorded before it appears in the AIDC? 06:57 06:57

MR. WEINSTEIN: Object to form. Outside the scope. Calls for speculation. 06:57 06:57

THE WITNESS: I can't be exhaustive. My understanding is that it was in the spreadsheet. But that doesn't directly answer your question. 06:57 06:57 06:57

BY MR. PRITT: 06:57

Q. Okay. Was Anna's Archive in the spreadsheet? 06:57

A. My understanding is that it was and that this 06:57

Q. Does "lots of text" refer to written works? 07:24

A. When you say "written works," isn't all text written? 07:24

Q. So is the answer, yes? 07:24

MR. WEINSTEIN: Object to form. 07:24

THE WITNESS: I think so. 07:24

BY MR. PRITT: 07:24

Q. Okay. Were all of the written works that the description refers to taken from Anna's Archive? 07:24

A. I -- 07:24

MR. WEINSTEIN: Object to form. Outside the scope. 07:24

THE WITNESS: I don't know. 07:24

BY MR. PRITT: 07:24

Q. Is that because the only way to actually know what the contents of this dataset named Anna's Archive are is to look at the underlying data contained elsewhere? 07:24

A. That's one of the reasons. 07:25

Q. What are the other reasons? 07:25

A. Well, you provided me with an overview page, and I don't have access to the rest of the system. 07:25

Q. Are you saying that there are other parts of the system in the AIDC which would tell you precisely what was in this dataset? 07:25

A. I'm not saying that. 07:25

Q. Okay. 07:25

A. I just can't rule it out. 07:25

Q. Okay. Fair enough. 07:25

What is the -- this has two usages listed, exploration and training. 07:25

A. Uh-huh. 07:25

Q. Are there any other -- are those approved usages? 07:25

A. That's a level of detail in the interworkings of the system I couldn't confirm without more research. I see that it matches the same fields on the next -- on the compliance page. And -- 07:25 07:26 07:26

Q. And I apologize. I wasn't asking -- I didn't mean to just interrupt you. But I wasn't asking specifically about this dataset. I was just asking in general, are exploration and training usages that are considered approved usages? 07:26 07:26 07:26 07:26

A. The approval would be for specific datasets for specific purposes. 07:26 07:26

Q. And what are the different purposes that a dataset can be approved for? 07:26 07:26

A. Well, certainly we know that members of that set include the phrase "exploration" and the phrase "training." I don't know what the rest of that set of 07:26 07:26 07:26

potential approved usages there are. But, again, we could find out by enumerating the field. Or the unique values of this field in AIDC. 07:27

Q. But sitting here today you are not prepared to tell me what all the different potential approved usages are for datasets in the AIDC? 07:27

MR. WEINSTEIN: Object to form. Scope. 07:27

THE WITNESS: I am not prepared to tell you that sitting here today. 07:27

BY MR. PRITT: 07:27

Q. You'll see there is a -- it says "model family." And next to it, it says "Llama 4." Some other AIDC sets might say Llama 3 or some variation of a model in a family. What does that signify? 07:27

A. It signifies that this -- this instance -- this is what I meant by rows when we were talking earlier. So if we think of Anna's Archive here as a row in the database, which, again, is a simplification, which was updated on what looks like January 8th, 2025, was when this row was updated being contemplated, I think is the right term, for use in the Llama 4 herd of models. 07:28

Q. Okay. And you'll see under sources the first field is hive filter, and it gives a path that says data hive, you know, data_mine and then crawl_bespoke. What 07:28

BY MR. PRITT: 07:36

Q. Okay. Does that spreadsheet have a name? 07:36

A. If it does, I don't know it. I have heard of it simply as the dataset review tracking spreadsheet. 07:36

Q. Okay. 07:36

A. But I don't know what its name actually is. 07:36

MR. PRITT: Will you let me know when we have ten minutes left? 07:36

THE VIDEOGRAPHER: We're at nine and a half minutes left. 07:36

MR. PRITT: Oh, can we take a break? That's five minutes so I can figure out what my last questions will be? 07:37

MR. WEINSTEIN: You're going to take a five-minute break? 07:37

MR. PRITT: Yes, I am. All right. 07:37

MR. WEINSTEIN: I'll let you go. I'm going to let this -- 07:37

THE VIDEOGRAPHER: Stand by. The time is 7:37 p.m., and we're going off the record. 07:37

(Off the record.) 07:37

THE VIDEOGRAPHER: The time is 7:43 p.m., and we are back on the record. 07:43

BY MR. PRITT: 07:43

Q. Mr. Beldock, the dataset identified in the 07:43

AIDC in the exhibit before you, is that dataset still approved today? 07:43

MR. WEINSTEIN: Which exhibit? 07:43

THE WITNESS: Which exhibit are you referring to? 07:44

BY MR. PRITT: 07:44

Q. Exhibit 15. 07:44

MR. WEINSTEIN: Thank you. 07:44

THE WITNESS: My understanding is that the system -- 07:44

MR. WEINSTEIN: Outside the scope. 07:44

You can answer. 07:44

THE WITNESS: -- still shows this as approved for exploration and training. 07:44

BY MR. PRITT: 07:44

Q. Is it, in fact, still approved for exploration and training today? 07:44

MR. WEINSTEIN: Object to form. Outside scope. 07:44

THE WITNESS: 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. 07:44

BY MR. PRITT: 07:44

Q. I wasn't trying to ask a hypothetical. I'm 07:44

just trying to ask is this dataset, whether it's been replaced or not, still approved for either exploration or training today? 07:44–07:45

