EXHIBIT 138



HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY

Transcript of Russ Altman, Ph.D., M.D.

Date: February 27, 2026
Case: Concord Music Group -v- Anthropic

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

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
DEPOSITION of
RUSS ALTMAN, Ph.D., M.D.
Palo Alto, California 94306
Friday, February 27, 2026

Reported Stenographically by:
MARY J. GOFF
CSR No. 13427
WA CSR No. 21030779
Job No. 620592
PAGES 1-309

2

HIGHLY CONFIDENTIAL--ATTORNEYS' EYES ONLY

VIDEOCONFERENCED and VIDEO-RECORDED DEPOSITION of RUSS ALTMAN, Ph.D., M.D., Volume I, taken on behalf of Plaintiff Music Publishers, Oppenheim + Zebrak, LLP, at WilmerHale, 2600 El Camino Real, Suite 400, Palo Alto, California 94306, beginning at 9:05 a.m. and ending at 6:37 p.m., on Friday, February 27, 2026, before MARY J. GOFF, California Certified Shorthand Reporter No. 13427 and WA CSR No. 21030779.

3

APPEARANCES:

For Plaintiff Publishers
Oppenheim + Zebrak, LLP
BY: COREY MILLER
Attorneys at Law
4530 Wisconsin Avenue
5th Floor
Washington, D.C. 20016
202-957-5043
corey@oandzlaw.com

For Defendant Anthropic PBC
WilmerHale
BY: ARI HOLTZBLATT
BECKY MAZ (remote)
Attorney at Law
2100 Pennsylvania Avenue NW
Washington, D.C. 20037
ari.holtzblatt@wilmerhale.com
202-663-6964

ALSO PRESENT: Natalie Naugle, in-house counsel Anthropic
Videographer: Philip Astor

4

INDEX

WITNESS	EXAMINATION
RUSS ALTMAN, Ph.D., M.D.	
Volume I	
ATTORNEY MILLER	7, 302
ATTORNEY HOLTZBLATT	284, 304

NUMBER	DESCRIPTION	PAGE
Exhibit 291	Expert Report of Russ Altman 12-19-2025	11
Exhibit 292	Expert Rebuttal Report of Russ Altman, 1-23-2026	13
Exhibit 293	Expert Reply Report of Russ Altman, 2-13-2026	14
Exhibit 294	The Adolescence of Technology article, January 20226, Amodei Tab 2	64
Exhibit 295	Spirit of Asilomar EnTreaty 2025.3.1 Tab 3	71
Exhibit 296	Hacker Used Anthropic's Claude to Steal Sensitive Mexican Data 2-25-26, Mercury News, Tab 4	80
Exhibit 297	LLMs in Healthcare and Medical Applications: A Review Maity, Jyoti Saikia article Pages 1-25, Tab 5	97

Pages 101-108

101

A It is context-dependent as -- as --
THE DEPONENT: Sorry.
THE VIDEOGRAPHER: Sorry. Can you just clip it on your right lapel?
THE DEPONENT: Right lapel?
THE VIDEOGRAPHER: Yes.
THE DEPONENT: I'm now putting my microphone on my right lapel. And now I'm going to have to ask you to repeat the question.
Q (BY ATTORNEY MILLER) What does human oversight verification and auditing require?
A So that is a context-specific question. The answer about appropriate oversight, verification, and auditing -- excuse me. The appropriate human oversight, verification, and auditing depends on the situation. In that sentence I was referring to myself as a user of these models and tools, and so as I am a user in different roles, I apply different oversight, verification, and auditing approaches.
Q All right. Could you turn to paragraph 54 of the document you have in front of you, please?
A This is paragraph 54 of my -- of my expert reply report.
Q Yes.

102

A Yes, I have it.
Q And in paragraph 54, you state that, "The value of any individual class of training material, including song lyrics" -- oops. That's what it says, but that's not what I meant to be -- withdrawn. Ah. Turn to paragraph 53.
A Okay.
Q I'm sorry.
A I'm now looking at 53.
Q This. And do you see the third sentence in paragraph 53 states that:
"LLMs derive their capabilities from training on and exposure to vast amounts of heterogeneous text pulled from diverse sources, reflecting the full range of human language use, including creative, metaphorical, and nonstandard expressions"?
A Yes.
Q Do you consider song lyrics to be creative, metaphorical, and nonstandard expressions?
A Lyrics certainly can be that, yes.
Q Now I want you to turn to paragraph 54.
A Okay.

