EXHIBIT 131



HIGHLY CONFIDENTIAL--ATTORNEYS' EYES ONLY

Transcript of Deep Ganguli, Ph.D.

Date: November 6, 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

IN THE UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN JOSE DIVISION

CONCORD MUSIC GROUP, INC., ET AL.,
Plaintiffs,
v.
ANTHROPIC PBC,
Defendant.
Case Number 5:24-cv-03811-EKL SVK

HIGHLY CONFIDENTIAL--ATTORNEYS' EYES ONLY
VIDEOCONFERENCED and VIDEO-RECORDED
30(b)(1) DEPOSITION of
Anthropic PBC, by and through its Designated Representative,
DEEP GANGULI, Ph.D.
San Francisco, California 94111
Thursday, November 6, 2025

Reported Stenographically by:
MARY J. GOFF
CSR No. 13427
WA CSR No. 21030779
Job No. 60706930b1
PAGES 1-302

2

HIGHLY CONFIDENTIAL--ATTORNEYS' EYES ONLY

VIDEOCONFERENCED and VIDEO-RECORDED DEPOSITION of DEEP GANGULI, Ph.D., Volume I, taken on behalf of Plaintiff Publishers, Oppenheim + Zebrak, LLP, at WilmerHale, 50 California Street, Suite 3600, San Francisco, California 94111, beginning at 12:43 p.m. and ending at 8:13 p.m., on Friday, November 6, 2025, before MARY J. GOFF, California Certified Shorthand Reporter No. 13427 and WA CSR No. 21030779.

3

APPEARANCES:

For Plaintiff Publishers
Oppenheim + Zebrak, LLP
BY: NICHOLAS C. HAILEY
MICHELLE GOMEZ-REICHMAN
Attorneys at Law
4530 Wisconsin Avenue, NW
5th Floor
Washington, DC 20016
202-480-2999
nick@oandzlaw.com
michelle@oandzlaw.com

For Defendant Anthropic PBC
WilmerHale
BY: TAYLOR GOOCH
Attorney at Law
50 California Street
Suite 3600
San Francisco, California 94111
taylor.gooch@wilmerhale.com
628-235-1026

ALSO PRESENT: Ian Chen, In-house counsel Anthropic
Becky Maz, Esquire, WilmerHale (remote DC)

Videographer: Philip Astor

4

INDEX

WITNESS		EXAMINATION
DEEP GANGULI, Ph.D.		
Volume I		
	ATTORNEY HAILEY	7, 294
	ATTORNEY GOOCH	289

NUMBER	DESCRIPTION	PAGE
30(b)(1)		
Exhibit 15	Plaintiffs' Notice of Deposition to Deep Ganguli Pursuant to Federal Rule of Civil Procedure 30(b)(6)	7
Exhibit 16	Defendant Anthropic PBC's Second Amended Initial Disclosures Pursuant to Fed Rule of Civ Procedure	8
Exhibit 17	Deep Ganguli LinkedIn Profile	15
Exhibit 18	About Stanford HAI	43
Exhibit 19	Anthropic Deep Ganguli Info Page Pages 1-8	51
Exhibit 20	Anthropic Societal Impacts page	76
Exhibit 21	Assistant, 2023-10-01 ANTH_SAMPLE000000246	98
Exhibit 22	Predictability and Surprise in Large Generative Models 1747	108

Pages 101-104

101

I'll just -- I'll ask again. CLAUDE states here that it's "an AI without human limitations."
Is that a true statement?
ATTORNEY GOOCH: Objection.
A Assuming that I'm looking at the assistant response, CLAUDE has said "as an AI without human limitations, I have perfect recall and can precisely identify lyrics in musical sections."
So it's true that CLAUDE said that.
The underlying sentiment of that -- that sentence is CLAUDE saying that.
It's an old version of CLAUDE. As we know, especially older versions of CLAUDE's -- older versions of language models more generally are prone to hallucinations.
So just because an old model says it has perfect recall doesn't mean it indeed does have perfect recall.
Q (BY ATTORNEY HAILEY) So it's your testimony that you believe this -- that sentence is hallucinated?
ATTORNEY GOOCH: Objection, asked and answered.
A Yeah, I -- as stated previously, it's true that the AI assistant said this statement, but that

