

Exhibit 17

CONFIDENTIAL

Page 1

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

IN RE: SOCIAL MEDIA) Case No. 4:22-MD-03047-YGR
ADOLESCENT ADDICTION/)
PERSONAL INJURY PRODUCTS) MDL No. 3047
LIABILITY LITIGATION)
_____)
THIS DOCUMENT RELATES TO)
ALL ACTIONS)

THURSDAY, JUNE 26, 2025

CONFIDENTIAL - ATTORNEYS' EYES ONLY -
PURSUANT TO PROTECTIVE ORDER

- - -

Remote Videotaped Deposition of LARS BACKSTROM, taken pursuant to notice and conducted at the offices of Covington & Burling, LLP, 3000 El Camino Real 5, Palo Alto Square, 10th Floor, Palo Alto, California, 94306, commencing at 2:05 p.m., PDT, on the above date, before Jennifer A. Dunn, Registered Merit Reporter, Certified Realtime Reporter, California (#14461), Illinois & Texas Certified Shorthand Reporter, and Missouri Certified Court Reporter.

Job No. MDLG 7444596

CONFIDENTIAL

Page 12

1 in advance of the deposition.

2 THE WITNESS: Yes, that is accurate.

3 BY MR. OLSZEWSKI-JUBELIRER:

4 Q Thank you for that clarification.

5 A I did not take any notes.

6 Q Okay. So we'll mostly be talking about the
7 Facebook and Instagram platforms today. If I refer to
8 Meta's platforms, I'll be talking about Facebook and
9 Instagram. Okay?

10 A Okay.

11 Q If there's a difference between how Facebook or
12 Instagram work in response to a question that's relevant to
13 your answer, please let me know.

14 A Okay.

15 Q Meta uses various machine learning models across
16 its platforms, including Instagram and Facebook, correct?

17 A Yes. That's correct.

18 Q And roughly speaking, a machine learning model is
19 a tool that's used to make predictions based on data, right?

20 A Yes. Roughly speaking.

21 Q And Meta relies on machine learning models to
22 engage users, right?

23 MR. CHAPUT: Object to the form.

24 THE WITNESS: We use machine learning models
25 for a broad variety of tasks.

CONFIDENTIAL

Page 13

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q One of those tasks is to drive Meta's
3 recommendation systems, right?

4 A Yes.

5 Q And in terms of recommendation systems, we're
6 talking about models that determine, for example, which
7 posts will show up in a particular order and each user
8 speed, correct?

9 A I think we generally refer to the connected
10 content within feed as a ranking sort of task. Although
11 that language is slightly loose versus recommendations of
12 unconnected content or unconnected means that I have not
13 explicitly sort of trended or followed the entity producing
14 the content.

15 Q Okay. And we can talk more broadly about ranking
16 and recommendation systems, but both involve machine
17 learning models that determine which posts show up in
18 particular orders in various feeds across Meta's platforms,
19 right?

20 MR. CHAPUT: Object to the form.

21 THE WITNESS: Yes. Ranking and
22 recommendation systems use machine learning models as
23 one of their components.

24 BY MR. OLSZEWSKI-JUBELIRER:

25 Q Okay. And Meta also uses machine learning models

CONFIDENTIAL

Page 14

1 in systems that determine which notifications to send to
2 users at any particular time?

3 MR. CHAPUT: Object to the form.

4 THE WITNESS: Yes. Machine learning models
5 are a component in the identification systems as well.

6 BY MR. OLSZEWSKI-JUBELIRER:

7 Q Meta's recommendation models are sometimes called
8 algorithms; is that right?

9 A Maybe colloquially. I'd say algorithms is a much
10 sort of more broad and general term. I don't think that we
11 would take algorithms to imply machine learning in general.

12 Q Okay. So let's, if we can, call these types of
13 ranking and recommendation models content recommendation
14 models for the purpose of our deposition.

15 A I think machine learning models would probably be
16 a little bit more precise.

17 Q Okay. Meta also has machine learning models that
18 are involved in determining which ads Meta will show to a
19 user, right?

20 A Yes. Meta is using ads as well.

21 Q Okay. And Meta also has machine learning models
22 that predict users' ages, right?

23 A Yes.

24 Q Okay. Let's talk about how machine learning
25 models work at a very high level.

CONFIDENTIAL

Page 15

1 Meta uses data to train machine learning models to
2 complete certain tasks, right?

3 A Yes.

4 Q And the model finds patterns and makes inferences
5 from data, right?

6 A I don't think I would describe it quite that way,
7 in that that implies that there's a sort of explicit
8 representation of a pattern within the model which I don't
9 think is generally the -- or necessarily the case.

10 Q Roughly speaking, the model makes predictions
11 based on data, right?

12 A Yes. The models make predictions based on the
13 input signals that are put into the models.

14 Q And input signals are also sometimes called
15 features in this context, right?

16 A Yes. I believe those two terms are often used
17 interchangeably.

18 Q Okay. Features are the basis of the models'
19 predictions, right?

20 A Yes.

21 Q And Meta develops machine learning models by
22 training them, right?

23 MR. CHAPUT: Object to the form.

24 THE WITNESS: Yeah. Machine learning models
25 are trained based on data, yes.

CONFIDENTIAL

Page 16

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q And training data contains a set of features or
3 inputs for the model and the label, right?

