EXHIBIT 116

(FILED UNDER SEAL)

1

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN JOSE DIVISION

CONCORD MUSIC GROUP, INC., ET AL., Plaintiffs, vs. ANTHROPIC PBC, Defendant. Case No. 24-cv-03811-EKL-SVK

***

HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY

***

REPORTER'S TRANSCRIPT
VIDEOTAPED DEPOSITION OF
DAWN DRAIN
Wednesday, November 12, 2025

Stenographically Reported by:
Shannon Borge, RPR,
CSR No. 9436

Job No. 606064
Pages 1-231

2

APPEARANCES

---oOo---

ON BEHALF OF THE PLAINTIFFS:

OPPENHEIM + ZEBRAK, LLP
BY: BRET MATERA, ESQUIRE
COREY MILLER, ESQUIRE
461 5th Avenue
19th Floor
New York, New York 10017
212.951.1833

ON BEHALF OF THE DEFENDANT AND THE WITNESS:

WILMER CUTLER PICKERING HALE AND DORR LLP
By: TAYLOR GOOCH, ESQUIRE
ALYSSA MILSTEAD, ESQUIRE (Via Zoom)
50 California Street
Suite 3600
San Francisco, California 94111
628.235.1026

ALSO PRESENT:

Gary Garcia, Planet Depos Videographer

Ian Chen, Anthropic Counsel

---oOo---

3

DEPOSITION SUPPORT INDEX

QUESTIONS INSTRUCTED NOT TO ANSWER:

PAGE	LINE
211	8

REQUEST FOR PRODUCTION OF DOCUMENTS

PAGE	LINE
16	22
191	22

STIPULATIONS

PAGE	LINE
(None)	

QUESTIONS MARKED

PAGE	LINE
(None)	

REPORTER'S NOTE:
QUOTATION MARKS ARE USED FOR CLARITY AND DO NOT NECESSARILY REFLECT A DIRECT QUOTE.

4

INDEX

---oOo---

	PAGE
EXAMINATION BY MR. MATERA	7
EXAMINATION BY MR. GOOCH	228

---oOo---

EXHIBITS

EXHIBIT	DESCRIPTION	PAGE
Exhibit 135	LinkedIn profile	19
Exhibit 136	Blog post (Anthropic_0000013558)	58
Exhibit 137	"Studying Data Repetition Through Scaling Laws and Mechanistic Interpretability (Anthropic_0000520123)	59
Exhibit 138	11/09/23 PowerPoint (Anthropic_0000500599)	85
Exhibit 139	"Exact Dedupe for [redacted]" (Anthropic_0000510628)	131
Exhibit 140	08/15/23 "Deduplication Algorithms" (Anthropic_0000507750)	146
Exhibit 141	"Dawn's Thoughts On Data Priorities" (Anthropic_0000216875)	165
Exhibit 142	"pranay dawn 1:1" (Anthropic_00052912)	195
Exhibit 143	Slack thread (Anthropic_0000024220)	197
Exhibit 144	Memorization Prompt (Anthropic_0000242229)	199

5

E X H I B I T S (Cont'd)

EXHIBIT	DESCRIPTION	PAGE
Exhibit 145	Garbage Prompt (Anthropic_0000242230)	202
Exhibit 146	Slack thread (Anthropic_0000016426)	204
Exhibit 147	Slack thread (Anthropic_0000352478)	218
Exhibit 148	"The Fall of Brome" Amazon listing	220
Exhibit 149	"The Fall of Brome" Excerpt (Digital format)	225

---oOo---

6

BE IT REMEMBERED that, pursuant to Notice of Taking Deposition, and on Wednesday, November 12, 2025, commencing at 9:00 a.m. PDT and ending at 5:45 p.m. PDT thereof, at 50 California Street, San Francisco, California, before me, Shannon Borge, a Certified Shorthand Reporter, Registered Professional Reporter, in and for the State of California, personally appeared,

DAWN DRAIN,

a witness called on behalf of Plaintiffs, pursuant to all applicable sections of the Code of Civil Procedure of the State of California, and who, having been first duly sworn by me to testify to the truth, was examined and testified as follows:

THE VIDEOGRAPHER: Here begins Media No. 1 in the videotaped deposition of Dawn Drain in the matter of Concord Music Group versus Anthropic in the United States District Court, Northern District of California, San Jose Division, Case Number 5:24-cv-03811-EKL-SVK.

Today's date is November 12th, 2025.

The time on the video monitor is 9:00 a.m.

The videographer today is Gary Garcia representing Planet Depos.

This video deposition is taking place at 50 California Street, San Francisco, California.

Would counsel please voice identify themselves

7

and state whom they represent starting with the taking attorney.