103

Q And in paragraph 54, you state that:
"The value of any individual class of training material, including song lyrics, cannot be evaluated in isolation, but must be viewed as part of a broader ecosystem of language data that collectively supports generalization and responsible use."
Do you see that?
A Yes.
Q Is it important that Anthropic include -- let me take -- withdrawn. Let me take a step back.
You understand that the copyrighted works that are at issue in this litigation are song lyrics?
A I -- yes, I have a rough understanding that song lyrics and copyright issues are at stake in this case, yes.
Q Is it important that Anthropic include the -- the song lyrics at issue in this case in Claude's training data?
ATTORNEY HOLTZBLATT: Objection, vague.
A Could you repeat the question?

104

Q (BY ATTORNEY MILLER) Is it important that Anthropic include the song lyrics at issue in this case in Claude's training data?
A So I did not -- as you know, I did not train Claude, and I really can't that question. In general, they -- it constitutes text, but the strategies by which Anthropic decided to include or not to include things in its training are opaque to me, and I do not know them.
Q Would Claude be less effective if its training data did not include the song lyrics at issue in this case?
A Unless that's been specifically tested, that's an empirical question that I don't know the answer to.
Q Do you agree that lyrics, as text, do not constitute high-quality datasets for general-purpose LLM training, as they lack lexical diversity in factual information?
A That was a mouthful, so I'm going to ask you to repeat the question. I'll listen even more carefully.
Q Do you agree that lyrics, as text, do not constitute high-quality datasets for general-purpose LLM training, as they lack lexical diversity and

105

factual information.
A I don't have an opinion about that.
Q Do you agree that compared to other sources of text within Claude's training corpora, lyrics in general and including Plaintiffs' lyrics, do not constitute a high-quality dataset specifically for general-purpose LLM training?
A I really don't have an opinion on that as well.
Q Do you agree that the effect of Plaintiffs' lyrics on Claude's learned parameters or outputs is extremely likely to be near zero, if present at all?
A That one requires some thought. The corpus -- my understanding, without any expertise or inside knowledge about the training of Claude, is that it is trained on a huge -- the word could even "humungous" corpus, and I don't know the size of the lyric corpus that may or may not have been included in the training of Claude.
It seems to me from common sense only that it is likely to be a tiny, minuscule fraction of the text that -- that was provided to Claude, and therefore, it's quite unclear to me how critical or -- or it -- or what kind of difference it would

106

make. I cannot comment on that, but that's this -- the way I think about your question.
Q Do you agree that lyrics have minimal value to training a general-purpose LLM?
A I think that's an empirical question as well that I don't have access to the data or the experience to know the answer to.
Q Do you agree that with respect to lyrics, any informational utility they provide could be obtained from many alternative sources, so lyrics text specifically may not be imperative for the training of a general-purpose LLM?
A So you used the word "may" and "could," which -- which then changes the nature of my answer, which is: Yes, it may or could be the case.
And it would need to be empirically tested.
Q And a couple of times today you mentioned the CRISPR-GPT model that you have worked on.
Does that model use Claude in any way?
A No. It used mostly the products from OpenAI.
Q Do you know whether the OpenAI products that that CRISPR model used were trained on song lyrics?

107

A I do not know that.
Q For the CRISPR model to work, was it necessary that the training data for the LLMs on which the CRISPR model relies, included song lyrics?
A It is extremely hard to and, in fact, arguably impossible to understand how any particular piece of data that is fed in at the very beginning of the process into a network of billions or trillions of interacting variables, how that affects the output.
So I'm only saying that to say it is very difficult for me to opine on what the exact effect is. I can't say it's there, and I can't say it's not there, in the same way that every experience we all have contributes to our own personal neural network of the world called the brain.
And then connections that are made are not fully understood, but could be quite -- what's the word -- surprising and idiosyncratic.
Q In your report, you describe work being done on FDA -- excuse me -- on using LLMs in connection with FDA bioequivalence?
A Yes.
Q Is that the -- the fine-tuning you were referring to --

