# EXHIBIT 135

Planet Depos
We Make It Happen™

**HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY**

# Transcript of Pranay Sangani

**Date:** November 10, 2025
**Case:** Concord Music Group -v- Anthropic

**Planet Depos**
**Phone:** 888.433.3767 | **Email:** transcripts@planetdepos.com
**www.planetdepos.com**
**Michigan #8598 | Nevada #089F | New Mexico #566**

1

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN JOSE DIVISION

CONCORD MUSIC GROUP, INC., ET AL.,
Plaintiffs,
v.
ANTHROPIC PBC,
Defendant.

Case No.: 5:24-CV-03811-EKL-SVK

VIDEOTAPED DEPOSITION OF PRANAY SANGANI

HIGHLY CONFIDENTIAL: ATTORNEYS' EYES ONLY

San Francisco, California

November 10, 2025

9:14 a.m.

REPORTED BY:

Tammy Moon, CSR No. 13184, RDR, CRR

2

APPEARANCES:

FOR PLAINTIFFS CONCORD MUSIC GROUP, INC., ET AL.:

OPPENHEIM & ZEBRAK, LLP
BY: COREY MILLER, ESQ.
BY: MICHELLE GOMEZ-REICHMAN, ESQ.
4530 Wisconsin Ave. NW 5th Flr
Washington, DC 20016
202.621.9027
corey@oandzlaw.com

FOR DEFENDANT ANTHROPIC PBC:

WILMIER CUTLER PICKERING HALE and DORR LLP
BY: ARI HOLTZBLATT, ESQ.
BY: ALYSSA MILSTEAD, ESQ.
2100 Pennsylvania Ave. NW
Washington, DC 20037
202.663.6964
ari.holtzblatt@wilmerhale.com

ALSO PRESENT: LIZ KIM, ANTHROPIC

NATALIE NAUGLE, ANTHROPIC

PHILIP ASTOR, THE VIDEOGRAPHER

3

INDEX TO EXAMINATION

PRANAY SANGANI

Monday, November 10, 2025

Tammy Moon CSR No. 13184, RPR, CRR

WITNESS: PRANAY SANGANI


| EXAMINATION | PAGE |
|---|---|
| BY MR. MILLER | 7 |
| BY MR. HOLTZBLATT | 279 |


4

INDEX TO EXHIBITS

PRANAY SANGANI

Monday, November 10, 2025

Tammy Moon CSR No. 13184, RPR, CRR


| MARKED | DESCRIPTION | PAGE |
|---|---|---|
| Exhibit 26 | Bates-stamped pages ANTHROPIC217620-ANTHROPIC217623 | 168 |
| Exhibit 121 | Bates-stamped page ANTHROPIC217164, Excel sheet | 93 |
| Exhibit 122 | Bates-stamped page ANTHROPIC218116, Excel spreadsheet | 100 |
| Exhibit 123 | Bates-stamped pages ANTHROPIC234139-ANTHROPIC234141 | 120 |
| Exhibit 124 | Bates-stamped pages ANTHROPIC12590-ANTHROPIC12590 | 145 |
| Exhibit 125 | Bates-stamped pages ANTHROPIC237281-ANTHROPIC237293 | 155 |
| Exhibit 126 | Bates-stamped pages ANTHROPIC2127620-ANTHROPIC2127620 | 179 |
| Exhibit 127 | Bates-stamped pages ANTHROPIC342979-ANTHROPIC342983 | 186 |
| Exhibit 128 | Bates-stamped pages ANTHROPIC346083-ANTHROPIC346087 | 202 |

# Pages 37-40

37

**Claude.**
Q. Okay.
**A. But, yes, there exists research models and commercially available models at Anthropic.**
Q. What does Anthropic use the noncommercial research models for?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: I think that's a pretty broad question. Can you clarify that a little bit more.
MR. MILLER:
Q. Sure. Why does Anthropic create models that it doesn't intend to commercially release?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: There are many reasons to train research models. Sometimes we train small models that there would be no reason to commercially release to understand a part about how models train in general and the dynamics by which certain relationships emerge in models. So generally, maybe we can bucket that into the category of, like, we train models to learn. There are probably other key reasons to train research models, but I think -- would love for you to clarify if -- if this is a good -- is a good branch of thought.
///

38

MR. MILLER:
Q. Mr. Sangani, do you use Claude?
**A. I do.**
Q. Did you use Claude before you joined Anthropic?
**A. At the time, I don't believe there was a commercially viable -- I don't believe we had launched Claude as a product.**
Q. As a Claude user, what types of things do you use Claude to do?
**A. I use Claude for everything from helping with software engineering/coding type of tasks; to understanding research papers; to maybe, like, broadly trying to understand things I'm less familiar with; to trying to understand aspects about how Claude works.**
Q. Do you use Claude to help you write professionally?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: I use Claude to help me refine my own thoughts and maybe, like, have that impact show up in better writing.
MR. MILLER:
Q. If you need to send an email to somebody, will you have Claude help you draft the email?