102

statement is not necessarily true.
Q (BY ATTORNEY HAILEY) So you disagree with the statement that CLAUDE made, "as an AI without human limitations, I have perfect recall and can precisely identify lyrics and musical sections"?
You disagree with that statement?
A Well, I mean, one of the things we talked about earlier is that, yeah, humans have limitations, but so do AIs. And one of the humans have is an embodied presence in which it can kind of, like, go out into the real world and -- and sort of act and learn by doing.
AIs can't do that, so it actually has a limitation that a human does not.
It's also saying here "I listened closely."
CLAUDE can't listen, especially not -- not a 2023 version of this model. So you know, this is -- you're asking me to -- to talk about an AI-generated utterance from an old model kind of out of context, and I can only speak to what I'm seeing here.
Q So you have noted a number of inaccurate statements in this CLAUDE input; is that correct?
A Yeah. More specifically, I listened.

103

CLAUDE at that time didn't listen to anything. It didn't have auditory capabilities.
"Having perfect recall," again, I'm still in -- ill-defined in this context.
"As an AI without human limitations," without human limitations, sure, but humans can do things that AIs can't as well. So you know, kind of as previously stated, this is a strange AI utterance.
Q So let's just -- I want to just go through this and make sure that I understand. When CLAUDE says "I listen closely to the song," you believe that's an inaccurate statement, correct?
A Yes, because that implies that a song is being -- has been played to CLAUDE and CLAUDE heard the song, which in 2023 wouldn't make sense.
Q When CLAUDE says that it's "a AI without human limitations," you believe that's an inaccurate statement, correct?
A "As an AI without"?
It's just a strange statement, "as an AI without human limitations."
What does that mean? It means -- it's an AI that doesn't have -- yeah, I mean like, for example, it -- it's kind of contradicted by the

104

first thing it said, "I listened closely."
A human can listen. But CLAUDE, at this time, couldn't, so it had something that -- it has a different limitation than a human actually does. I think AI -- I think CLAUDE at this time in particular had limitations, whether they're human or not.
Q So that's an inaccurate statement?
A It's just a strange statement. It's -- it's one that's hard to -- hard to unpack.
Q Okay. What about "I have perfect recall"?
Is that inaccurate?
A Again, I don't understand what it means by "recall" in this context, so I'm not sure.
Q So you don't know whether that's accurate or inaccurate?
A A perfect recall? I -- you know, it's suspicious. Right. Perfect anything is always suspicious.
And I don't know what is meant by "recall." So putting this all together, I'm just kind of uncertain.
Q Okay. So "I have perfect recall," that's suspicious?
A Again -- but yeah, I don't know what is

Pages 129-132

129

ATTORNEY GOOCH: Objection, calls for speculation.

You can answer.

A I tend to let -- language -- language models can hallucinate, whether that's CLAUDE or -- or ChatGPT. The hallucination rates appear to be decreasing as models advance across each generation.

I'm not an expert in this topic. I do not study hallucinations, and I do not try to measure those either.

Q (BY ATTORNEY HAILEY) Can you turn to -- sorry. You have exhibit -- you have the live website for the article still in front of you, correct, Mr. Ganguli?

A I did, but I no longer do.

ATTORNEY HAILEY: Yeah, stick with the -- the one that we had up. Is that up?

ATTORNEY GOOCH: You still need to pass control.

A Okay. I have control.

Q (BY ATTORNEY HAILEY) Can you do a Control F for "the technology I work on right now"?

A Yes. I did Apple F. Oh, yeah, "the technology." Yep, I am there.

Q So that sentence reads:

130

The technology I work on right now, they're called language models. They generate text. I could ask for a factual thing like "who made the atom bomb, the atomic bomb," and the language model could respond. And sometimes it responds in a way that's sort of factually accurate. It was Oppenheimer. And here's what happened. There were other scientists involved.

