EXHIBIT 74

IN THE UNITED STATES DISTRICT COURT

FOR THE SOUTHERN DISTRICT OF NEW YORK

----------x

IN RE: OPENAI, INC. : Case No.

COPYRIGHT INFRINGEMENT LITIGATION : 1:25-md-03143

----------x

HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY

Videotaped Deposition of OPENAI

By and through its Designated Representative

JOHANNES HEIDECKE

San Francisco, California

Thursday, February 12, 2026

9:13 a.m.

Reported Stenographically by:
MARY J. GOFF
CSR No. 13427
WA CSR No. 21030779
Job No. 619815
PAGES 1-267

INDEX

WITNESS		EXAMINATION
JOHANNES HEIDECKE		
Volume I		
	ATTORNEY MATCH	10
	ATTORNEY BAYLEY	--
	ATTORNEY PRITT	107
	ATTORNEY BAYER	--
	ATTORNEY NAGGAR	--

NUMBER	DESCRIPTION	PAGE
Exhibit 1	Copyright Policy OPCO_MDL_000999211-19	36
Exhibit 2	2-9-24 email from Pronin Option to include Bing snippets from blocked domains OPCO_MDL_002129144	46
Exhibit 3	Slack message, 3-2-23 tmp incorrect training data responses OPCO_MDL_001936736-40	58
Exhibit 4	Slack message, 11-29-23 Johannes OPCO_MDL_001669756	61
Exhibit 5	Antiregurgitation, Harmony 6-26-2023 OPCO_MDL_002138928-41	68

ATTORNEY BAYLEY: Edward Bayley, Keker, Van Nest & Peters on behalf of OpenAI and the witness. And with me from OpenAI is Nora Puckett. 09:14:11 09:14:14 09:14:19

THE VIDEOGRAPHER: The court reporter today is Mary Goff, representing Planet Depos. The witness will now be sworn. 09:14:22 09:14:23 09:14:23

THE COURT STENOGRAPHER: My name is Mary Goff. I am a California Certified Shorthand Reporter, CSR Number 13427. Today's proceedings are being captured by stenographic means. 09:14:23 09:14:23 09:14:23 09:14:23

JOHANNES HEIDECKE,

Being first duly sworn or affirmed to testify to the truth, the whole truth, and nothing but the truth, was examined and testified as follows: 09:14:23 09:14:23 09:14:23

EXAMINATION 09:14:23

BY ATTORNEY MATCH: 09:14:41

Q Good morning, Mr. Heidecke. 09:14:42

Is that how you pronounce it -- 09:14:44

A That's good, yes. 09:14:46

Q -- roughly? My name is Stephen Match. I'll be taking your deposition today on behalf of the News Plaintiffs. 09:14:46 09:14:48 09:14:51

So right now you understand that you're testifying in your capacity as a corporate witness for OpenAI, correct? 09:14:53 09:14:55 09:14:59

actually happen. 09:26:11
Q (BY ATTORNEY MATCH) Okay. You have -- 09:26:12
OpenAI has used post-training mitigation measures 09:26:19
for -- well, you -- actually, you tell me. 09:26:24
How long has OpenAI been using 09:26:26
post-training regurgitation prevention measures? 09:26:30
ATTORNEY BAYLEY: Objection, form. 09:26:46
A Part of this is hard for me to answer 09:26:47
because we employed human data campaigns that might 09:26:48
be relevant to this topic. The targeted work that 09:26:53
I'm familiar with started in or around March 2020 -- 09:26:59
I believe March 2023 and took a couple of months. 09:27:08
So maybe until June, July of 2023. 09:27:14
Q (BY ATTORNEY MATCH) Okay. So since 09:27:17
March -- so roughly -- a little bit less than three 09:27:19
years, roughly? 09:27:21
ATTORNEY BAYLEY: Objection, form. 09:27:23
A Under the same caveats that I made before, 09:27:26
I -- that is a correct calculation. 09:27:28
Q (BY ATTORNEY MATCH) Okay. So in those 09:27:32
almost three years, has OpenAI discovered that its 09:27:33
other mitigation measures are sufficient to prevent 09:27:35
regurgitation? 09:27:41
ATTORNEY BAYLEY: Objection, form. 09:27:42
A I think the answer to this question 09:27:46