39

**A. From time to time, yeah.**
Q. Outside of work, do you use Claude personally?
**A. Sometimes.**
Q. What kind of tasks do you use Claude for personally?
**A. Mostly to, like, understand aspects about Claude itself or just, like, things that I'm unfamiliar with to try to kind of get, like, baseline/ramped-up understanding on something.**
Q. Have you ever used Claude for any uses relating to songs or song lyrics?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: Probably.
MR. MILLER:
Q. Have you ever asked Claude to search for song lyrics?
**A. Probably. Again, I don't remember off the top of my head.**
Q. Have you ever asked Claude to try to identify a song based on lyrics?
**A. I might have, but I'm not sure.**
Q. When you've used Claude, has Claude ever generated output that's not factually accurate?
MR. HOLTZBLATT: Objection. Form.

40

THE WITNESS: I believe that's referred to as hallucinating. And, yes, I've seen instances of Claude hallucinate.
MR. MILLER:
Q. Can you give an example?
**A. I don't remember off the top of my head, but I think there -- there are definitely times that, like, I remember seeing an example of hallucination. But I don't -- I don't remember the example off the top of my head. I remember this with other models that I use as well.**
Q. When you've used Claude, has Claude ever regurgitated text from its training data?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: Can you define "regurgitated" in this sense.
MR. MILLER:
Q. Sure. When you use Claude, has Claude ever substantially reproduced verbatim text from its training data?
MR. HOLTZBLATT: Objection. Form.
THE WITNESS: Sorry. I'm just trying to understand. Like, it would be helpful if -- if you could kind of clarify what, like, "substantially reproduced verbatim" means.

# Pages 169-172

169

it says: "Current highest priority direction is increasing repetition." Do you see that?
**A. I do see that.**
Q. Do you know what that means?
**A. My best guess at what I think Kipply, if she wrote this document, may have meant -- I'm just kind of, like, picking up some context clues, but I still don't have a great idea who wrote the document -- would be that, like, trying to reduce harm from repetition is probably the -- the, like, direction that she sees to be the most promising.**
Q. And then there's a subbullet -- or dash under that that says: "We currently have very little understanding of memorization." Do you see that?
**A. I do.**
Q. In a large language model, what is memorization?
**A. That's a great question. I think there are vibes. I don't think there's a, like, clear consensus definition of memorization.**
Q. When Anthropic employees discuss memorization, what typically do they mean?
MR. HOLTZBLATT: Objection. Calls for speculation. Vague.

170

THE WITNESS: My understanding is it could refer to, like, a large contiguous block of tokens being memorized verbatim.
MR. MILLER:
Q. And when something -- when -- in -- withdrawn.
When an LLM has memorized a large continuous block of tokens, does that increase the likelihood that the model will output or regurgitate that block of tokens?
MR. HOLTZBLATT: Objection. Vague.
THE WITNESS: Can you define "regurgitate."
MR. MILLER:
Q. Yes. Produce an output in response to a prompt that includes the -- the verbatim block of tokens?
MR. HOLTZBLATT: Objection. Vague.
THE WITNESS: Sorry. Can you -- like, can you refine that a bit more.
MR. MILLER:
Q. What is the source of your confusion?
**A. I -- maybe can you restate the question in its entirety.**
Q. Sure. So you said that memorization could refer to a large continuous block of tokens being

171

memorized verbatim. When that happens, does that increase the likelihood that the model in response to a prompt will output that block of memorized tokens?
MR. HOLTZBLATT: Objection. Misstates prior testimony. Vague.
THE WITNESS: I think it depends on the prompt, and it depends on what the model is, like, being elicited to -- to do. If you could refine the question, I'm happy to try to answer it.
MR. MILLER:
Q. Do you have an understanding of what regurgitation is?
**A. Can you -- can you define what you mean by "regurgitation."**
Q. Well, I want to know do you have a -- what do you understand regurgitation to mean?
**A. I can guess. I don't think there's a term of art here that is, like, consensus industry definition. My best understanding could be that it is a verbatim reproduction of tokens in a large contiguous block.**
Q. That definition works for me.
**A. That's great.**
Q. Given that definition of regurgitation,

