EXHIBIT 105

1

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

- - - - - - - - - - - - - - -x
CONCORD MUSIC GROUP, INC., ET AL., : Civil Action No.
Plaintiffs, : 5:24-cv-03811-EKL
v. : SVK
ANTHROPIC PBC, :
Defendant. :
- - - - - - - - - - - - - - -x

Videotaped Deposition of JEFFREY BILMES, Ph.D.
* * * HIGHLY CONFIDENTIAL * * *
* * * ATTORNEYS' EYES ONLY - SOURCE CODE * * *
Seattle, Washington
Friday, March 6, 2026
9:00 a.m.

Job No.: 620607
Pages: 1 - 255
Reported By: Connie Recob, CCR 2631, RMR, CRR

2

Deposition of JEFFREY BILMES, Ph.D., held at the offices of:

Cooley LLP
1700 Seventh Avenue, Suite 1900
Seattle, Washington 98101

Pursuant to notice, before Connie Recob, CCR 2631, RMR, CRR, Notary Public in and for the State of Washington.

3

A P P E A R A N C E S

ON BEHALF OF PLAINTIFFS:
COREY MILLER
OPPENHEIM + ZEBRAK, LLP
4530 Wisconsin Avenue, Fifth Floor
Washington, D.C. 20016
(202) 957-5043
corey@oandzlaw.com

TIMOTHY CHUNG (via videoconference)
OPPENHEIM + ZEBRAK, LLP
461 Fifth Avenue, 19th Floor
New York, New York 10017
(212) 951-0521
timothy@oandzlaw.com

ON BEHALF OF DEFENDANT:
TAYLOR GOOCH
WILMERHALE
50 California Street, Suite 3600
San Francisco, California 94111
(628) 235-1026
taylor.gooch@wilmerhale.com

4

A P P E A R A N C E S C O N T I N U E D

ON BEHALF OF DEFENDANT:
ELLEN STEPHANO
WILMERHALE
60 State Street
Boston, Massachusetts 02109
(617) 526-6375
ellen.stephano@wilmerhale.com

ALSO PRESENT:
Parker Simpson - Videographer
Ian Chen - Anthropic PBC

5

EXAMINATION INDEX

EXAMINATION BY:		PAGE NO.
BY ATTORNEY MILLER		7

EXHIBIT INDEX

EXHIBIT NO.	DESCRIPTION	PAGE NO.
Exhibit 319	Opening Expert Report dated 12/19/25	6
Exhibit 320	Rebuttal Expert Report dated 1/23/26	6
Exhibit 321	Reply Expert Report dated 2/13/26	6
Exhibit 322	AI Model & API Providers Analysis|Artificial Analysis	57
Exhibit 323	Don McLean Lyrics, Bates No. Plaintiffs_0000000564	204

WITNESS INSTRUCTED NOT TO ANSWER
Pages 247-248

INFORMATION REQUESTED
(None)

6

P R O C E E D I N G S

(Exhibit Nos. 319-321 marked for identification.)

THE VIDEOGRAPHER: Here begins the videotaped deposition of Professor Jeffrey Bilmes, in the matter of Concord Music Group, Incorporated, et al., versus Anthropic PBC, in the United States District Court, Northern District of California San Jose Division, Case No. 5:24-cv-03811-EKL-C -- or SVK.

Today's date is March 6th, 2026. The time on the video monitor is 9 o'clock Pacific time. The videographer for today is Parker Simpson representing Planet Depos, headquartered at 451 Hungerford Drive, Suite 400, Rockville, Maryland 20850.

This video deposition is taking place at Cooley LLP, Seattle, 1700 Seventh Avenue, Suite 1900, Seattle, Washington 98101.

Would counsel please voice-identify themselves and state whom they represent.

ATTORNEY MILLER: Corey Miller from O+Z LLP for the music publisher plaintiffs. Also present on the Zoom with me is my colleague, Tim Chung.