A Yeah. Apologies. Can you repeat the question -- 09:39:40 09:39:44

Q Yes -- 09:39:45

A -- for me? 09:39:45

Q What were the post-training methods that OpenAI employed to reduce regurgitation of copyrighted training data? 09:39:45 09:39:48 09:39:53

A My understanding is there were two parallel workstreams, one on the general model behavior and one targeted at the browsing capability of the model. 09:40:01 09:40:03 09:40:06 09:40:09

Q Okay. 09:40:11

A Which one would you like me to start with? 09:40:14

Q Why don't you start with general model behavior, and then we'll move to browsing. 09:40:16 09:40:18

A Yeah. The general model behavior targeted post-training campaign from the first half of 2023 was a synthetic data campaign where we created both demonstrations that the model would learn from, as well as preference data that we train a -- a broad model -- a preference model with that can give training signal during the reinforcement learning stage. 09:40:22 09:40:26 09:40:28 09:40:35 09:40:38 09:40:43 09:40:47 09:40:50

Q Okay. So can you tell me what this demonstrations -- what these demonstrations were? 09:40:51 09:40:56

A Those demonstrations were pairs of inputs to the model that are relevant to copyright and regurgitation paired with ideal model behaviors according to our copyright policy. 09:41:03 09:41:07 09:41:11 09:41:16

Q So -- okay. 09:41:23

And what about the preference data? What are the -- what is the preference data? 09:41:27 09:41:30

A The preference data has a similar or possibly the same distribution of inputs combined with multiple completions for each input of varying quality or varying levels of adherence to the copyright policy. 09:41:34 09:41:38 09:41:45 09:41:49 09:41:54

These multiple completions are then both scored with a Likert score, which is just an integer value where -- between 1 and 7 where "higher" means better, as well as ranked from best to worst. 09:41:56 09:42:00 09:42:04 09:42:07

Q So let me just see if I'm understanding this. And correct me if I'm not. 09:42:15 09:42:18

On this approach, OpenAI would supply a prompt to the model, supply a sequence of human-written possible outputs; it would rank those outputs and it would score those outputs according to the copyright policy? 09:42:24 09:42:32 09:42:38 09:42:41 09:42:44

ATTORNEY BAYLEY: Objection, form. 09:42:48

A I think that is not fully accurate. 09:42:49

in Exhibit 1 is that we do allow the model to repeat text that is limited to text that was exactly inputted by the user already. 02:07:36 02:07:42 02:07:46

Q Okay. Thank you. 02:07:49

Can you turn to Slide 30? The title of this slide is "Red-Teaming and Robustness." 02:07:50 02:08:02

Are you there yet? 02:08:06

A Yes. 02:08:07

Q Is red-teaming part of post-training? 02:08:08

A I think not in any typical interpretation of the term "post-training." 02:08:13 02:08:16

Q Well, then I guess how does red-teaming and post-training interact? 02:08:18 02:08:24

ATTORNEY BAYLEY: Objection, outside the scope. 02:08:27 02:08:27

A Red-teaming is the general activity of proactively trying to identify failures or gaps in a safety or mitigation stack which can be applied to post-training mitigations, but can also be applied to entire systems or to other levels of mitigations. 02:08:32 02:08:37 02:08:42 02:08:47 02:08:49

Q (BY ATTORNEY PRITT) Meaning red-team can happen at other phases of training? 02:08:53 02:08:58

ATTORNEY BAYLEY: Objection, form. 02:09:03

A I think it's less about the time of when it happens and more about the object that the 02:09:06 02:09:09

a user's request and refusing it. 03:40:17

Q And how is a refusal different from the BLOOMFILTER? 03:40:20 03:40:22

ATTORNEY BAYLEY: Objection to form. 03:40:24

A Under this current definition of a "post-training mitigation," the refusal is outputted as tokens by the model itself, whereas the BLOOMFILTER is a system-level or inference-level mitigation whose output can be used to, for example, completely stop the model from outputting any further tokens. 03:40:30 03:40:32 03:40:37 03:40:41 03:40:43 03:40:48 03:40:52