108

A Yes.
Q -- earlier?
Is -- does that project use Claude?
A No.
Q That uses Qwen, Q-W-E-N; is that correct?
A Let me -- if I can expand.
There are two pieces to that project. The first piece is the creation of a high-quality benchmark. As you may know, whenever you're building a system, it's actually good practice to predeclare what you expect it to be good at and, like, commit to that and then build the system and see how good you are.
So the first thing we did is we built a benchmark of questions and high-quality answers in collaboration with the FDA, and that's going to be a separate -- that's going to be a paper about this benchmark.
That benchmark can then be used for two things. It can be used for the fine-tuning, and -- but of course, you take some of that and put it aside, don't make it available to the LLM, and use it to test.
So when -- we built the benchmark, and in the benchmark we did use LLMs to help generate

Pages 129-136

129

page 3.
A Okay. Yes. Thank you. That's. That's right.
Q And if you would please turn to page 11.
A Yes.
Q In the "Discussion," I want to direct your attention to the last sentence of the last complete paragraph.
A On page 11?
Q Yes. And do you see that it says:
"We kept our grading scheme lenient because of the amount of hallucination and poor overall performance of the assessed models"?
Do you see that?
A Yes.
Q All right. Can we go back your report, please, to paragraph 57?
A Yes.
Q Beginning on page 23 and continuing on to page 24, you state that:
"Our results increased my confidence in Claude's potential medical or life science

130

applications."
Do you see that?
A Yeah.
Q So these results were all models performed poorly overall, and you had to keep your grading scheme lenient because of the amount of hallucination and poor over -- poor overall performance of the assessed model increased your confidence in Claude's potential for medical and life sciences applications?
ATTORNEY HOLTZBLATT: Objection, mischaracterizes the document.
A Could you just repeat the question?
Q (BY ATTORNEY MILLER) So these results where all models performed poorly overall and you had to keep your grading scheme lenient because of the amount of hallucination and poor overall performance of the assessed models increased your confidence in Claude's potential for medical and life sciences applications?
ATTORNEY HOLTZBLATT: Same objection.
A Yes, it did. And I can understand your question.
This had not been evaluated yet, so the very act of turning a wonder about whether an LLM

131

might be useful in pulling out this information, which I could describe to you, if you're interested, why recalling facts about biological reactions is important.
But we set up an experiment. I -- I discussed earlier the importance of benchmarks whenever you build a system. So a lot of my confidence is actually not about the performance of any individual LLM, which were all kind of uniformly not that great, but the fact that we now have a system for evaluating.
And in fact, the reviews -- it has occurred to me since you last asked me about the reviews, one of the reviews was: These models are changing all the time, and you're giving us this little snapshot, which will be of very limited use as soon as the next version of Claude or the next version of ChatGPT comes out, and so we -- we're working to figure out how to respond to that.
But that's why I didn't heavily rec -- lean on the actual performance of these. And I think the source of my confidence is that we did the experiment. We have numbers. And as we now try to do clever things to improve performance, I have a way to know when I'm getting better. And -- and

132

that comb -- I guess I'm being optimistic. I now have a system.
In engineering we say: You cannot intervene unless you can measure.
And now I can measure performance, which will allow me to intervene.
Q Starting at paragraph 58, you describe seven use cases for Claude in which Claude delivers what you call "measurable benefits in scientific discovery; is that right?
A Yes.
Q And these are the use cases you identified according to the -- in the methodology in Footnote 59 that we discussed earlier?
A That's right.
Q And the first use case beginning at paragraph 60 that you describe is "managing large multimodal datasets"; is that correct?
A Yes.
Q If Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to assist with managing large multimodal datasets?
A Could you repeat the question?
Q If Claude's training corpus did not

133

include song lyrics, would that have any effect on Claude's ability to assist with managing large multimodal -- multimodal datasets?

A So again, that -- that -- as I have said before, that's an empirical question. You would have to do an experiment.

Even the use of the word "lyrics" in your question could be interpreted multiple ways because there's written lyrics. But in a multimodal setting, it might be learning the lyrics from watching a performance or -- so "it's not tested" is my answer. And "lyrics" would have to be defined more clearly in order -- even to do that experiment.

Q But you're not offering any opinion that Claude's training corpus needed to include song lyrics in -- in either written or any other form in order for Claude to be able to assist with managing large multimodal datasets?

A I said in the report that "a diverse nonhomogeneous corpus is absolutely critical."

The degree to which it contributes to that, it's contributing. The magnitude needs to be tested separately.

Q In paragraph 61, you describe a potential use of LLMs as "interpreting health insurance

134

policies"?

A Yes.

Q Have you ever used an LLM for this purpose?

