EXHIBIT 1

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

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)

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.

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

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

33

ATTORNEY MILLER: Okay.
THE WITNESS: So yes.
BY ATTORNEY MILLER:
Q And then do you have a copy of the third report you submitted in this case in front of you?
A Yes.
Q We'll designate that in the record as Plaintiffs' Exhibit 321. Is Exhibit 321 a complete and accurate copy of the expert report you submitted in this case on February 13, 2026?
A Yes.
Q Okay. Did you use AI in connection with the preparation of any of your expert reports in this case?
A No.
Q Why not?
A I -- why did I not use AI?
Q Uh-huh.
A There were several reasons: One, I didn't need to, and I was informed not to.
Q Okay. Did you speak to an Anthropic employee named Benjamin Mann in connection with preparing your opinions in this case?
A There were very brief instances where there were conversations that I had with counsel

34

present to Ben Mann.
Q And how many times did you speak to Mr. Mann?
A I don't remember exactly, but it was approximately two or three times.
Q How long did those conversations last?
A They lasted no more than five, perhaps at most ten minutes but probably closer to five minutes.
Q Did you take notes during those conversations?
A No.
Q Did you speak with any other Anthropic employees in connection with preparing your opinions in this case?
A No.
Q Do you know anyone who works at Anthropic?
A I'm just thinking about your question. Sorry for the delay. Surprisingly, no, other than people in this room.
Q In connection with preparing your opinions in this case, did you speak to anyone who works in the music publishing business?
A No.
Q In your reports, you discuss Anthropic's

35

process for pretraining its Claude models, correct?
A I'm sorry. Could you point me to where in the reports this is the case?
Q For example, starting on Page 66 of your opening report. In Section 5(a) of this report you discuss LLM pretraining, right?
A So Section 5(a) in my report is regarding general LLM training.
Q And that also includes -- that section of your report also includes a discussion of how Anthropic pretrains its Claude models; is that right?
A Is there a specific paragraph that you can point me to? This section is primarily about general LLM pretraining, but I'm happy to answer a question about a specific paragraph or section or subsection within 5(a).
Q Does this section not also apply to how Anthropic trains Claude?
A It applies to all large language model training method strategies in general, so it's a very wide general description of LLM training.
Q Did you review any of the specific processes Anthropic uses for pretraining its

36

Claude models?
A So as part of my report, I reviewed a large number of sources with regarding general LLM training which includes all of the ways that all of the entities who produce these large language models train those large language models.
Q Did you review any of the code Anthropic has used for pretraining its Claude models?
A Are you asking if I looked at specific code of specifically Anthropic?
Q Uh-huh.
A In terms of pretraining? I have not looked at the specific code that Anthropic uses for pretraining.
Q Did you review any of the data mixes Anthropic has used for pretraining its Claude models?
A May I ask for clarification?
Q Yes.
A What do you mean precisely by "data mixes"?
Q Well, what's a -- what -- are you -- you understand that Anthropic uses large datasets to pretrain its Claude models?
A My understanding is that anyone who wants

105

can occur during the training process but it's impossible to know whether it's occurred?

A If one is being true to oneself, one can't 100 percent know because of the properties that I just mentioned, because one can't imagine that a model might be able to produce something that exists out there that has been assembled together from distorted fragments and lots of other pieces of information so that the behavior of the model could be that it behaves very similarly when that information is present or absent.

It doesn't mean that the model behaves identically, but it would mean that for that particular instance, it would behave similarly enough.

So as a result, one can't really know for certain. One can speculate. One can surmise. One can guess, but I don't think one can truly know, at least not yet.

Q If you can't truly know whether memorization is occurring, why go through all the effort to attempt to reduce it?

ATTORNEY GOOCH: Form.

THE WITNESS: There are many reasons. One might want to reduce the appearance of

106

memorization. Even if memorization is not there, because one can't truly know, one doesn't want it to be appearing to memorize, and so, therefore, one goes through efforts to reduce that chance.

And the other answer to that question is related specifically to de-duplication. There are other adverse effects that redundant imbalanced data can have on the model. Redundancy is generally bad in training data. Diversity and coverage and bits of everything is generally most preferred in training data that usually we have found, the community has found, provides the best model.

And so it's sort of by de-duplicating one gets not only the reduction of any appearance or form of memorization which might be happening, it reduces the appearance which would mean if it is happening, it reduces the actual instance, and if it's not happening but it appears to be happening, it would be reducing the appearance. But also, you get these other benefits of de-duplication which is that you have more diverse as a result of the de-duplication process, you get more reverse data. You get less redundant data. You get more balanced data. And overall, you get better,

107

cleaner, less buggy data. And those efforts are generally seen as the overall process of what is known sometimes as data hygiene.