MR. WEINSTEIN: Scope. 07:45

THE WITNESS: We wouldn't use it for the reasons I've explained in my prior testimony. 07:45

BY MR. PRITT: 07:45

Q. Okay. There are many datasets in the AIDC that state unknown in the usage field. 07:45

A. Okay. 07:45

Q. If a dataset in the AIDC states unknown in the usage field, does Meta know what the usage of that dataset is? 07:45

MR. WEINSTEIN: Object to form. Lacks foundation. 07:45

THE WITNESS: I think your question misapprehends the actual use of this field. 07:45

BY MR. PRITT: 07:45

Q. What is the actual use of the usage field? 07:45

A. The approved usage field contains information about what is approved in the instance of the review in question. For this dataset. It -- it doesn't track the -- the actual usage. 07:46

Q. And what does track the actual usage of each dataset that is reflected in the AIDC? 07:46

A. I don't know that there is a single system that does that. 07:46

Q. Okay. So how would Meta identify what the actual usage is for each data that is -- dataset that is reflected in the AIDC? 07:46

MR. WEINSTEIN: Object to form. Scope. 07:46

THE WITNESS: It would be -- it would have to search through the records of all of the training that was done with the dataset in question if that's possible. 07:46

BY MR. PRITT: 07:47

Q. Is it possible that a dataset reflected in the AIDC would not be contained in any records of all of the training that was done? 07:47

MR. WEINSTEIN: Object to form. 07:47

THE WITNESS: Excuse me. Could you repeat that? I didn't quite follow you. 07:47

BY MR. PRITT: 07:47

Q. Yeah, I was trying to use your words. My question was, is it possible that a dataset reflected in the AIDC would not be contained in any records of all of the training that was done? 07:47

A. Oh. 07:47

MR. WEINSTEIN: Object to form. Foundation. 07:47

THE WITNESS: There -- I -- I believe there 07:47

are datasets that are approved but not used. So, yes. 07:47

BY MR. PRITT: 07:47

Q. This is Beldock Exhibit 21. 07:48

(Exhibit 21 was marked for identification.) 07:48

BY MR. PRITT: 07:48

Q. And I just want to point you -- sorry, it's Meta Kadrey 00244487. The dataset name is LibGen. 07:48

A. Is LibGen? 07:48

Q. Is LibGen. 07:48

A. I see. Okay. 07:48

Q. I just want you to look at a field on the page ending in 488. 07:48

A. 488. 07:48

Q. 488. 07:48

A. Okay. So that's -- excuse me, that's -- that's the second page. All right. I'm with you. 07:48

Q. Yeah. The second kind of main field says "Download." Then it says, "Submit a request," appears to be a link, "for the data team to consider downloading this dataset for you." Do you see that? 07:48

A. I do see that. 07:48

Q. Does -- is this referring to re-downloading the dataset or just getting a copy of the dataset that already exists? 07:48

A. That's a good question. I interpret -- 07:49

I don't know for sure. I can tell you what this says. 07:49
And I can interpret it to mean that it was, in fact, a 07:49
request to download in the first place. 07:49

Q. And does Meta keep a record of the requests that are submitted through this process in the AIDC to request from the data team to consider downloading the dataset? 07:49

MR. WEINSTEIN: Object to form. 07:49

THE WITNESS: Do we keep a record -- there -- there likely would be a record. Is it one that we formally keep, I do not know. 07:49

BY MR. PRITT: 07:49

Q. Okay. 07:49

A. Is it likely that a form that is submitted would have some record somewhere, yes. 07:49

Q. Okay. And on the very first page of this, you'll see towards the bottom there is an airstore field, and then it lists what is actually SciTech in the name. Does this signify that this dataset, you know, with the name SciTech underscore, et cetera, et cetera, was used or intended to be used in an in-production model? 07:49 07:50 07:50 07:50 07:50

A. The existence of an airstore URL does not map one-to-one to production models. You can't infer one from the other. 07:50 07:50

I have read the foregoing deposition transcript and by signing hereafter, subject to any changes I have made, approve same.

Dated: _____

_____

JAMES BELDOCK, 30(b)(6)

REPORTER'S CERTIFICATE

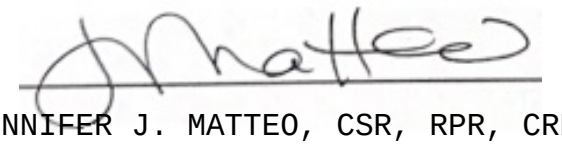
I, JENNIFER J. MATTEO, hereby certify that the witness in the foregoing deposition was by me sworn to tell the truth, the whole truth, and nothing but the truth in the within-entitled cause;

That said deposition was taken in shorthand by me, a disinterested person, at the time and place therein stated, and that the testimony of the said witness was thereafter reduced to typewriting, by computer, under my direction and supervision;

That before completion of the deposition, review of the transcript [ ] was [X] was not requested. If requested, any changes made by the deponent (and provided to the reporter) during the period allowed are appended hereto.

I further certify that I am not of counsel or attorney for either or any of the parties to the said deposition, nor in any way interested in the event of this cause, and that I am not related to any of the parties thereto.

Dated: September 14, 2026



JENNIFER J. MATTEO, CSR, RPR, CRR
CSR License No. 12139