Exhibit 11



HIGHLY CONFIDENTIAL
ATTORNEYS' EYES ONLY

Transcript of Dawn Drain

Date: November 12, 2025
Case: Concord Music Group -v- Anthropic

Planet Depos
Phone: 888.433.3767
Email: transcripts@planetdepos.com
www.planetdepos.com

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

A P P E A R A N C E S

---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

I N D E X

---oOo---

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

---oOo---

E X H I B I T S

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 Teal" (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

Pages 65-72

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

69

testimony and vague.
THE WITNESS: No, I wouldn't -- I wouldn't say that. I wouldn't say that there's intentional memorization, no.
BY MR. MATERA:
Q. Fair enough. I'm just trying to nail down these two types of --
A. Yeah.
Q. -- memorizations you mentioned. There's one -- if I'm understanding you correctly, you're saying there's one memorization that occurs through training and then there's another kind of memorization. Is that correct?
MR. GOOCH: Same objections.
THE WITNESS: Um, um, yeah, I'm saying that there's -- um, um, there's pathological memorization which can happen during training where the model overfits to a piece of text --
(Court reporter requested clarification.)
THE WITNESS: -- overfits to a piece of text by memorizing it in a superficial damaging way, um, and then separately there's a, um, capability where the model could, um, generate a -- a known piece of text like a famous quote or like a -- a famous excerpt from the Bible.

70

BY MR. MATERA:
Q. That capability you're describing where a model could generate a known piece of text like a famous quote or like a famous excerpt from the Bible, how does that capability work?
MR. GOOCH: Same objections.
THE WITNESS: Um, um, I mean, mechanically, the model, um, will -- um, is capable of, um, generating that piece of text, um...
BY MR. MATERA:
Q. Does it do that on its own or do you tell the model what to memorize specifically?
A. Um, so -- so, again, in general, we're trying to reduce memorization, um, because it's harmful for the model, um, because it makes the model dumber and it works at generalizing, um. Yeah. Sorry. I don't know of any -- any text that we've intentionally had the model memorize.
Q. I understand. I'm -- I'm asking more about the process. You gave some specific examples of "famous quotes and excerpts from the Bible," and I'm wondering how the model learns to memorize those specific examples that you listed?
A. Um, yeah, it --
MR. GOOCH: And same objection. Calls for

71

speculation and vague.
THE WITNESS: Um, um, yeah, so I don't -- I don't know that the model is -- is memorizing those things, um, yeah. But, in general, if a model is trained on a piece of text many times, then it will memorize the text, which in a typical case makes the model -- model dumber. It's a form of overfitting.
BY MR. MATERA:
Q. I understand that part. I'm talking about the specific example you gave of a model being capable of reciting, for example, a -- a bible excerpt. And I want to know how it does that and if it does that in a way that's different than the other memorization that you have discussed?
A. Um --
MR. GOOCH: Objection. Vague and asked and answered.
THE WITNESS: Yeah, I think -- I think I've answered this question.
BY MR. MATERA:
Q. I don't think you have. I'm asking specifically about this second form of memorization that you've been describing, and I just want to know how that works, who decides in a hypothetical situation what to specifically memorize?

72

MR. GOOCH: Objection. Vague. Compound. Calls for a hypothetical and speculation.
THE WITNESS: Could you ask a -- a narrower question?
BY MR. MATERA:
Q. If you wanted to train an AI model to memorize a specific bible excerpt, how would you do that?
MR. GOOCH: Objection. Calls for speculation and an expert opinion.
THE WITNESS: Um, again, in general, um, you could use, um, either pretraining or supervised learning or reenforcement learning.
(Court reporter requested clarification.)
THE WITNESS: Supervised learning.
But I don't -- I don't know of a way to do that which wouldn't, um, be damaging to the model. It's -- it's a hard problem because generally if you, um, just train on a piece of text, then it will result in this very superficial form of memorization, which is harmful to the model.
BY MR. MATERA:
Q. Earlier you said, quote, "I think it's good for the model to have memorized things like famous quotes that you or I might memorize." Do you recall that testimony?