It's interesting that you ask this question because this is something that I teach in my machine learning classes, and the reason why I put a little bit more effort into this is because I think data hygiene is such an important part of the overall machine learning process.

There are many in academia who spend lease time on data hygiene and they would wish to assume that data always comes clean, but in practice, it doesn't. And so data hygiene, things like de-duplication, processing of the data to remove bad stuff or irrelevant stuff or redundant stuff or stuff that makes it imbalanced or the duplications, all of those things are a really important process to the overall machine learning goal which is to produce a model that truly generalizes, that addresses and can deal with new situations, new problems, new scenarios, new inputs that the model has never before seen.

BY ATTORNEY MILLER:

Q Could you turn to Paragraph 10(a) of your rebuttal report, please, which is on Page 7?

108

A Did you say Paragraph 10 on Page 7?

Q Yes.

A Okay.

Q 10(a) in particular. Are you there?

A Yeah, Page 7, Paragraph 10, yes, I'm there.

Q And do you see in bold type you state there, "The widely accepted goal of LLM training including Anthropic's is generalization, learning statistical patterns in language that transcend the LLM's training data."

Do you see that?

A Yes.

Q What does it mean to -- withdrawn.

Do you stand by that statement?

A Yes.

Q What does it mean to "transcend the LLM's training data"?

A That's also a very good question. So that is a way of talking about forms of generalization.

One way of looking at it could mean what I've already said which is that the model can do things based on inputs that it has never before seen in its training data.

For example, the training data itself, if

129

include things like Dolma, which is an open source tool, and other open source tools so it could be the case that they have. I would be surprised if there was no engineer that has never adopted these tools or at least tested them, again assuming they were appropriately licensed, but I don't know for certain precisely which of these tools, which engineer who works or has worked at Anthropic has ever tested, tried out or possibly adopted.

BY ATTORNEY MILLER:

Q In Paragraph 113, as you mentioned, you cite to Anthropic's -- excuse me. You cite to the deposition testimony of Anthropic witnesses; is that right?

A So in Paragraph 113, Paragraph 113 has a relatively large number of footnotes, some of which refer to -- give me a moment -- various different notes and interrogatories and depositions including some from Ben Mann as well as other related documents, for example, a document that explains the concept of what a bag of words, set of features is.

Q Did you do anything to confirm whether Anthropic's witnesses in the deposition testimony you cite in this paragraph have accurately

130

described Anthropic's data cleaning and processing processes?

ATTORNEY GOOCH: Form.

THE WITNESS: Doing such would not have affected my conclusions in any way, and the reason is because of my background in machine learning and artificial intelligence. And my background in artificial intelligence and machine learning that's relevant to me coming to the conclusion that it wouldn't have affected my conclusions in any way is that I know that a lot of the things that are mentioned in Paragraph 113 or things like the open source tools mentioned in Paragraph 112 or just describing the general processes of data cleaning and data preprocessing in Paragraph 111, I just know that those are a necessary element of producing a functional resulting machine learning model.

And given that in this particular case, I know that Anthropic's models are functional general purpose models that are capable of doing an enumerable number of things, it just doesn't seem reasonable to me to think that there would be any chance that they could possibly have achieved what they've achieved without putting significant

131

efforts, significant -- quite significant efforts into data cleaning and data preprocessing.

BY ATTORNEY MILLER:

Q That didn't answer my question at all --

A Apologies.

Q -- so I'll ask you to please answer my question.

Did you do anything this confirm whether Anthropic's witnesses in the deposition testimony you cite in this paragraph have accurately described Anthropic's data cleaning and processing processes?

ATTORNEY GOOCH: Objection. Form. Asked and answered.

THE WITNESS: Yeah, I apologize if that seems like I didn't answer your question, but I want to make sure that I'm providing good and accurate answers.

So my answer is that based on my experience in machine learning and artificial intelligence over many decades and having experienced what it takes in order to produce one of these models, one of these kinds of models, a functional general purpose machine learning model of a variety of different sorts, I know that what

132

is necessary to do that is to have good clean data so that you don't experience the phenomenon of garbage in, garbage out and as a result of that, and a result of my conclusion that it's not reasonable to not do these things, the answer to your question which is -- which is what I was saying a little bit earlier is that I did not -- I do not believe it is necessary to do that because it would not have changed my conclusion.