172

does a model's -- if a model is engaged in memorization, does that increase the risk of potential regurgitation?
MR. HOLTZBLATT: Objection. Vague.
THE WITNESS: I would not frame regurgitation as a risk. I think the risk is memorization. And if a large contiguous block of tokens has been memorized, it's -- it's possible that for some combination of prompts, those tokens will be reproduced. Maybe they would not be regurgitated; maybe they would. It's hard to say, but it's not unreasonable that that could occur.
MR. MILLER:
Q. Since -- strike that.
Do you see this document is dated August 3rd, 2023?
**A. That's right.**
Q. Since then, has Anthropic's understanding of memorization improved?
MR. HOLTZBLATT: Objection. Vague.
THE WITNESS: I would say we have learned more things that evolve our understanding of memorization.
MR. MILLER:
Q. You see there's a reference here to

# Pages 213-216

213

MR. HOLTZBLATT: Okay.
MR. MILLER: When we download the document for use as an exhibit, it doesn't preserve all of the metadata that's reflected in the production file.
MR. HOLTZBLATT: There is a way to do that, to preserve the metadata --
MR. MILLER: Okay.
MR. HOLTZBLATT: -- you agree?
MR. MILLER:
Q. Mr. Sangani, I'll represent to you that this document has the file name "Honoring robots.txt" followed by an underscore and a lot of random characters. Does that refresh your -- refresh your recollection at all about what this document might be?
MR. HOLTZBLATT: Objection. The questioning attorney is testifying, not the witness.
THE WITNESS: I think based on the dates here that are actually referenced in the document, you see that all the dates here are prior to my time on the team -- on the data team. So I joined roughly September of 2023. The most recent date here is August of 2023. I can guess about what this document is, but I don't, like -- this is probably

214

not created from material that I produced, and I -- I don't really remember this document. I'm sorry about that.
MR. MILLER:
Q. Okay. Can we go to Tab 12, please. And for the record, what's being marked as Plaintiff's Exhibit 130 bears the Bates No. ANTHROPIC342806.
**A. Okay.**
(Exhibit 130 was marked for identification.)
Q. Do you have Plaintiff's 130 in front of you?
**A. I -- sorry. Can you repeat that number again.**
MR. HOLTZBLATT: It's a lot of different numbers going on. You have got a tab number, a Bates stamp number, and then an exhibit number that's not placed on the document.
MR. MILLER: Yes.
MR. HOLTZBLATT: So there's some confusion here.
MR. MILLER:
Q. Yes. So for the record, Exhibit 130 -- you can ignore the tab numbers, which are just an internal reference number for us.
**A. Okay.**

215

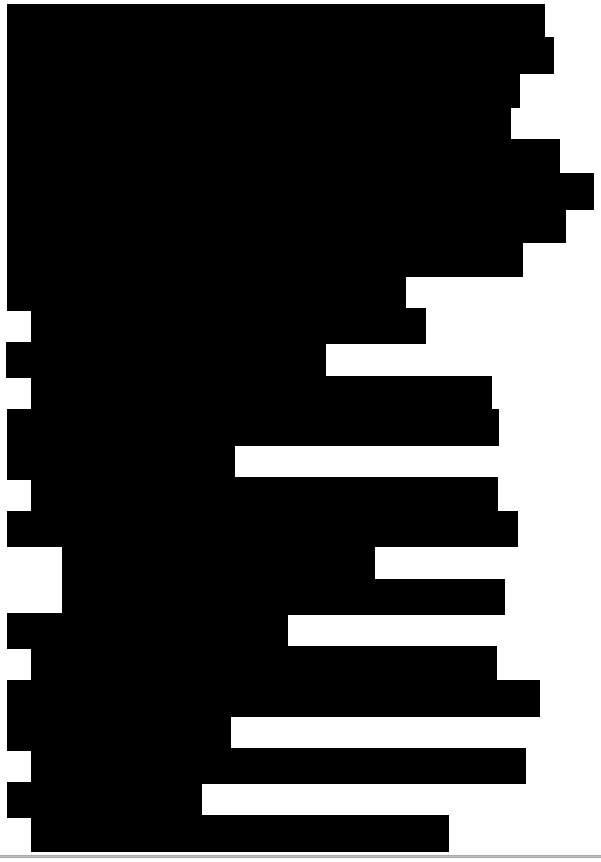
Q. Exhibit 130 should have the Bates number on the first page of ANTHROPIC342806.
**A. Sounds good.**
Q. And does this appear to be a notification from Slack to you dated November 22nd, 2023?
**A. Sure.**
Q. And it's a notification that Gyula sent you a message, right?
**A. That is -- I believe this is an email version of a notification with a lot of links and so on. But, yes, I think that is --**
Q. Yeah. I apologize that it's a little difficult to read, but this is the format in which it was produced to us.
**A. Yeah, I can see this doc.**
Q. Okay. Do you see -- I want to start in the middle of the page where it -- do you see where it says: "One more question"?
**A. I see that, yeah.**
Q. Do you remember this conversation with Gyula?
MR. HOLTZBLATT: Objection. Form.
[redacted]

216

[redacted]
MR. HOLTZBLATT: Objection. Hearsay. Form.
[redacted]

# Pages 261-264

261



262



263



Q. Does Anthropic use -- withdrawn.

While you were head of the pretraining data group, did Anthropic use its extractors to remove certain information or certain pieces of data from the -- for example, from the websites that it collected?