7

ATTORNEY GOOCH: Taylor Gooch for WilmerHale. Also with me in the room is our client's in-house attorney, my colleague Ellen Stephano and the witness on behalf of Anthropic.

THE VIDEOGRAPHER: And the court reporter today is Connie Recob representing Planet Depos.

The witness will now be sworn.

JEFFREY BILMES, Ph.D., having been first duly sworn, deposed and testified as follows:

EXAMINATION

BY ATTORNEY MILLER:

Q Good morning, Professor Bilmes.

A Good morning.

Q My name is Corey Miller and I'll be taking your deposition today. Have you ever been deposed before?

A No, I haven't.

Q So let me just briefly walk through how today is going to proceed. I'll be asking you questions and it's your job to answer my questions

8

as truthfully and completely as you can.

From time to time your counsel may object. Even if he objects, you still have to answer my questions unless he specifically instructs you not to answer.

Do you understand that?

A Yes.

Q We have a court reporter here today who is writing down everything we say so that there's a record of today's proceedings. For her benefit, it's very important that you give verbal answers to questions. Shakes of the head or "uh-huhs" or that kind of thing don't produce a clear record of what your answer was so I'll ask that you give verbal answers.

Similarly, if we talk over each other, it makes it much harder for the court reporter to create a clear record, so I'll ask that you wait until I finished asking my question before answering and in return I will try to wait until you finished answering my question before asking the next question.

Does that make sense?

A Yes.

Q If you do not understand a question that I

101

Q If a piece of content is present in a model's training data and the model outputs a near exact reproduction of that piece of content or a substantial portion of that piece of content, isn't that evidence that the model has memorized that piece of content?

ATTORNEY GOOCH: Objection. Form. Asked and answered.

THE WITNESS: I don't agree with that, that that would be evidence.

BY ATTORNEY MILLER:

Q So you don't think that the near exact reproduction of a substantial piece of training data present in the model's training dataset indicates that the model has likely memorized that piece of training data?

ATTORNEY GOOCH: Form. Asked and answered.

THE WITNESS: So when one thinks about the capabilities of these models which I think we've talked about as being very, very large, almost unimaginably large with capabilities that we haven't -- we as the collective community of AI scientists and researchers have not yet fully fleshed out and are aggressively studying, even if

102

it's the case that there's something in the training data, one might think it could be the case but there could be other reasons because of the fact, as I was saying a little bit earlier, that there's no true available known causal model that can allow us to establish the cause of a particular output as being a particular piece of training data.

It's a very complicated process and it's a very subtle distinction, and to be honest it's a very interesting distinction. But if one wants to really know something, one needs to think about these things and study them, and this is what, as I was saying, the collective community of artificial intelligence and machine learning, scientists and researchers and engineers are studying.

So one can't know that. Maybe some day, maybe it's 50 years. Maybe it'll be ten years. Maybe there will be the development of a true causal model that one can deduce from one of these many hundreds of billion parameter models to be able to establish the true cause of a particular output, but that is not yet in the capability of the community of machine learning scientists and

103

researchers or AI scientists and researchers.

BY ATTORNEY MILLER:

Q Has Claude memorized any of the works in suit that are at issue in this case?

ATTORNEY GOOCH: Asked and answered. Form.

THE WITNESS: So any model, any large model or even smaller models, but this becomes even more challenging when the models get larger, one can't know for certain whether or not a particular output is caused by a particular piece of training data.

There is evidence to suggest that a model can output something that seems like it's something that exists out there even without being in the training data. The challenge really is the bigness and the interaction and the complexity of the system. If there are say -- if there is a model with, say, 16 trillion tokens of training data, one could potentially in theory do an experiment where you train 2 raised to the 16 trillion times. That's 2 to the 16th trillion, so that would mean doing that many re-trainings to study the behavior of all of the resulting 2 raised to the 16 trillion model. 2 to the 16

104

trillion is a ridiculously large number. The number of atoms in the universe is estimated to be less than 2 to the 100. This is a number that's 2 raised to the 16 trillion trainings, each of which would include or not include some subset of tokens, and each of which would require many months of training using tens of thousands of high performance GPUs and possibly also varying the underlying model architecture to really develop such a thing.