It wouldn't have been anything that would have provided new information and it wouldn't have ended up changing anything that I concluded in any of these three reports.

And, again, I want to apologize because if it seems like I'm not answering your question I'm happy to -- to try again if you're interested.

BY ATTORNEY MILLER:

Q Did Anthropic filter data for Claude 2 the same way it did for Claude 3.5?

A Those are -- may I ask precisely the years and the dates when those models were active and developed?

Q Do you know when -- the years when those models were active and developed?

ATTORNEY GOOCH: Form.

137

many different actual meanings at multiple levels of granularity.

So generally, the notion of concept along with the notion of granularity, along with the notion of the degree of semanticity, namely is it monosemantic, polysemantic, and if it's polysemantic, then how polysemantic is it. All of those things are things that are fairly regularly discussed, not just in the sparse autoencoder literature which is a very small subpart of the neural network literature which is a subpart of the machine learning literature, but it's generally discussed fairly widely.

And so in that particular sentence that you pointed out, the last sentence of Paragraph 116, that sentence is referring specifically to the idea of that redundancy and imbalance is bad. The sentence reads, "No single important concept or linguistic structure should be appreciably overrepresented."

So the idea of overrepresentation is the same idea as duplication. If something is duplicated many times, it's overrepresented. There's too much representation of that concept because -- and we -- the reason why we can say too

138

much it's because it ends up hurting the model.

The example of fruit that we were talking about earlier when we had the duplications of the banana. We had the good -- the two pictures of a banana that were still good and edible and the two pictures of the banana that were bad and should be taken off the shelf. So that would be an example of a concept namely in that particular case at that particular granularity. It's the concept of a fruit. In particular the concept of a banana, and it's not infeasible that in a neural model there could be learned a unit that corresponds to that concept of the banana. It could also be the case that that unit corresponds to aspects of parts of a banana like the concept of yellowness which normally is -- I mean, there are other kinds of bananas, there are red bananas, but the concept of yellowness would normally be something that would be associated with a banana.

So, therefore, it's completely fair in my mind to use the word "concept" when we're referring to various different aspects of machine learning, especially in the case when one -- when one wants to express notions of balance, when one wants to express notions of lack of

139

overrepresentation. When one wants to express notions of diversity and when one wants also to express notions of a dataset not having duplicates in it because, as we've established earlier, duplication generally makes the model perform worse.

So these are all various different ways of referring to certain -- certain things that are fairly similar to each other. But I do want to add that actually that's another good question and I really mean that.

Q In Paragraph 117, you state, "Anthropic employs substantial data filtering and de-duplication techniques to maximize its models ability to generalize beyond their pretraining data and to minimize memorization."

Do you see that?

A Sorry. Give me one moment. I'm sorry.

I see that, yes.

Q And then you have a footnote and you cite the deposition testimony of Dawn Drain.

Do you see that?

A This is Footnote 230 on Page 79?

Q Yes.

A I see that, yes.

140

Q Did you review any of the code for Anthropic's de-duplication processes?

ATTORNEY GOOCH: Form.

THE WITNESS: So as we were talking about earlier and the whole notion of de-duplication is this idea of data cleaning, and in order to produce a machine learning model that functions well, I know that data cleaning, data processing and data de-duplication is absolutely necessary and so therefore as a result, I didn't feel that it would have changed my conclusions at all. In fact, I knew it wouldn't have changed my conclusions at all because these are things that I know from my experience in machine learning are necessary to produce a functional model and so -- because I want to make sure that I -- that I properly answer your question.

So I felt I didn't need to review any specific de-duplication code because it wouldn't have changed my outcomes, my conclusions and so as a result I did not.

BY ATTORNEY MILLER:

Q Okay. Did you compare any of Anthropic's pre de-duped datasets to the purportedly post de-duped datasets to see what the effect of

141

Anthropic's de-duplication processes was?
ATTORNEY GOOCH: Form.
THE WITNESS: In my experience working for -- in machine learning and artificial intelligence for many, many decades, I have experienced firsthand the effect of redundancy removal or diversity enhancement of datasets in many, many different contexts. This includes examples in computer vision. It includes examples in machine -- statistical machine translation. It includes examples in predomics and genetics which are two branches of computational biology. It includes examples of digital pathology which is essentially the problem of trying to diagnose cancer images, H&E stained slides trying to detect cancer and also in many, many other datasets, for example, tabular data, things like housing, price datasets, machine tool wear datasets, financial datasets.
I have looked at many, many, many different kinds of data and in all of those cases over the years, I have found that the effects of de-duplication are highly beneficial. And in general, the effects of redundancy removal and diversity enhancements, the effects of all of that