4 A That's right.

5 Q The label is a piece of information that describes
6 the outcome the model is being trained to predict?

7 A I guess whether it's an outcome or in some
8 other way it's the expected output of the machine learning
9 model.

10 Q Okay. So just, for example, if you're training a
11 model that's designed to predict whether a user will watch a
12 particular video, the training data would consist of videos
13 with labels representing whether a user has watched that
14 video, right?

15 MR. CHAPUT: Object to the form.

16 THE WITNESS: It wouldn't contain the videos,
17 per se. It would contain metadata about the videos
18 perhaps. And then the label would be as you described.

19 BY MR. OLSZEWSKI-JUBELIRER:

20 Q Okay. And the training data would also have all
21 of the other types of data that the model uses as a feature
22 or input, right?

23 MR. CHAPUT: Object to the form.

24 THE WITNESS: I'm not sure what you mean by
25 "all of the other data."

CONFIDENTIAL

Page 17

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q Well, we established that models make predictions
3 based on features, right?

4 A Yes.

5 Q And so those features would be present also in the
6 training data, right?

7 A Generally, yes. Although at times you have to
8 sort of deal with missing values either in the training data
9 or in the prediction task, but I'd say generally the same
10 data is available at training time, as well as inference
11 time.

12 Q And inference time, just for the jury's
13 understanding, that's the time when the model is being used
14 to make predictions based on new or different data, right?

15 A Yes, that's correct.

16 Q Okay. How often does Meta train its machine
17 learning models that are involved in its ranking and
18 recommendation systems?

19 MR. CHAPUT: Object to the form.

20 THE WITNESS: The specific cadence depends on
21 the specific model.

22 BY MR. OLSZEWSKI-JUBELIRER:

23 Q Can you give me an example?

24 A Some of the models are trained on a sort of not
25 quite continuous, but on an ongoing basis and would be

CONFIDENTIAL

Page 18

1 deployed more frequently than once per day.

2 Other models that may be either less important or
3 they may have a sort of more stable and sort of evergreen
4 distribution of data could be trained much less frequency --
5 frequently.

6 For example, a model might try to predict the,
7 you know, high level topic for a piece of content, which is
8 a task that does not change very much from one day to
9 another.

10 And so a model like that might be trained
11 infrequently versus a model based on, you know, that was
12 used for content recommendation with a sort of ever evolving
13 corpus of content, as well as a broad sort of diverse user
14 base where changing preferences would benefit from more
15 frequent training, and those ones are often done at a higher
16 sort of daily or, in some cases, more frequent basis.

17 Q And when you're talking about a more frequent
18 training basis for these models that are sort of
19 continuously updated, are you talking about online training?

20 A Yes. We refer to them online, I think they're
21 often trained online, but then maybe deployed -- or, sorry,
22 at a somewhat slower than realtime cadence.

23 And so a model might be trained as new data comes
24 in, but then to distribute that new model to all of the
25 servers worldwide that it is being used on might happen at a

CONFIDENTIAL

Page 19

1 not realtime, but some other frequency, you know, in the
2 order of, I think, hours in many cases.

3 Q When Meta trains its machine learning models, do
4 the models retain the inferences they learn from prior sets
5 of training data?

6 MR. CHAPUT: Object to the form.

7 THE WITNESS: They retain some of the
8 inferences, but they generally do not retain any of the
9 sort of specific data that was used to make those
10 inferences.

11 BY MR. OLSZEWSKI-JUBELIRER:

12 Q Meta uses user data to train its machine learning
13 models, right?

14 A Not all of them, but certainly, you know, the
15 ranking and recommendation models are -- their purpose is to
16 provide a personalized user experience and provide people
17 with the content that is most relevant to them, which
18 requires using their personal data.

19 Q And models that decide which ads to show to users,
20 those are also included in ranking and recommendation
21 models, right?

22 MR. CHAPUT: Object to form.

23 THE WITNESS: I didn't mean to include them
24 when I said "ranking and recommendation models," but
25 they do also use personal data.

CONFIDENTIAL

Page 20

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q Okay. Let's talk about the types of user data
3 Meta uses to train its machine learning models.

4 Meta uses data on users' engagement on the
5 platform, such as what accounts or posts they have viewed,
6 to train its machine learning models?

7 MR. CHAPUT: Object to form.

8 THE WITNESS: Meta does use data about what
9 accounts people have connected with. And, I'm sorry, I
10 don't remember the second thing you said in that
11 example.

12 BY MR. OLSZEWSKI-JUBELIRER:

13 Q Posts that people have viewed, for example?

14 MR. CHAPUT: Object to form.

15 THE WITNESS: Yes. Meta uses data about the
16 posts that people have viewed.

17 BY MR. OLSZEWSKI-JUBELIRER:

18 Q Meta uses information about users' posts on the
19 platforms such as their photos or comments that train
20 machine learning models, right?