MR. MATERA: Bret Matera from Oppenheim & Zebrak LLP on behalf of the music publisher plaintiffs and joined by my colleague Corey Miller.

MR. GOOCH: Taylor Gooch with -- from Wilmer Hale. Also with me here is Ms. Alyssa Milstead representing Anthropic. In the room with us are in-house counsel for Anthropic as well as the witness.

THE VIDEOGRAPHER: The court reporter today is Shannon Borge, representing Planet Depos. The witness will now be sworn.

THE COURT REPORTER: Please raise your right hand.

(Whereupon, the witness was sworn.)

THE COURT REPORTER: And for the record, my name is Shannon Borge, California Shorthand Reporter, License Number 9436.

DAWN DRAIN,

having been first duly sworn,

testified as follows:

EXAMINATION

BY MR. MATERA:

Q. Okay. Good morning, Miss Drain.

A. Good morning.

8

Q. My name is Bret Matera and I'll be taking your deposition today. As I stated earlier, I represent the plaintiff music publishers in this case. Thank you for taking time out of your schedule to answer my questions today. We appreciate it.

To begin, can you please start by stating your full name and address?

A. Yeah. My name is Dawn Drain. My address is [redacted]

Q. Thank you. Miss Drain, have you been deposed before?

A. Yeah. I had a brief deposition, which was like 30 minutes, when I was in high school related to a car accident.

Q. And that's the only time you've been deposed before?

A. Yeah.

Q. You've never testified at trial before?

A. No.

Q. Okay. Before we get started, I would like to just go over a few ground rules today so that we can go quickly and smoothly. I'm going to ask you a series of questions that I would like you to answer truthfully and accurately. If there's anything about any of my questions today that you do not understand, please let

65

A. Um, a FLOP is a floating point operation, like an add or a multiply.
Q. And so when a model memorizes something, that memorization has a tendency to impact the model's compute; is that correct?
A. Um, so the way that -- that I was trying to -- to frame things is, um, if -- if the model memorizes some sub-distribution of the text, according to the previous technical definition, then, um, the model's overall performance will be degraded as if you had used less compute to train it.
Q. And as a user of the model, how is that degradation you just described seen?
A. Um, so we have -- um, we have -- Anthropic has three different offerings of model sizes that vary by how much compute has gone into them, and if a model is made dumber, then it will be as if it was performing more similarly -- similarly to a model from a smaller class.
Q. And you mentioned there's a compute budget; is that correct?
A. Um, when I -- when I use the phrased "compute budget," I was saying the definition of the compute efficient frontier for -- for -- if you were given a fixed compute budget, there's a -- an optimal loss you

66

can achieve given -- like given a codebase or a training codebase.
Q. And so in your role, your objective is to find the optimal loss; is that correct?
A. I -- think it's more that we're trying to shift the compute efficient frontier to make training more efficient so you can achieve a better loss with the same amount of compute.
Q. I see. So, broadly speaking, in your work you want to reduce memorization; is that correct?
A. Um, that isn't something that I'm working on now, but, um, especially in 2023, I was looking into ways to, um, reduce memorization because memorization generally makes the models dumber and we want to reduce it.
Q. You said you worked on that in 2023?
A. Yeah.
Q. And did you succeed in your work?
A. Um, yeah, we -- we developed techniques to reduce duplication and, um, make it so that there was less text that the, um, models were memorizing so that they were, therefore, not being harmed as much or made dumber by memorizing that text.
Q. I'd like to come back to the techniques you developed to reduce duplication, but in the meantime, I

67

do want to ask does Anthropic decide what its AI models memorize?
MR. GOOCH: Objection. Calls for speculation.
THE WITNESS: Um, um, yeah. So -- so, in practice, um, based on the training data, um, if some piece of text is repeated many times and the model memorizes it, which is harmful to the model's performance, it makes the model dumber. Um, so by improving deduplication, um, we can reduce the number of things that are memorized.
BY MR. MATERA:
Q. Conversely, if you were to add repeated data, would you be able to increase the probability that a model memorizes that repeated data?
A. Yeah, you -- you wouldn't want to do --
MR. GOOCH: Sorry. Just give me --
THE WITNESS: Sorry.
MR. GOOCH: -- five seconds to object. No worries.
Same objection and also, um, asked and answered.
Go ahead.
THE WITNESS: Um, yeah, we wouldn't want to do that. Um, you could train a model on a heavily repeated dataset and have it memorize the text in that dataset,