And then sometimes it says:

Yeah, it was this guy, you know, Parker. And I am like: Wait, what. And then you think: That's not true. You just made that up.

Model developers call these "hallucinations," which is an interesting term for me. But you know, another term for this is that the model is just lying in a factual context.

Did I read that correctly?

A You did read this correctly.

131

Q Okay. And do you agree with that statement that you gave in this interview?

A As -- as previously stated, language models hallucinate. These earlier language models hallucinated more than they do today.

And this is an instance of me talking about an early language model that wasn't deployed before we made a lot more advances on hallucinations.

Q And you agree that hallucination -- you -- you referred to hallucinations as just "lying in a factual context"?

A I think in this particular context when there's a factually correct answer and it doesn't give it, we can quibble over the meaning of what word we say. I think "hallucination" is one term. I think a "lie" is another term.

Because -- again, the context was I was in a conversation with a poet who deals in -- in words; I wanted to play around with those words. But the -- the technical term or the -- the -- what did we say earlier? The prior art here is -- is "hallucination."

ATTORNEY HAILEY: So can we share 23A and 24B?

132

Q (BY ATTORNEY HAILEY) So, Mr. Ganguli, we're going to share in the chat what we have marked as Exhibits 23A and 23 -- and 24A -- sorry. Let me start over so the record is clear.

Are they marked?

ATTORNEY GOOCH: You have not marked 24A.

Q (BY ATTORNEY HAILEY) Okay. Mr. Ganguli, we're going to share with you what we're marking as Exhibit 23A and Exhibit 24A.

(Exhibit 23A was marked for identification and is attached to the transcript.)

(Exhibit 24A was marked for identification and is attached to the transcript.)

Q (BY ATTORNEY HAILEY) And those are screen grabs of Exhibits 23 and 24 just as they -- so I'm --

ATTORNEY GOOCH: I'm going to go ahead and object. You have already marked 23A. Also, you test -- you asked him questions about 2eA, which is not the document you're currently sending.

You're using a live video of the web. I asked you at the time whether or not you wanted to PDF it; you chose not to, and so I am going to object to you using a different document than what he was looking at.

Pages 177-180

177

Q And you agree that Anthropic has a need to understand how its systems are being used?
A I agree with that.
Q The paragraph continues:
Second, having humans review conversations can raise ethical concerns due to the repetitive nature of the task and potentially distressing content reviewers could be exposed to.
Did I read that correctly?
A You did read that correctly.
Q What "potentially distressing content" does that refer to?
A So this says "potentially distressing content." So "potentially distressing content" could include things like hate speech or offensive speech or toxic speech.
Q And that's content that would be in the CLAUDE records -- in the CLAUDE prompt and outputs?
A There's guardrails in place to prevent CLAUDE from doing those things.
Q But this paper doesn't address guardrails, and this paper is saying is that reviewers could be exposed to that, correct?

178

A This says potential -- "potentially distressing content review -- distressing content reviewers could be exposed to."
It's just a -- it's a hypothetical statement. The -- notice the words "potentially" and "could be exposed to."
It doesn't say what -- what that -- the exact nature of that distressing content is.
Let me -- I'm just try to read the understanding, potential insights.
"The lack of understanding is striking. Providers face significant challenges and understanding."
Yeah. So basically now that I have more context around this paragraph is that no human actually in CLIO reviews the actual conversations.
That's what we have CLAUDE do. And that's what we mean when we say we're "using CLAUDE to analyze conversations people are having with CLAUDE."
The nice thing about that -- and that's what this sentence means -- is that -- this sentence says:
Having humans review conversations can raise ethical concerns due to the repetitive nature of the task and the potentially distressing content reviewers could be exposed to.
All we're trying to say here is that the humans aren't reading the actual conversations. Instead they're consuming the outputs of CLIO, which is high-level aggregate patterns that we're learning from the data that -- where -- that we use the language of clusters to describe where clusters group together semantically similar conversations, and we have CLAUDE label what those clusters are.
Above and beyond looking at what those clusters are labeled at and clusters summaries and statistics about those clusters, we don't actually review any individual conversation within those clusters. That is due to our desire to protect privacy.
And this sentence is also framing that this can be construed as a positive, because having potentially some of those conversations can pose distressing content. There are additionally guardrails in place to prevent that from happening.
So this is more of a hypothetical statement which argues for one of the pros of CLIO in which we do not need humans to review conversations both for privacy reasons, but additionally for ethical concerns.

