EXHIBIT 57

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

BENJAMIN MANN

Conducted Virtually

Friday, February 20, 2026

9:06 a.m.

Stenographically reported by:
Erica Field, RDR, CRR, CA-CSR, TX-CSR,
GA-CSR, WA-CSR, NM-CCR, IL-CSR, NJ-CCR,
FL-FPR, OR-CSR, NY Notary, FL Notary,
AK Notary, DC Notary
Job No. 620904

INDEX OF PROCEEDINGS

WITNESS	PAGE
BENJAMIN MANN	
DIRECT EXAMINATION BY MR. FRAWLEY	7
CERTIFICATE OF REPORTER	105

EXHIBITS

EXHIBIT	DESCRIPTION	PAGE
Exhibit 1	OPCO_MDL_004412023	10
Exhibit 2	OPCO_MDL_000495790	14
Exhibit 3	OPCO_MDL_00479314	17
Exhibit 4	OPCO_MDL_001686982	22
Exhibit 5	OPCO_MDL_001938901	36
Exhibit 6	OPCO_MDL_000478522	41
Exhibit 7	OPCO_MDL_000494603	43
Exhibit 8	OPCO_MDL_000492829	53
Exhibit 9	OPCO_MDL_000479741	59
Exhibit 10	OPCO_MDL_000479619	71
Exhibit 11	OPCO_MDL_000482605	77
Exhibit 12	ANT0000014	90
Exhibit 13	ANT0000149	90

Keker, Van Nest & Peters, for OpenAI. 09:07:13
MR. SCHWARTZ: Robert Schwartz, 09:07:20
Orrick, Herrington & Sutcliffe, on
behalf of Microsoft. 09:07:24
MR. NISHIMOTO: Ryan Nishimoto 09:07:24
of Arnold & Porter on behalf of the 09:07:26
witness. With me today physically in 09:07:27
the room is Karl Johnston, in-house 09:07:28
counsel for Mr. Mann's current 09:07:30
employer, Anthropic. For the record, 09:07:32
I'm also physically present in the 09:07:35
room with Mr. Mann. 09:07:37
THE VIDEOGRAPHER: The court 09:07:39
reporter is Erica Field representing 09:07:40
Planet Depos, and the witness will now 09:07:42
be sworn. 09:07:46

Whereupon,

BENJAMIN MANN,

having been first duly sworn or affirmed, was
examined and testified as follows: 07:06:33

DIRECT EXAMINATION 07:06:33

BY MR. FRAWLEY: 09:07:59

Q. Good morning, Mr. Mann. 09:08:04
A. Good morning. 09:08:07
Q. My name is Alex Frawley. 09:08:07

on datasets? 10:05:47

A. So in general, for scientific 10:05:47

purposes, we wanted to be sure that when we 10:05:56

ran evaluations on our models that there 10:06:01

wasn't a contamination between what we 10:06:01

trained on and what we evaluated on because 10:06:04

we were looking for generalization, not 10:06:08

memorization. And if we trained on 10:06:11

something, then it's possible we would just 10:06:14

measuring memorization. 10:06:17

And so a big project that I did 10:06:19

over many, many months was what we called 10:06:23

this overlap analysis. And you can see in 10:06:26

these further bullets scale overlap analysis 10:06:29

for big datasets, run overlap analysis for 10:06:32

big and new datasets. Those are both 10:06:35

referring to this overlap analysis which 10:06:38

ended up in the GPT-3 paper that we published 10:06:40

in a few months after this. 10:06:46

So training a Bloom filter is 10:06:48

describing our method at the time or one of 10:06:52

the methods that we used to do this overlap 10:06:55

analysis and doesn't refer to doing 10:06:58

generative fine-tuning but, instead, just 10:07:04

taking the words in the dataset in sort of 10:07:07

So what happens is you take a batch of training data, you run it through the model in what's called the forward pass. You compute the loss in the model's ability to predict the next token, and then you backwards propagate that through the model. 10:54:48 10:54:49 10:54:54 10:54:57 10:54:59 10:55:02

And you do that a bunch of times, computing all these different loss signals, aggregating them and then computing the weights, what the weights should be in the next step. And the process of model training is take many, many, many of these steps. 10:55:06 10:55:08 10:55:11 10:55:13 10:55:15 10:55:17

And a snapshot comes from some periodic recording of what the weights are in the model at a given time and serializing them to some storage like cloud storage. 10:55:20 10:55:22 10:55:27 10:55:32

So it's quite arbitrary what a model is because it could be any of those steps, and there's no theoretical reason, and, in fact, there are many theoretical and practical reasons that you wouldn't want to enforce that you only ever train forward in time. 10:55:37 10:55:39 10:55:43 10:55:46 10:55:49 10:55:50 10:55:55

And we have a lot of infrastructure to make sure that those 10:55:55 10:55:56

CERTIFICATE OF REPORTER

UNITED STATES DISTRICT COURT)
SOUTHERN DISTRICT OF NEW YORK)

I, ERICA FIELD, RDR, CRR, certify that I was authorized to and did stenographically report the deposition of BENJAMIN MANN, pages 1 through 104; that a review of the transcript was not requested; and that the transcript is a true and complete record of my stenographic notes.

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

DATED this 3rd day of March, 2026.

Erica Field

Erica Field, RDR, CRR