68

but you wouldn't want to do that because it would -- it would make the model dumber.
BY MR. MATERA:
Q. Is there anything you would want the model to memorize?
A. Um --
MR. GOOCH: Same objections.
THE WITNESS: Um, um, yeah. So there's this, um, definition of memorization where the model has this pathologically high probability of generating some piece of text that it's been trained to memorize in this superficial way that, um, makes the model dumber, um, and then there's a more colloquial sense in which we want the model to -- um, there's a more colloquial sense or like another -- another definition where the model's simply capable of generating a sequence. Um, and, yeah, we want -- yeah. I think, um, it's good for the model to have memorized things like famous quotes that like you or I might memorize.
BY MR. MATERA:
Q. So are you saying there's a incidental memorization that comes with training and then a more deliberate memorization that you described as colloquial?
MR. GOOCH: Objection. Misstates prior

81

MR. GOOCH: Objection. Misstates prior testimony.
THE WITNESS: Um, yeah. So I -- I don't know -- I don't know that Claude has memorized the Gettysburg address beyond at least the like "Four scores and seven years ago" part, um, but one -- one definition of memorization you can use is if top -- top-one sampling could reproduce that same piece of text at like maximum likelihood sampling.
BY MR. MATERA:
Q. What's "top-one sampling"?
A. Um, so for each token, um, that the model is generating, it has a probability distribution over all of the possible tokens in the vocabulary --
(Court reporter requested clarification.)
THE WITNESS: -- probability distribution over all the tokens in the vocabulary, um, and normally, um, you sample directly from the probability distribution, but, instead, you could just look at the token with the highest probability and use that one and that would be top-one sampling or greedy sampling or greedy decoding.
BY MR. MATERA:
Q. "Greedy sampling," is that what you said?
A. Yeah.
Q. That's the same as top-one sampling?

82

A. Yeah.
Q. How are the probability distributions of tokens determined?
MR. GOOCH: Objection. Vague.
THE WITNESS: Um, yeah. There's a whole -- whole training process. Um, you -- you have pretraining and then fine-tuning and then you can do a forward pass through the model and then that will give you the probabilities.
BY MR. MATERA:
Q. Are you able to set the probability distribution for specific excerpts of data?
A. Um, not really. Not -- yeah. It's a more abstracted process than that.
Q. In your experience, does Anthropic want its AI models to be able to recite the Gettysburg address?
MR. GOOCH: Objection. Calls for speculation.
THE WITNESS: Um, yeah, I -- I can't really talk on the behalf of Anthropic [indiscernible utterance].
BY MR. MATERA:
Q. I'm asking for your experience at Anthropic. Is that generally something that they would want the AI models to be able to rec- -- recite?
MR. GOOCH: Same objection. Calls for

83

speculation.
THE WITNESS: Um, yeah, I don't -- I don't know about Anthropic in general.
BY MR. MATERA:
Q. But you personally believe that Anthropic's AI models should be able to recite the Gettysburg address?
A. Um, maybe not like the whole -- whole thing but -- and maybe like not 100 percent perfectly but -- but yeah.
Q. Does Anthropic -- strike that.
In your -- in your experience, does Anthropic do anything to ensure that its AI models can recite the Gettysburg address?
A. Um, in -- in my experience, we -- um, yeah, we don't take any steps to -- um, to make it -- to make it memorize. Um, in general, we try and reduce memorization because it's harmful to the model and makes it dumber.
Q. So to be clear, you're not aware of anything that Anthropic does to ensure that its models memorize the Gettysburg address?
A. I mean, I don't know that it's a memorizable thing. I haven't -- I haven't tested that, um, but, yeah, I don't -- I don't know of any efforts to try and get the model to memorize the Gettysburg address. I'm

84

generally trying to reduce memorization because it makes the model dumber.
Q. Do you recall earlier testifying about works in the public domain?
A. Um, I -- I remember saying that I thought the Constitution and the Gettysburg address were in the public domain, but I wasn't an expert on it, yeah.
Q. Do you think AI models should memorize certain works in the public domain?
MR. GOOCH: Objection. Asked and answered maybe five times at this point.
THE WITNESS: I've -- I've already answered this.
BY MR. MATERA:
Q. So you're not going to answer my question?
A. I've already answered it.
Q. I don't believe you have.
MR. GOOCH: Same objection.
THE WITNESS: It's -- I mean, you keep asking the same question.
BY MR. MATERA:
Q. It's not the same question. Are you going to answer my question, yes or no?
A. I think it's the same as the previous question so...

153

lot more in improving our deduplication algorithms, this was a bigger problem.

Q. I'm sorry. You said this is looking at a couple definitions of what?

A. Of memorization.

Q. What are those definitions?

A. Um, yeah. So that -- um, the -- the main two definitions here are, um, if you have, um, a single token with an extremely low loss or if you have --

(Court reporter requested clarification.)