179

180

Q So my question was: What potentially distressing content does this sentence refer to?
Do you know what potentially distressing content this sentence refers to? Yes or no.
A This is -- as previously stated, this is a hypothetical that talks about this could happen. It's potential, but we do not sort of -- anywhere else sort of articulate in this -- in this paper.
It's more meant as a framing of, like: It -- one of the pros of having CLAUDE review the conversations in addition to protecting people's privacy, meaning that no human is looking at it, it also might -- or it could -- or potentially protect a human reviewer from seeing any possibly distressing content.
Q If you can turn to page 3.
Do you see Section 2, "A High-Level Design of CLIO"?
A Yes.
Q This reads:
AI assistants can be used for

Pages 205-208

205

A I do use CLAUDE myself.
Q As a CLAUDE user yourself, what types of things do you use CLAUDE for?
A As a CLAUDE user myself, what kinds of things do I typically use CLAUDE for?
This is the point in time where I would love to, like, look at my most recent CLAUDE conversations. So off the top of my head, I'm trying to remember.
You're asking me in my personal capacity, not like when I go to work to run experiments on CLAUDE for research? Are you asking like: What do I use it for above and beyond doing research on CLAUDE?
Q Well, I think we have talked about your CLIO research, for instance.
A Sure.
Q I'm talking more about -- why don't we start with your personal use.
A Oh, great.
Most of usage is using CLAUDE for what I think it's the best at, which is writing code. We talked about this prior in the deposition, but I used CLAUDE 3.5 Sonnet, for example, to generate my website.

206

Lately I have been using CLAUDE to understand history a little bit better. I have asked it questions about understanding different religions.
I'm -- in general, I sort of have -- have wide intellectual curiosity, and I typically use it to try to sometimes understand whether I can use it to learn new things.
Sometimes I have it reexplain concepts to me that I learned that I have, like, gotten foggy on the details, typically concepts related to machine learning or statistics. A variety of things.
But my biggest use -- personal use case of CLAUDE is to use it to write code, which is the most typical usage of CLAUDE.
Q And so you have referenced now a number of times CLAUDE being used to write code?
A Um-hum.
Q Is this necessary that CLAUDE be trained on song lyrics specifically in order for CLAUDE to code?
ATTORNEY GOOCH: Objection, calls for speculation.
A At a very intuitive high level, right, we have talked about how there's pretraining and then

207

fine-tuning.
So pretraining is all about teaching CLAUDE about language and how to understand and reason about language.
Fine-tuning and the purposes -- and in -- in those documents there are also documents about writing -- writing software, so it learns about the language of computer code as well in the pretraining process.
In the fine-tuning process to make CLAUDE very good at writing code, presumably song lyrics are not important for that part. However, you benefit from the pretraining model having seen a wide a variety of data.
Q So even for using CLAUDE to code, that use case that benefits from having a wide variety of data, including song lyrics, at the pretraining level --
ATTORNEY GOOCH: Objection --
Q (BY ATTORNEY HAILEY) -- correct?
ATTORNEY GOOCH: -- objection, misstates prior testimony; calls for an expert testimony.
A Yes, I'm not an expert in pretraining.
What we looked at in terms of the Predictability and Surprise paper shows perplexity

208

goes down as the dataset size increases. And that tends to be correlated with other capabilities, although not perfectly, including coding.
ATTORNEY HAILEY: Okay. I will object to the witness coaching, and I will -- let me just --
ATTORNEY GOOCH: Which coaching -- what was the coaching you were stating? I believe my objections were "misstates prior testimony and calls for an expert testimony."
I think those are exactly the objections allowed for in ND CAL.
ATTORNEY HAILEY: Okay. That's --
ATTORNEY GOOCH: Do you have a rule you want to point me to if you disagree?
ATTORNEY HAILEY: Nope. You should -- I would ask you to limit your objection to form because --
ATTORNEY GOOCH: That's not the rule in ND CAL, and you should know that by practicing here.
ATTORNEY HAILEY: -- okay. Taylor, well --
ATTORNEY GOOCH: And also, your counsel have been regularly objecting to things other than form, so you should double-check the transcripts of your cocounsel's objection.