A As I said at the very bottom, I did have a small project with a large health insurance provider for this problem, but it was pre-LLM.

Q Okay. Are you aware of any published research showing that LLMs can be used to interpret health insurance policies?

A I'm not aware of any literature on that. I'm not sure the insurance policies would publish that, but I'm not aware -- to answer your question, I'm not aware.

Q In paragraphs 60 and 61, you don't cite to any papers from PubMed, BioRxiv, or ArXiv, do you?

A It looks like not.

Q And in fact, in paragraphs 60 and 61, you don't cite anything, do you?

A It appears no.

Q Okay.

A I -- these observations were made from my general impressions from attending conferences, reading the literature. It's -- it falls within -- in this case this falls well within my kind of

135

professional activities, and so this was a summary statement by me based on my expertise, not supported by citations in this case.

Q All right. The second use case or category of use cases you discuss beginning at paragraph 62 is "assisting with the development of research protocols."

Do you see that?

A Yes.

Q Have you used Claude to assist with the development of research protocols?

A The -- CRISPR-GPT, as I said before, was -- was an effort using the OpenAI tools.

Biomni is another tool generated by a colleague of mine. But I know that students in my lab have tested Biomni, and Biomni is at least partially based on Claude.

Q Have you ever used Biomni?

A I have not used Biomni. My knowledge of Biomni, as for many technical low-level tools, is usually based on deep conversations with my students about their experiences and their successes and failures.

Q In paragraph 64, you describe a study in which Claude enabled -- in paragraph 64, on -- let

136

me start over.

In paragraph 64, on -- I want to -- on the top of page 27, you describe a study using Biomni. And in middle of that paragraph, you state:

"Throughout these tests, Claude enabled Biomni, maintained all of the relevant work data, including code, figures, and logs so that it could be later audited and reproduced."

Do you see that?

A Yes.

Q Do you know whether that data was ever, in fact, audited?

A There is a published paper describing this work. It's -- it's the subject of Footnote 72, and that is where I would look.

So the answer to your question is: No, I don't know. And where I would look would be -- I would reread a citation, 72. I have also heard a talk about this work, and I just don't recall if they used these logs to audit and reproduce.

Q And I take it then that you don't know what results, if any, an audit of that data would have revealed?

Pages 141-144

141

not clear to me that I would do a PubMed-type query to Claude. It's a conversational chatbot. It doesn't kind of represent itself as a source of literature.

So I could do that and it might be okay, but I would instruct my students to be a little bit more like clever in their us of the -- to take advantage of what it's good at and take advantage of what the other resources are good at.

Q In order for Claude to answer a question -- "Who is doing research in my field on this topic -- on a particular topic?" -- does Claude need to have access to the underlying research papers?

A It almost certainly has access to PubMed, as -- as I do. It also access to the websites of the investigators -- or most academics have a little website with summaries of their research projects and their -- and their lab members and what not.

So I -- I think that it would help do it better job if it had some knowledge of the publications that are associated with those reviewers.

But that's not the only data source. It's one of the valuable ones.

Q If Claude's training corpus did not

142

include song lyrics, would that have any effect on Claude's ability to assist with scientific literature reviews?

A I don't know that that's been tested. It does provide diverse and nonhomogeneous text that gives it a better understanding, however incremental, of -- of the use of language and how to interpret language -- sentences and text.

Q And similarly, with the previous use case we were discussing, assisting with the development of research protocols, if Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to assist with the development of research protocols?

A I don't think it's been tested, and so I don't know, except through the fact that it's a nonhomogeneous -- it is a nonhomogeneous and diverse dataset, which in general, to my understanding, improves the performance of LLMs.

Q In paragraph 66, you discuss FutureHouse; is that right?

A Yes. That's correct.

Q And you have in paragraph -- in the first sentence -- withdrawn.

Paragraph 66 has four footnotes with

143

citations? Do --

A Yes.

Q -- you see that?

And the first one is Footnote 75, correct?

A Yes.

Q And it points -- Footnote 75 cites three sources; is that right?

A Let me just read through it. Give me a moment, please.

I have read the footnote. And it may be that it refers to four, but it's at least three. There's one at the very bottom.

Q Okay. You're correct. I missed one, but it doesn't change -- so let's go through them.

The first one is an article from Anthropic's website; is that correct?

A Yes.

Q The second is FutureHouse's website, right?

A Yes.

Q The third is an article from TechCrunch?

A That's right.