THE WITNESS: An extremely low loss.

Or if you have a large span of tokens, um, with a pathologically low loss. Yeah, those are the two main definitions here.

BY MR. MATERA:

Q. Are there other definitions of memorization?

A. Um, yeah, we -- we talked earlier about the greedy sampling definition.

Q. Any others?

A. Um, um, um, and you can define -- you could set the -- the threshold for the loss to be based on, um, the typical loss for the model. You could renormalize it that way. I've used that definition before.

Q. Any others?

A. Um, um, yeah. It's pretty much just those.

154

Yeah. I mean, there are probably some others in the literature, but I can't -- I can't think of any.

Q. None other that you're aware of at this moment?

A. Yeah.

[redacted]

Q. "cc.p8" is a Common Crawl band; is that correct?

A. Yeah.

[redacted]

155

[redacted]

MR. GOOCH: Objection. Vague.

[redacted]

BY MR. MATERA:

[redacted]

A. So, yeah, I would say that a sixth would be a large fraction.

[redacted]

Do you see that?

A. Yeah.

Q. Can you explain that in more layman's terms?

A. Um, yeah. Um, [redacted]

156

[redacted]

BY MR. MATERA:

[redacted]

177

A. Yep.

[redacted]

MR. GOOCH: Objection. Calls for speculation.

[redacted]

178

[redacted]

(Court reporter requested clarification.)

[redacted]

BY MR. MATERA:

[redacted]

MR. GOOCH: Objection. Asked and answered.

[redacted]

(Court reporter requested clarification.)

[redacted]

179

[redacted]

BY MR. MATERA:

Q. Can you turn to the next, please. I believe it ends in 882.

A. Okay.

[redacted]

180

[redacted]

229

all of which have pathologically low loss, and we do a forward pass through the model.

Q. And, now, in your time at Anthropic with your work on various experimental models or production models, how do you view memorization?

(Court reporter requested clarification.)

BY MR. GOOCH:

Q. Now, at your time at Anthropic and your work there with various models, whether experimental production, how do you view memorization?

MR. MATERA: Object to form.

THE WITNESS: Um, yeah, in general, memorization is harmful. It's something that we try and reduce using deduplication techniques. Um, memorization is a form of overfitting where the model can get very low train loss, but this actually makes the test loss worse. It makes the model dumber and -- at test time and worse at generalization.

MR. GOOCH: Thank you. That's all my questions.

MR. MATERA: Miss Drain, thank you for being here today.

Also, thank you to Gary and Shannon for helping us.

No redirect or no recross.

230

THE VIDEOGRAPHER: This marks the end of the deposition of Dawn Drain. We are going off the record at 5:45 p.m.

(Whereupon, proceedings were adjourned.)

---oOo---

231

CERTIFICATE OF CERTIFIED SHORTHAND REPORTER

STATE OF CALIFORNIA)
) ss.
COUNTY OF SACRAMENTO)

I, SHANNON BORGE, RPR, CSR No. 9436, Certified Shorthand Reporter, do hereby certify:

That prior to being examined, the witness in the foregoing proceedings was by me duly sworn to testify to the truth, the whole truth, and nothing but the truth;

That said proceedings were taken by me in shorthand and thereafter transcribed into typewriting under my direction and supervision;

That I am neither counsel for, nor related to, any party to said proceedings, nor in any way interested in the outcome thereof.

I further certify that I am not a party to any stipulation, if made, that would waive my duties mandated by Court Reporters Board of California.

In witness whereof, I have hereunto subscribed my name.

Dated: Monday, November 17, 2025

Shannon Borge

Shannon Borge, RPR, CSR
Certificate No. 9436

ERRATA TO THE DEPOSITION TRANSCRIPT OF
DAWN DRAIN
TAKEN ON **November 12, 2025**

I have reviewed the transcript of my testimony of **November 12, 2025**, and hereby make the following changes in form and/or substance:

PAGE	LINE(S)	CHANGE	REASON FOR CHANGE
13	14	"privileged" to "privilege"	Transcription error
27	25	"love" to "left"	Transcription error
50	4	"data BPI" to "GitHub API"	Transcription error
55	7	"constant to find in the codebase" to "defined in the codebase"	Transcription error
130	5-6	"common crawl" to "Common Crawl"	Capitalization error
168	10	"Reenforcement" to "Reinforcement"	Spelling error
223	23	"ISPN" to "ISBN"	Transcription error

I, **DAWN DRAIN**, read the foregoing deposition and hereby state that the foregoing is true and correct, except as noted herein.

Dawn Drain
Dawn Drain

12/17/2025
Date