Pages 89-92

89

A. Um, um, yeah, it was -- it was a couple years ago and there were a lot of remote people so I don't remember that any of the founders attended. Um, yeah. Yeah, I don't remember.

Q. People tuned in remotely is that what you said?

A. Yeah. There -- it was in person and also people tuned in remotely.

Q. Okay. What is the alignment science team?

A. Um, yeah, it's a team that Ethan and Ansh are on and they're -- yeah, they're mostly working on making the models more aligned and then also briefly they were helping out with this deduplication project.

Q. Okay. Did any of the people listed here help you create this PowerPoint?

A. Um, yeah, I mean, I used some graphs from Devi's data, some graphs that I made with Ethan, but, yeah, I'm not one who -- who wrote the PowerPoint -- put together the PowerPoint.

Q. If you could flip to the second page there. It ends with the Bates -- it ends in 600 for the Bates stamp.

A. Okay.

Q. Do you see that? And it just says, "Memorization is Terrifying." Did you write that?

A. Um, yeah, I wrote that.

90

Q. What'd you mean by that?

A. Um, so this is referring to, um, this thing I discovered or noticed where our models were -- um, were memorizing in this harmful superficial way, um, large amounts of text which suggested that our datasets weren't as deduplicated as we wanted. Um, so this was saying that, um, if we didn't do a better job to deduplicate our datasets, we risked the models memorizing and overfitting to -- to sub-distributions of the dataset, which -- which, yeah, would have made them a lot dumber, which was, yeah, kinda scary.

Q. Scary because why?

A. Um, well, you know, I mean, we -- were were afraid of the -- the models becoming dumber but -- I don't know -- one thing that was like especially scary about it was that it seemed that the effect was getting worse with scale, that the -- the bigger models I looked at --

(Court reporter requested clarification.)

THE WITNESS: The effect was getting worse at scale.

It seemed like the bigger models were memorizing more, which was, um, yeah, a risk of making them significantly dumber.

BY MR. MATERA:

91

Q. On page 6 ending in the Bates No. 604, do you see where it says, "Memorization looks *very different* from natural scaling"?

A. Um, yeah, I see that.

Q. What did you mean here?

A. Um, so generally larger models are -- are smarter and will get lower loss on any given piece of text, um, but, um, if you memorize a span of text, you get this pathologically low loss on it, which deviates from that trend, um, because the loss just collapses to zero because the model is memorizing it in a superficial way.

Q. And on page 13, which ends in Bates stamp 611, do you see where it says, "How bad is memorization/overfitting/duplication"?

A. Um, I see that.

Q. And below there, it says, "tl; dr very bad." Do you see that?

A. Yeah.

Q. And "tl; dr," that's "too long; didn't read"? Is that correct?

A. Yeah.

Q. Why -- what are you trying to say here?

A. Like I've said many times, if your dataset is duplicated, um, because -- because you did a bad job de-

92

-- deduplicating it, then the model can memorize and overfit to a sub-distribution of your dataset which was duplicated, which makes the model a lot dumber, which is what we strive to avoid.

Q. And you were just reiterating that on this slide?

A. Um, well, yeah, I mean, as in reiterating it like today, but this seems like the first time I say it -- said it within the presentation itself.

Q. I see.

A. I guess I said it was terrifying earlier. So, yeah -- yeah, I'm reiterating it within the presentation, yeah.

Q. Okay. If you could show me on page 26, which ends in 624, see where it says, "Tradeoffs -- Losing data is bad, mkay"?

A. I see that.

Q. Is that a South Park reference?

A. Yeah.

Q. Do you like South Park?

A. Yeah.

Q. Why do you say, "Tradeoffs," here?

A. Um, so with a perfect deduplication, we would only remove documents that are generally de -- deduplicated, but you can have false positives or