142

is highly beneficial. That in my mind is a given. It's something that one does when one is developing a machine learning process. One needs to make sure you have good data. One needs to essentially process and analyze and clean, de-dup and --
BY ATTORNEY MILLER:
Q Professor, excuse me. But you're not answering my question.
ATTORNEY GOOCH: Why don't you let him finish.
ATTORNEY MILLER: Because he's not answering the question.
ATTORNEY GOOCH: Why don't you let him finish.
ATTORNEY MILLER: Because he's not answering the question.
BY ATTORNEY MILLER:
Q I will ask the question again which is a simple yes-or-no question. Did you compare any of Anthropic's pre de-duped datasets to the purportedly post de-duped datasets to see what the effect of Anthropic de-duplication processes was?
ATTORNEY GOOCH: Professor, why don't you finish your last answer to the question and then

143

Corey can ask his next question and you can answer that one, but don't feel like you can't finish your question or answer. That's what you're here for.
THE WITNESS: First of all, again, I want to apologize if it seems like I'm not answering your question. But what I honestly am wanting to do is give you a complete, clear and accurate answer to your question and I will come back to that specific point, but I think in order to answer that question accurately, it's necessary to provide context, and so I'm just providing the context in order to finally answer your question. But I will get back to that.
So -- so -- so continuing where -- where I was before, and I'll restate your question just so that you know that I'm thinking about your question. So your question is: Did I compare the effects of the de-duplication process both before and after de-duplication on the models performance; is that correct?
BY ATTORNEY MILLER:
Q Not on the model's performance. Did you look at the datasets both before the de-duplication is run and after de-duplication was

144

run to see what was removed during the de-duplication process?
A Okay. To see --
ATTORNEY GOOCH: Same -- real quick. Same objections.
THE WITNESS: Okay. To see what was removed from the -- Okay. So that question is in my mind.
So continuing what I was saying a little bit before, that -- and I'll just summarize. To my understanding of, and my experience, my firsthand experience, is that one cannot get a functional model unless one does successful proper de-duplication of data. And so it would not have changed my analysis and my conclusions and my opinions if I had looked at exactly what, in the original raw, unclean dataset was removed because of my experience in machine learning and data science and artificial intelligence.
It could not in my mind have been possible to produce a powerful general purpose model that's so good at so many different things if it wasn't the case that proper de-duplication wasn't done because the evidence to me in my experience is that it was done, the strong evidence. And, like

169

Plaintiffs' copyrighted works for training its LLMs is transformative as transformative is used in the context of fair use law?

ATTORNEY GOOCH: Why don't you point him to his section in his report, Corey?

THE WITNESS: Yeah, I would be very happy to explain a particular section in the report. I am not a lawyer, and so the particular legal definition of transformativeness is not something that I can formally express any expertise in.

I can express expertise in other things, like the effect of duplication -- the adverse effect of duplication on the performance of machine learning models. But in this particular case, it's -- you know, as a -- as a nonlawyer, I -- I can't really express those notions or have an opinion on those notions.

BY ATTORNEY MILLER:

Q On the page right in front of you, do you see there's a big header "Roman VI," and it states, "The transformative nature of generative AI large language models"?

Do you see that?

A Yeah.

Q When you use the word "transformative"

170

there, are you using it to mean the legal concept of transformative as that concept has a meaning in fair use law?

A This is used in its -- in its general sense, and it includes a variety of different sentences. And what the text in the document clarifies is that there are -- it may be in this document. It may be in one of the other documents. It might take me a moment.

But there are instances in these documents where I talk about the various different ways that transformativeness is used.

Q And which way are you using it here?

ATTORNEY GOOCH: And you should feel free to review your report. He's asking you questions about your report, and while he doesn't want to direct you to specific things, Paragraphs 288 through 291.

ATTORNEY MILLER: I'll ask you not to coach the witness, please.

ATTORNEY GOOCH: I'm not coaching the witness. You're asking questions about 600 pages of reports and refusing to point him to any one. This isn't a memory test.

THE WITNESS: So Section 6, the title is

171

just the general notion. And if you look at Section 6A, for example, it talks about how LLM training transforms text into abstract statistical relationships.

So this report is comprehensive and it's meant to convey the general sense in which that word might be used. And from counsel, I came to an understanding of -- from -- as written in Paragraph 286, I believe the definition that you're referring to. But if you look at the same section, Subsection A of that section, for example, it talks about -- in Paragraph 293, it says, "Text data used for LLM training is transformed through heavy computational processing."