**A. It's inherently the job of the extractor to remove things that are cursory to its purpose of, like, extracting content; maybe extracting links; maybe, like, filtering for things like "Here is an ad" or something. And maybe that ad gets filtered out as part of the extraction process. I think it**

264

**sort of depends on the extractor. But it's sort of the purpose of the extractor to remove this type of cursory information.**

Q. Was one type of cursory information that Anthropic used extractors to remove the copyright information at the bottom of a web page?

MR. HOLTZBLATT: Objection. Vague.

THE WITNESS: I don't know if the extractor removed that or not. I -- I won't dispute it either way. Like, it's plausible that this is something that extractors removed. It's plausible that the extractors did not remove this.

MR. MILLER:

Q. Is copyright information on a web page something Anthropic would seek to remove as not helpful for training the model?

MR. HOLTZBLATT: Objection. Vague. Calls for speculation and corporate testimony.

THE WITNESS: It's my understanding that large repeated contiguous blocks of text that show up verbatim across many different sites that maybe just, like, add something like a copyright paragraph at the end or something -- yeah, I think generally this is not useful training information.

MR. MILLER: Okay. Can we take a short

# Pages 277-280

277

**A. Yes.**
Q. Do you believe that artists should receive compensation for the music that they create?
**A. I think if their music is good and people like listening to it, then they should have the ability to, like, participate in the market and -- and sort of do that. I don't think I should get, like, money for music that I may write because I'm not as good as Kendrick Lamar.**
Q. Neither am I. Do you believe that artists should receive credit and attribution for the music they create?
**A. I don't know what the, like, impactive derivative works there could look like. Like, it's not clear. I think in general, if it's a unique work of art that they create and there's a mechanism by which to afford them this credit, then, yeah, I support people getting credit for the work they -- they do in general.**
Q. Do you have any music streaming subscriptions?
**A. My partner might. I don't think I do unless it's covered by one of the, like -- unless it's, like, also a part of something else. Like, Amazon just gives you their thing -- or something.**

278

Q. Do you ever buy vinyl records or CDs or nondigital forms of music?
**A. I've probably purchased them at some point in the past. I don't regularly do that.**
Q. In your role as leading the data pretraining team -- is it data pretraining or pretraining data?
**A. It's pretraining data.**
Q. Okay. I think I mixed it up some --
**A. All good. I got you.**
Q. In your role as leading -- in leading the pretraining data team, did you have any responsibilities for fine-tuning?
**A. Maybe, like, incidentally but not directly. There's a separate group of people who do fine-tuning.**
Q. All right. Thank you for your time today, Mr. Sangani. I appreciate your patience. I have no further questions at this time. I pass the witness.
MR. HOLTZBLATT: Great. Let's take a short break. I can look at my notes, and then we can decide whether we have more to do.
THE WITNESS: Sounds great.
THE VIDEOGRAPHER: We are going off the record at 18:48.

279

(Break taken.)
THE VIDEOGRAPHER: We are back on the record at 18:56.
EXAMINATION BY MR. HOLTZBLATT, COUNSEL FOR DEFENDANT
MR. HOLTZBLATT:
Q. Okay. Mr. Sangani, just a few questions for you. During the time you led the pretraining data team, to what extent was it Anthropic's goal when training Claude to have it regurgitate large blocks of text verbatim?
**A. I don't think that was the goal at all.**
Q. Why not?
**A. Part of what we are trying to do when we train Claude is to have it learn the underlying patterns in the data as opposed to memorize the data verbatim. You can think of this as maybe, like, the difference between the purpose of what you -- what Google does when you sort of visit Google as a search engine. It's trying to give you an accurate outlink of sort of some of the pages that it's ranking so that you get the, like, content within those pages as sort of the thing that Google is doing when it's ranking those pages.**
**Training a large model is not that. We want Claude or other labs training other models want**

280

**their models to learn about the underlying dynamics and build a -- an understanding for the universe that allows the model to be useful in different contexts by essentially predicting the next token in the sequence. Memorizing data is kind of antithetical to that and is just probably not, like, what most ML practitioners and definitely not what we are looking for here.**
Q. What are some steps, if any, that Anthropic has taken to ensure that its models don't memorize or regurgitate verbatim the data it's training on?
MR. MILLER: Objection.
[redacted]
I'm just making half here as a -- as just sort of like to illustrate the point, not a specific