Q And then the fourth is an article from FutureHouse's website, correct?

A Yeah. That's right.

144

Q And then Footnote 76 cites an article from Anthropic's website, right?

A Yeah, it's actually the same -- it's the first -- it's the first citation in 75. I think it's the same one.

Q And Footnote 77 cites that same article from Anthropic's website?

A Yeah. These are just different little quotes that we pulled out.

Q And Footnote 78, which is on the next page, cites the same article from FutureHouse's website that was cited in Footnote 75?

A That's right.

Q Okay. In the second sentence of paragraph 66, you say that:

"In reviewing FutureHouse's work, I found they developed four specialized agents to address different aspects of workflow," and then you identified the four specialized agents --

A Yes.

Q -- right?

Did you actually review any of FutureHouse's work?

Pages 153-156

153

used Bluenote to accomplish, do you?
A I don't know that.
Q And you don't know what Eledon Pharmaceuticals used Bluenote to accomplish, do you?
A I do not know precisely what either of them did. I had a sense of it from the beginning of that paragraph, but I did not confirm it. I relied on the contents of the report.
Q Okay. You can -- if Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to streamline regulatory processes?
A I think that's a -- a matter of empiricism. You would have to check it empirically. I would only say that lyrics would represent diverse and nonhomogeneous data that is the type of data that is required for building a high-quality LLM.
Q And you say that "lyrics represent diverse and nonhomogeneous data," what do you mean?
A So diverse and nonhomogeneous text. You're training an LLM. And essentially an LLM is a very large statistical model of language. And in order to have a good statistical model, you need to have lots and lots -- literally millions, billions of examples of text -- from which it can estimate a

154

probability distribution.
And if you choose -- but it -- just having billions of sentences in the ridiculous extreme, it would be one sentence repeated billions of times -- would lead to terrible LLM.
And there's a sense -- and I think it's known empirically from building LLMs that the more different the text is using words in different contexts, saying different kinds of statements with sentences of different lengths and on topics of a wide variety leads to a more robust statistical model of what words mean and how they work together to create sentences. And -- and that's critical for the LLM.
And so when I say "diverse and nonhomogeneous data or text," I mean fulfilling that criteria of not being entirely repetitive and offering no kind of incremental advantage over just repeating the same sentence millions of times.
Q How are lyrics diverse and nonhomogeneous?
A So diversity obviously is something you don't evaluate on a single item. You look at a corpus. Oh, and when I say "corpus," I mean a collection of text. And that is how you determine diversity.

155

So you look at the corpus -- in the example that I gave, the -- the extreme example, but perhaps useful for our conversation, if I'm repeating the same sentence a million times, that sentence doesn't have diversity or lack of diversity in it. It is the corpus of a million mentions of that sentence that would lead me to the conclusion: This is not diverse.
On the other hand, if I have a collection of data that included lyrics -- and I have never had that -- but if I did, then I would not just say: Is this lyric diverse?
I would say: As I look at the other text, which might include, for example, all of Dr. Altman's published papers, is that diverse?
And then you might be able to come up even with quantitative metrics of diversity that would give you confidence that as you're trying to estimate this probability distribution in the LLM, you're do -- going to get a good probability dist -- distribution that is robust for many situations.
Q So when you say -- excuse me.
When you said that "lyrics are diverse and nonhomogeneous," in what sense do you mean that they are diverse and nonhomogeneous?

156

ATTORNEY HOLTZBLATT: Objection, mischaracterizes prior testimony.
A When I -- when I used the shorthand of "lyrics are diverse and nonhomogeneous," having given you this explanation, now I can say: In the context of lots of other text that is different from those lyrics, they may well represent some useful diversity to improve the quality of the statistical model that the LLM is being trained -- that is being trained.
So that's what I mean by "diverse and nonhomogeneous" compared to the other elements of the corpus.
Q (BY ATTORNEY MILLER) In paragraph 70 in your report, you state:
"When used as a writing assistant, LLMs including Claude, can offer similar benefits to creative and academic writing workflows."
Do you see that?
A Yes.
Q Have you ever used Claude for a creative writing workflow?
A Let me start out by saying that I'm

Pages 165-168

165

Q If Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to facilitate drug development and discovery?

A I think that's primarily empirical question about the impact, which you would have to test.

However, in the context of a larger corpus containing many other different documents, the introduction of lyrics could be considered the introduction of diversity and nonhomogeneity into the corpus, which in general, leads to a more robust LLM.