Q (BY ATTORNEY PRITT) So when a copyright refusal occurs, does the model generate regurgitated content that is not displayed to the user? 03:40:54 03:40:57 03:41:02

A Under -- under our current definition -- and there's multiple definitions for what people call a "refusal" -- the refusal is outputted by the model, and thus, it cannot counterfactually output any copyrighted text. 03:41:08 03:41:10 03:41:13 03:41:17 03:41:22

Q So the model -- so in that context the model never generates any answer that might contain copyrighted text; it simply outputs a refusal; is that correct? 03:41:23 03:41:26 03:41:33 03:41:36

A When the model works as intended under the Copyright Policy of Exhibit 1, then yes, the model 03:41:37 03:41:41

I, MARY J. GOFF, CSR No. 13427, Certified Shorthand Reporter of the State of California, certify;

That the foregoing proceedings were taken before me at the time and place herein set forth, at which time the witness declared under penalty of perjury; that the testimony of the witness and all objections made at the time of the examination were recorded stenographically by me and were thereafter transcribed under my direction and supervision; that the foregoing is a full, true, and correct transcript of my shorthand notes so taken and of the testimony so given;

That before completion of the deposition, review of the transcript ( ) was (xx) was not requested: ( ) that the witness has failed or refused to approve the transcript.

I further certify that I am not financially interested in the action, and I am not a relative or employee of any attorney of the parties, nor of any of the parties.

I declare under penalty of perjury under the laws of California that the foregoing is true and correct, dated this 24th day of February, 2026.

Mary J. Goff
MARY J. GOFF

ERRATA SHEET

CASE NAME: *In re OpenAI, Inc. Copyright Infringement Litigation*, 1:25-md-3143-SHS-OTW (S.D.N.Y.)

DATE OF DEPOSITION: February 12, 2026

DEPONENT NAME: Johannes Heidecke (30b6)

PAGE	LINE	CORRECTION FROM	TO	REASON
13	15	locked	log	Transcription Error
20	16	defense and	defense in	Transcription Error
21	7	nonpost-training	non-post-training	Transcription Error
21	11	as aware	as I'm aware	Transcription Error
33	14	directed	direct	Transcription Error
35	1-2	there's hypotheticals in there in order	this hypothetical scenario	Transcription Error
35	13	words wherein such a word	worlds wherein such a world	Transcription Error
35	24	You're	Your	Transcription Error
55	19	that is also the	that there is also the	Transcription Error
57	11	model	moral	Transcription Error
62	25	his	its	Transcription Error
68	21	But mitigation	But the mitigation	Transcription Error
73	13	for certainty	for full certainty	Transcription Error
75	12	called by	caught by	Transcription Error
91	24	eval car regurg	eval card regurg	Transcription Error
97	12	cybersecurity personal	cybersecurity	Transcription Error
100	5	What actually by	What do you mean by	Transcription Error
103	1	sickofancy	sycophancy	Transcription Error
105	5	are run	are running	Transcription Error
110	10	serious date	serious stage	Transcription Error
123	9	idea	ideal	Transcription Error
131	2	who is	rules	Transcription Error
149	18	columns	column	Transcription Error
153	5	list --	listed	Transcription Error
157	20	defense and depth	defense in depth	Transcription Error
157	21	resilience of robustness of reliability	resilience or robustness or reliability	Transcription Error
159	12	defense and depth	defense in depth	Transcription Error
178	2	irreversible	reversible	Transcription Error
179	10	irreversibility	reversibility	Transcription Error
190	6	strokes	drops	Transcription Error
194	13	"eval bot"	"eval board"	Transcription Error
194	20	evaluations	evaluation	Transcription Error
201	18	pertraining	pretraining	Transcription Error
207	2	where	were	Transcription Error
226	1	seen	see	Transcription Error

PAGE	LINE	CORRECTION FROM	TO	REASON
235	16	Proximate	Proximal	Transcription Error
236	12	rates	weights	Transcription Error
246	13	hedge a voice	hedge or voice	Transcription Error
254	12	a presentation	representation	Transcription Error

I DECLARE UNDER PENALTY OF PERJURY THAT THE FOREGOING IS TRUE AND CORRECT.

Johannes Heidecke

Johannes Heidecke

6/11/2026

DATE