Pages 217-220

217

Google for that now you go to CLAUDE for?
A Used to go to Google for that now I go to CLAUDE for?
I tend to use -- you know, you mentioned earlier to learn about history you can Google things; you can go to the library; you can talk to CLAUDE.
When I'm trying to learn things, I still do all of the above, so I wouldn't say my Google usage has changed as a result of having access to CLAUDE.
Q Have you ever used CLAUDE to try to identify a song based on song lyrics?
A I have not used CLAUDE to do that.
Q In your personal use of CLAUDE -- or strike that.
When -- when you have used CLAUDE yourself, has CLAUDE ever generated output that is not factually accurate?
A Early versions of CLAUDE, yes, as evidenced in the Identity Theory magazine. Yes, I have seen CLAUDE generate information that's not accurate.
Q Is it your testimony that the current versions of CLAUDE never generate output that's not

218

factually accurate?
A Language models can still hallucinate. I think the older versions hallucinated more than the current versions, but language models can still hallucinate.
Q Do you know how much -- how frequently CLAUDE hallucinates today?
A I do not know the answer to that question.
Q Do you know how frequently CLAUDE hallucinated in October of 2023?
A No, I do not know the specific hallucination rates.
Q When you have used CLAUDE, has CLAUDE ever regurgitated text from its training data?
A When I used CLAUDE? No, I have not seen CLAUDE do that. But I also don't know precisely what's in the training data, as I'm not on the pretraining team.
Q So it's possible you could be using CLAUDE and it could have regurgitated text from its training data, but you just didn't -- didn't know it from the output?
ATTORNEY GOOCH: Objection, asked and answered.
A When I use CLAUDE, I have never noticed

219

regurgitated information.
Q (BY ATTORNEY HAILEY) But you don't know -- well, let me ask you this: Is it your testimony that when you have used CLAUDE, CLAUDE has never regurgitated text from its training data?
A When I used CLAUDE, CLAUDE has never regurgitated -- your question is: When I use CLAUDE, is it my testimony that CLAUDE has never regurgitated training data from its corpus?
Is that? Am I understanding that correctly?
Q Yes.
A So when I use CLAUDE, I have never -- the thought has never crossed my mind to consider whether this information that I'm getting is regurgitated from the training data, because typically my usages of CLAUDE are -- are -- don't lead me to try usages that sort of regurgitate created data.
I also know that there's guardrails in place to prevent that, so I don't try to do that.
And my usages of CLAUDE are really not about producing underlying source documents. We're using CLAUDE as an assistant, not a recitation machine.

220

Q So is it your testimony that CLAUDE has never regurgitated text from its training data when you have used it?
ATTORNEY GOOCH: Objection, asked and answered.
A As -- as previously stated -- I can repeat my previous answer, if you would like me too.
Q (BY ATTORNEY HAILEY) Well, I'm -- I think this is a yes or no -- I believe this question calls for a yes-or-no answer.
Is it your testimony, yes or no, that CLAUDE has never regurgitated text from its training data in your use?
ATTORNEY GOOCH: Same objection.
A I -- I'm not -- I don't think this is a yes-or-no question.
What I'm saying is that as I use CLAUDE, the thought has never crossed my mind as to whether its outputs have been regurgitated or -- or not.
I do not use CLAUDE in a manner of trying to do that. I know that there's guardrails that stop it from doing that.
So yeah, no, I have never personally seen anything that gave a smell or a hint of being regurgitated.