Doing things like that are completely infeasible. The reason why I say that it's an interesting question is because one can only, you know, speculate but one doesn't want to speculate because one can't about what might come out of such an experiment. So, therefore, as a result, because of the infeasibility of doing anything like that, the community of machine learning scientists and researchers are trying hard to come up with strategies that might be able to answer your question more accurately. But again, I do want to reiterate that these are very interesting questions.

BY ATTORNEY MILLER:

Q So is it your view then that memorization

177

you could think of as distortions. And there could be various different amounts of each one of these various different types of distortions.

So the first part of your question is a hypothetical, which would not be something necessarily -- that would not be something possibly to define or even do because of all of those things.

Secondly -- so the second part of your question was what would be the consequence of doing something that is effectively impossible to do. So the consequence -- let's say that we are in a universe where the impossible is possible. And then the model, Anthropic's Claude, because the training data has changed, would behave differently.

Now, the reason is because this goes back to what we were saying a little bit earlier that, especially when you have models as powerful as state-of-the-art large language models deployed by commercial entities such as OpenAI, Google, Anthropic and other entities that have the abilities to produce these large models, one never really knows the effect that one piece of data has on everything else that the model does. One knows

178

that the model behaves differently. One knows that every training data can change the billions of floating point numbers -- floating point numbers that comprise the underlying model's parameters. And the resulting model is different. It will behave differently.

And it's very counterintuitive at first, because one might think, as an example, that a recipe for, say, lasagna would have no effect on one's ability to get instructions on how to fix a carburetor in a car, but if you removed that recipe for lasagna, it's the case that the precise mechanism that goes into producing the instructions for fixing a carburetor on a car would be different even though these two concepts to us humans, to us people, who would not think of those two things as being related to each other, would not find a lasagna recipe useful for fixing a car.

These models work in very complicated ways, and the community is collectively working really, really hard and really hard to try to understand precisely mathematically how are they doing what they are doing.

There are some theories. Most of the

179

theories that I've seen are based on tiny little chunks of the model. So as good as the existent mathematical theories are, they're theories about a very small chunk. They're theories about the small -- and they themselves would be first to admit that, despite their underlying mathematics being correct, they don't know if their theories and their results would themselves still apply once you scale up those results to the sized models that -- you know, Claude or ChatGPT or any of the models that, you know, comprise the state-of-the-art large language models are.

So therefore, getting back to your question, if all of the copies of the works in suit were removed from the training data, how would it affect the resulting model or would it effect the resulting model? I can't remember which of those two variants it was. But both sides of those questions, to really accurately answer it would need to be teased apart.

First is really the infeasibility and impossibility of doing the first part. And secondly, the fact that any piece of data can have an effect on any other ability that the model might possess. And it's -- it's quite remarkable.

180

I mean, it really is.

Q Do you disagree that the effect of Plaintiffs' lyrics on Claude's performance is extremely likely to be near zero?

ATTORNEY GOOCH: Objection. Asked and answered.

THE WITNESS: I'm sorry. Might you -- would you mind rephrasing the question?

BY ATTORNEY MILLER:

Q Do you agree that the effect of Plaintiffs' lyrics on Claude's performance is extremely likely to be near zero, if present at all?

ATTORNEY GOOCH: Objection. Asked and answered.

THE WITNESS: My understanding is that not just lyrics, but every piece of data, every iota of data has some effect. It might be extremely small, but every piece of data has some effect on the model's behavior.

The models can do so many things in so many different ways that changing the data in any way changes the model.

BY ATTORNEY MILLER:

Q If you turn to Paragraph 328 of your

181

opening report, please, I want to direct your attention to the last sentence of this paragraph where you write, "At this magnitude, the effect of Plaintiffs' lyrics on Claude's learned parameters or outputs is extremely likely to be near zero, if present at all."