Q All right. Let's turn to -- the sixth use case you identify starting on page 33 is "augmenting healthcare delivery and patient care"?

A Yes.

Q And you begin by referring -- by -- excuse me. You begin by stating that:

"LLMs have been used for knowledge extraction such as protocol and literature synthesis";

is that right?

A Yes.

Q Is -- the "literature synthesis" you

166

reference here, is this different from the literature review we were discussing earlier?

A Yes.

Q What is literature synthesis, as you used it here?

A Literature synthesis here is -- give me a moment to refresh, although -- I'm -- I'm just going to take a moment.

Literature synthesis is more focused than literature review. This is where you have several papers on the same topic and you want to combine their findings for a kind of -- to generate -- in order to generate a summary of what they say as -- as a whole instead of individual studies.

So you would only synthesize things that were similar and comparable so that you could come up with kind of a consensus statement of what they showed.

Q And what is protocol synthesis?

A Protocol synthesis is a similar idea, except now you're looking at protocols that are meant to do some task.

Like: What is your protocol for treating a patient or making a diagnosis?

And you can imagine that there would be

167

several publications from specialists about how they do it.

And you might want to say: Okay, let's generate a summary of what is the current best practices among these several experts who -- who have published a protocol?

The protocols may not be identically the same. That's the challenge. But it may be that when you extract them out into a synthesis, they share lots of features that give you a good idea of the general best practices of people who are expert at these protocols.

Q If Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to augment healthcare delivery in patient care?

A I don't know. I think that's an empirical question that one would have to test. I do think that lyrics provide -- in the context of the entire training corpus used to create an LLM, potentially provide diverse and nonhomogeneous compared to the other elements of the training corpus; and therefore, they could contribute to a better LLM.

And when I say "a better LLM," I mean a better LLM that better approximates a statistical --

168

a statistical distribution of how language can and is used.

Q When you say "that's an empirical question that one would have to test," you haven't tested it?

A I have not.

Q And I think earlier you suggested that there may be sort of formal statistical measures of diversity within or among datasets?

A I could --

ATTORNEY HOLTZBLATT: Objection, mischaracterizes prior testimony.

A -- I can imagine statistical measures of diversity. I know that they exist in fields that I work in, very disparate fields, but sometimes those measures are quite general purpose.

Q (BY ATTORNEY MILLER) You have not attempted to calculate any statistical measures of diversity with regard to any of the training corpora used for Claude, have you?

A No. We did use these ideas in our fine-tuning of Qwen that I have discussed in response to other questions because we knew that the diversity was important. And we -- we introduce diversity, and it's possible that my student

Pages 173-176

173

So the seventh use case you identify starting on page 36 is "supporting effective time management" --

A Yes.

Q -- is that correct?

A Yes.

Q If Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to the support effective time management?

A I don't know. And I consider that an empirical question that would -- could be answered with experiments -- with controlled experiments.

I can only say that LLMs, to the extent that they require a diverse and nonhomogeneous corpus of documents, would benefit from lyrics marginally if those lyrics, compared to the rest of the corpus, were different -- that's what I mean by "diverse" -- and different in different ways, which is what I mean by "nonhomogeneous."

Q I would like to direct your attention to paragraph 84 where you discuss "supervision of patients who are unable to care for themselves such as elders or children"?

A Yes.

174

Q And in this paragraph you don't describe anything beyond prototypes, do you?

A Let me read the paragraph to make sure that that's true.

That's right. I am aware -- I am exposed to ambient intelligence academic work quite a bit. It's an area of interest at Stanford.

And I have -- and many of my colleagues are experts in this area, so I didn't refer to specific systems. I kind of integrated my knowledge -- my preexisting knowledge of the kinds of things that are happening in ambient intelligence and the sources for my enthusiasm about them.

Q Are you aware of any ambient intelligence systems that have actually progressed to the stage where they're available for use?

A Yes.

Q Which ones?

A The most impactful ambient intelligence systems so far in medicine have been the ones that listen to an exchange between a physician and a patient in the room and basically take notes.

This has been so far the killer app in medicine, and we hope there will be better killer apps. And "killer" is not a good word to use in the context of medicine. I want to acknowledge that. It's a phrase.

But this has been a remarkable early win for LLMs.

Q Do you consider yourself an expert on AI safety generally?

A In academics we like to be specific, so to say things are general is very difficult.