It's worth noting that when we say "LLM," what that really means is a transformer. And a transformer is a particular architecture of neural network that includes many, many layers of multiheaded tension layers, followed by multilayer perception layers, various different other forms of normalization layers, as well as a variety of other architecture features, neural architectural features.

And the reason why it's called a

172

transformer is because of its transformative abilities. And I came to know that term certainly long before this case began, and I came to know that term and that term in general, "transformer," and the fact that the transformer, despite its very name is transformative, has the ability to transform.

And the community, thousands of people, tens of thousands of researchers and scientists around the world use this notion of a transformer and how the transformer transforms information and data into new -- generally, in the context of language, statistical abstract, semantic, linguistic forms that, as we were referring to earlier, can transcend some of the information that exists in the original data.

And it's interesting that you say this, because it gets back to this notion of transcend -- transcending the original data. The transformer, arguably, is transformative in its more informal sense, because it can do -- it exhibits behaviors that transcend the behavior of the original training data.

So this is a very -- this is meant to be a complete section, not just a specific narrow

173

section about this idea of transformations.
BY ATTORNEY MILLER:
Q You're not suggesting that early technology that uses a transformer architecture is necessarily transformative for the purposes of fair use law, are you?
ATTORNEY GOOCH: Form.
THE WITNESS: I am not in the position to comment on that or have an opinion on that because I am not a lawyer. Because when it comes to notions of fair use law, I'm not a lawyer. I'm a computer scientist. I'm a machine learning and artificial intelligence scientist. And I can clarify an answer to your questions faithfully about specifics associated with the technology behind a transformer and the degree to which and how it's reasonable to use the English word "transform" and "transformation" and "transformativeness" as -- in English to refer to the remarkable capabilities of these models. And they truly are remarkable.
And it may even be the case that -- and this is -- this is something that is very interesting. It may even be the case that the English language doesn't have sufficient

174

superlatives to refer to how remarkable these models and the capabilities that these models have and how they transcend the original training data. It is truly amazing.
BY ATTORNEY MILLER:
Q So even though in your report you cite the legal definition of transformative and then you talk about how LLMs are transformative, you're not offering an opinion that LLMs meet the legal definition of transformative?
ATTORNEY GOOCH: Objection. Misstates the testimony. Inconsistent with the facts.
THE WITNESS: So in Paragraph 293, which is in Section 6, that has that header that you referred to earlier, Section 6 on Page 195, Paragraph 293 on Page 179 starts out by saying, "Text data used for LLM training is transformed."
So that is not -- that use of the word "transformed" is using it in the general sense that I and many other researchers in the machine learning community would use that word without expressing an opinion on anything having to do with fair use or what other -- whatever other legal definitions might be relevant to this particular conversation.

175

BY ATTORNEY MILLER:
Q If Anthropic excluded all copies of the works in suit from Claude's training data, what effect would that have on Claude's performance?
ATTORNEY GOOCH: Form.
THE WITNESS: There's several aspects that I need to go through to answer that question. But may I repeat your question so that you can be sure that I heard your question and I'm working to answer your question?
BY ATTORNEY MILLER:
Q Yes.
A So your question was, if Anthropic excluded all copies of the works in suit from their training data, would that have an effect of -- on their model; is that right?
Q Yes. I can repeat the question, if it would be helpful.
A Yeah, if you don't mind repeating the question.
Q If Anthropic excluded all copies of the works in suit from Claude's training data, what effect would that have on Claude's performance?
A Okay. There's two sides to this answer. What effect would it have? Number one,

176

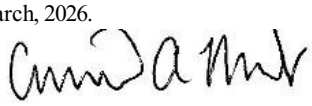
the first part of the question is if Anthropic excluded all copies. So this relates back to something that we were talking about a little bit earlier with regard to what would it mean and is it a well defined problem to exclude all copies, because one would need to identify precisely what is a copy. One would need to have the definitive copy. And then with that, one would also need, if one had that information, to know the idea and ways in which one can -- there can be variations of a particular copy. There can be missing sections. There can be misspellings. There can be extra punctuation. There can be inserted words. There can be deleted words. There can be permuted words. There can be reversals of words. There can be reversals, reversals of the entire task. There could be ASCII-level translations of the original ASCII characters, various different encodings. And then also to all of those things that I just mentioned, there can also be fragments or portions or pieces.
And there can be various distortions in all of those things; reversals, misspellings, punctuations, inverses, deletions, substitutions, associated with each one of those fragments, which

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