Do you see that?

A I do see that.

Q Do you agree with that statement?

A I agree with that statement.

Q How is that consistent with what you just testified, that every piece of data has some effect on the model's behavior?

A Because near zero is not zero. Sometimes in mathematics one can say that you could be near zero, but that's completely different than being actually zero. You can divide by a number that's near zero, and you don't get an infinity, but if you divide by zero, you get infinity.

So it's just not -- they're very different near zero and zero. Despite near zero sounding like it's zero, it's not. Near zero is not zero. So while it may be the case -- and I can explain how it's consistent. While it may be the case that the effect is near zero, that doesn't mean

182

it's zero. And what I was saying earlier was that every piece of data has some effect on the behavior. It may be very small. It may be near zero. It may be immeasurably small using our current mechanisms of measurability in a given fixed amount of time, but that doesn't mean it's zero necessarily. It does not. So near zero is different than zero.

Q Would excluding the works in suit from Anthropic's training data, assuming that's possible, which we'll address separately in a moment, have any effect that would be large enough to be noticeable in a practical sense?

ATTORNEY GOOCH: Form. Asked and answered.

THE WITNESS: So it depends what one's -- what -- what one would mean by "in a practical sense." So what I'm saying is it changes the model. It's near zero. It's not zero. It may be immeasurably small. These things are so large. You might run a number of tests over a period of time to see the difference, but because these things are so large, the tests wouldn't -- may not be enough time to -- to observe the difference. It may be that to -- for the test to show

183

difference.

This is what we mean by near zero. We would need to run the test for longer than the amount of time, maybe many hundreds of years before we might see a particular difference. But the point is that we know that every piece of training data does have an effect on the resulting model.

There's a notion of diminishing returns and there's a notion of scaling laws in the machine learning community. And the scaling laws use mathematical models to express the notion of diminishing runs. And the functions, the mathematical functional form that they use is such that the effect of any piece of training data -- this is something that has been tested many, many times in machine learning community. The functional form is such that the effect of any piece of data, no matter how small, on these mathematical models on top of existing data is never zero. It's near zero, but it's never zero.

And I could -- if you're interested, I would be happy to talk a little bit more about the specific mathematical forms of these functions of scaling laws, or I could, if you want, give you

184

another example of a slightly simpler mathematical form of a function that would also have that particular behavior.

BY ATTORNEY MILLER:

Q You -- you said that the effect of any piece of data, no matter how small, is near zero but it's never zero. Is saying that the effect is near zero saying that a piece of data, no matter how small, impacts the LLM's performance?

A I'm sorry.

ATTORNEY GOOCH: Objection. Form. Go ahead.

THE WITNESS: Sorry. Might you be willing to point me to a particular sentence or paragraph so that I could clarify it? I believe you were not referring to a particular paragraph. When you were saying -- you were suggesting that I was saying something, and I was wondering if you wouldn't mind pointing to a paragraph.

BY ATTORNEY MILLER:

Q I was looking at your previous testimony. And you said -- you testified that the effect of any piece of data, no matter how small, is never zero.

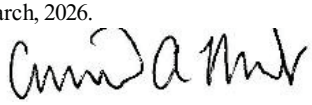
Is that the same as saying that any piece

253

REPORTER'S CERTIFICATE

I, CONNIE A. RECOB, the undersigned Certified Court Reporter, authorized to administer oaths and affirmations in and for the States of Washington, Oregon, Utah and Idaho, do hereby certify that the sworn testimony and/or proceedings, a transcript of which is attached, was given before me at the time and place stated therein; that any and/or all witness(es) were duly sworn to testify to the truth; that the sworn testimony and/or proceedings were by me stenographically recorded and transcribed under my supervision, to the best of my ability; that the foregoing transcript contains a full, true, and accurate record of all the sworn testimony and/or proceedings given and occurring at the time and place stated in the transcript; that a review of which was requested; that I am in no way related to any party to the matter, nor to any counsel, nor do I have any financial interest in the event of the cause.