I think I am -- I have quite a bit of knowledge about AI safety in the areas within which I work, which is medicine, biotechnology. But -- and scientific research.

And I think it -- it trails off in terms of my familiarity with the full academic literature. It kind of trails off as -- as expected.

So with regard to -- when we're talking about safety and when we're talking about principles and especially if those principles are present and relevant to the areas that I do know, then I feel on solid ground even talking to a physicist or a chemist or some -- something -- or a social scientist outside of my expertise because we're talking at a high level about shared challenges.

But when those same experts go to detailed challenges of safety in their field, that's when I would have to say: No, that is not my expertise.

So there is an element of generalizability to my expertise, which is founded in my actual specific expertise in those specific areas that we -- that I mentioned.

Q Do you have any experience working on developing safeguards or guardrails for AI models?

A So I feel that the answer is yes. And let me explain that. When we were building our LLM for the FDA, in that case and in that specific context, the risks are giving a reviewer at the FDA bad information, and so the entire reason that we did a training of a special model that was domain specific and narrow was to try to mitigate.

And so in that -- I know that -- I re -- an FDA reviewer getting bad information might not consider -- be considered a safeguard, but I think -- I can make the case that they are reviewing data that might lead to the approval of a drug that might be taken by millions of Americans or people globally.

And so even in that small example I consider myself building safeguards, and that's why I say "Yes."

Q When you were building the model for use

Pages 257-260

257

in helping it generate responses to your queries?
A I think if -- if song lyrics were included in Claude, then they contributed to Claude's performance. As I have said, when you train an LLM, you give a huge corpus of text. It builds an elaborate statistical model of language. And then when you query it, it uses that model to create output.
So -- so I do think that all of the training data in a -- in this -- in a distributed sense contributed to the output, and that is always the case modulo alignment.
Q In paragraph 115, you discuss Claude's ability to make recommendations of music in response to queries?
A Yes.
Q In Claude's training corpus did not include song lyrics, would that have any effect on Claude's ability to make music recommendations?
A I don't know. I don't -- because the LLM is a very complex statistical model and because there's multiple inputs, including lyrics, song reviews, album reviews, dis -- critic discussions of songs, indexes of songs that are available for -- for purchase, it is very hard to know the answer to

258

that question, so I don't know.
Q Are you a historian?
A No, I don't think so. No.
Q Do you have any academic training as a historian?
A Only the history classes that were taken as part of my general education for my B.A.
Q You have never taken any graduate-level history classes?
A I have not taken graduate-level history classes.
Q Have you ever published any hysterical research?
A No.
Q Prior to this -- your work as an expert in this case, had you ever conducted any research on the history of technological development?
A Could --
ATTORNEY HOLTZBLATT: Objection, assumes facts not in evidence.
Go ahead.
A -- could -- could you ask it again, please?
Q (BY ATTORNEY MILLER) Yeah.
Prior to your retention as an expert in

259

this case, had you ever conducted any research on the history of technological development?
A The only reason I'm pausing is my field of biomedical informatics is a relatively new field. We are aware of that. It's spun out of computer science.
And so I have written about that. And I -- but I have written it as a participant, and I don't think that would be bona fide historical analysis.
I wouldn't tell my -- I wouldn't tell my colleagues it was a bona fide historical analysis, but it is a history of the technological gen -- development of this subspecialty of computer science which is looking intensely at biological and medical data. So in that sense I have published a little bit.
Q Prior to your retention in this case, had you ever conducted any research on the history of the use of technology in education?
A No, I don't think I have done that.
Q In your rebuttal report you cite several historical examples of the introduction of new technologies that had significant risks associated with them?

260

A Yes. I believe I did, yes.
Q And the -- the first one is the "rollout of electricity"?
A Yes.
Q And the second is "the discovery of x-rays and radiation"?
A Can you give me the paragraph just so I can refresh what I said in those areas?
Q Sure. So beginning at paragraph 25 --
A Thank you.
Q -- and continuing through the next several paragraphs.
A Yes. The choice of those examples was made to make them easily accessible and understandable as examples.
I, of course, have examples of technological rollouts in my fields, but they're very specific. They're not of general interest. And they would not, if you would, hold a reader's attention. And they would be foreign.
And the reader would say: Well, I don't know. Is this true? Is this really how it went?
Relying on historical -- some historical documents and the fact that these are very accessible examples, I thought it would be useful as