WITNESS MY HAND and SIGNATURE this 10th day of March, 2026.



/s/CONNIE A. RECOB, RMR, CRR
Washington CCR No. 2631

254

DEPOSITION ERRATA SHEET

Our Assignment No. 620607
Case Caption: CONCORD MUSIC GROUP vs. ANTHROPIC

DECLARATION UNDER PENALTY OF PERJURY

I declare under penalty of perjury that I have read the entire transcript of my Deposition taken in the captioned matter or the same has been read to me, and the same is true and accurate, save and except for changes and/or corrections, if any, as indicated by me on the DEPOSITION ERRATA SHEET hereof, with the understanding that I offer these changes as if still under oath.

Signed on the _____ day of _____, 2026.

_____
JEFFREY BILMES, Ph.D.

255

DEPOSITION ERRATA SHEET

Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____
Page No._____Line No._____Change to:_____
_____
Reason for change:_____

SIGNATURE:_____DATE:_____
JEFFREY BILMES, Ph.D.

ERRATA TO THE DEPOSITION TRANSCRIPT OF
JEFFREY BILMES, PH.D.
TAKEN ON **March 6, 2026**

I have reviewed the transcript of my testimony of **March 6, 2026**, and hereby make the following changes in form and/or substance:

PAGE	LINE(S)	CHANGE	REASON FOR CHANGE
11	4	"analysis" to "analyses"	Transcription Error
13	9	"at" to "of"	Transcription Error
14	2-3	"microelectronic machines. It's various" to "Micro-Electro-Mechanical System (MEMS). It uses various"	Transcription Error / Clarification
14	5-6	"electrical computing engineering department are to think" to "electrical and computer engineering department is to think"	Transcription Error
14	13	"microelectronic machine" to "Micro-Electro-Mechanical Systems"	Transcription Error / Clarification
14	19	"computing" to "computer"	Transcription Error
16	1	Omit "and"	Transcription Error
39	20	".xlxs" to ".xlsx"	Transcription Error
42	18	"hallucination" to "LLMs"	Clarification
49	10	"number" to "model"	Transcription Error
82	24	Omit "me hear"	Transcription Error
90	16-17	"not that one would ever want." to "not one that one would ever want."	Transcription Error
94	18-19	"four bad bananas" to "two bad bananas"	Transcription Error
103	21	"train 2 raised" to "train models 2 raised"	Transcription Error
103	22	"16 trillionth" to "16 trillionth models"	Transcription Error
104	3	"2 to the 100" to "10 to the 100"	Clarification
106	22	"diverse" to "diversity"	Transcription Error

ERRATA TO THE DEPOSITION TRANSCRIPT OF
JEFFREY BILMES, PH.D.
TAKEN ON **March 6, 2026**

PAGE	LINE(S)	CHANGE	REASON FOR CHANGE
106	23	"reverse" to "diverse"	Transcription Error
107	10	"lease" to "less"	Transcription Error
130	22	"enumerable" to "innumerable"	Transcription Error
141	11	"predomics" to "proteomics"	Transcription Error
171	20	"tension" to "attention"	Transcription Error
171	21	"perception" to "perceptron"	Transcription Error
177	7	"possibly" to "possible"	Transcription Error
178	2	"every training data" to "every piece of training data"	Transcription Error
179	17	"effect" to "affect"	Transcription Error
200	19	"tamber" to "timbre"	Transcription Error
237	13	"AP" to "REP"	Transcription Error
242	3	"Komi" to "inktomi"	Transcription Error

I, **JEFFREY BILMES, PH.D.,** read the foregoing deposition and hereby state that the foregoing is true and correct, except as noted herein.

Jeffrey Bilmes, Ph.D.

04/09/2026
Date