21 MR. CHAPUT: Object to form.

22 THE WITNESS: We've certainly used the
23 metadata, like the number of comments, or whether or
24 not a piece of content is a photo.

25 The extent to which we use the -- the default

CONFIDENTIAL

Page 21

1 photo or the full text, I think, varies widely.

2 BY MR. OLSZEWSKI-JUBELIRER:

3 Q Okay. But for some of Meta's machine learning
4 models, Meta uses the photo to train its models?

5 MR. CHAPUT: Object to the form.

6 THE WITNESS: I don't believe that we use the
7 pixel data of the photos for ranking and recommendation
8 systems. If it's not in the sort final -- in the
9 prediction models that predict user outcomes, such as
10 whether or not you would like a piece of content, for
11 instance.

12 BY MR. OLSZEWSKI-JUBELIRER:

13 Q Meta uses photos in a earlier stage of the
14 recommendation system?

15 MR. CHAPUT: Object to form.

16 Characterization. Scope.

17 THE WITNESS: I think we process the
18 photos in that we attempt to extract sort of topics
19 which I believe use the photo information to
20 understand what's in the photo in a way that's
21 more easy for the machine learning algorithms to
22 use.

23 And so it might map from a photo to a topic
24 ID, for instance. And then that topic ID would be used
25 by the machine learning algorithm.

CONFIDENTIAL

Page 22

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q Meta uses information about users' interests to
3 train its machine learning models?

4 MR. CHAPUT: Object to the form.

5 THE WITNESS: Can you be more specific about
6 what you mean, data about users' interests?

7 In some sense, I guess every piece of content
8 that one likes is some data about -- expressed in
9 interest in a piece of content.

10 BY MR. CHAPUT:

11 Q Okay. In addition to that sort of broader
12 engagement data that you described, Meta also uses
13 information identifying users' interests to train its
14 machine learning models?

15 MR. CHAPUT: Object to the form.

16 THE WITNESS: I think interests
17 sometimes means or topical interests could also mean
18 the sort of people who I'm interested in like sort of
19 types of comments that I'm interested in, I guess
20 would have been sort of -- without more specificity,
21 I would say, in general, given that my behavior, or
22 one's behavior on the platform, you know, has positive
23 engagement with the things that I'm interested in
24 and, yes, we use data about what users are interested
25 in.

CONFIDENTIAL

Page 23

1 BY MR. OLSZEWSKI-JUBELIRER:

2 Q Meta uses users' IP addresses to train its machine
3 learning models?

4 MR. CHAPUT: Object to the form.

5 THE WITNESS: I don't believe that IP
6 addresses are in the training data used by our models.

7 BY MR. OLSZEWSKI-JUBELIRER:

8 Q Does Meta use representations of users' IP
9 addresses to train its machine learning models?

10 MR. CHAPUT: Object to the form.

11 THE WITNESS: Not as I understand
12 representations to mean. In that there's not a sort of
13 one-to-one mapping between an IP address and any other
14 representation that is used in the machine learning
15 algorithms. Or in the machine learning training data,
16 I meant to say.

17 BY MR. OLSZEWSKI-JUBELIRER:

18 Q Meta uses representations of a users' most
19 frequently used IP addresses as a feature in some of its
20 machine learning models?

21 A Again, I don't believe that we use a
22 representation that has a one-to-one mapping with an IP
23 address as a feature in any of our models.

24 Q Does Meta use a representation that has a
25 different mapping that's not one-to-one of users' IP

CONFIDENTIAL

Page 24

1 addresses as a feature of its machine learning models?

2 MR. CHAPUT: Object to the form. Scope.

3 THE WITNESS: I think one way in which IP
4 addresses are used is to ascertain a high-level
5 location, or I should say estimates of high-location
6 since I think the IP to location mapping is far from
7 perfect.

8 And we do use the sort of high-location data
9 in the training data.

10 BY MR. OLSZEWSKI-JUBELIRER:

11 Q Let's talk about the types of users whose data
12 Meta uses to train its models.

13 Meta uses user data collected from users who are
14 logged into their accounts on Meta's platforms to train its
15 machine learning models?

16 A Yes. We use data from logged-in users.

17 Q Does Meta use data from people who visit Meta's
18 platforms without being logged in to train its machine
19 learning models?

20 MR. CHAPUT: Object to form. Scope.

21 THE WITNESS: I'm not sure if we do or not.

22 To the extent that we do, it would be in a sort of
23 aggregated form, like I could imagine a feature that
24 would be, you know, the number of logged out views of a
25 piece of media which could be a useful feature as an

CONFIDENTIAL

Page 25

1 indicator for how, you know, sort of broadly and
2 externally popular a piece of content was.

3 BY MR. OLSZEWSKI-JUBELIRER:

4 Q Meta uses data collected from users who are logged
5 into their accounts on Meta's platforms regardless of their
6 age to train its machine learning models?

7 MR. CHAPUT: Object to the form.

8 THE WITNESS: In general, to provide sort of
9 personalized recommendations in ranking, we're using
10 the data from the sort of viewer to both inform
11 retrieval, you know, at the simplest level, have to use
12 data things like which friends a user has or which
13 pages they follow.

14 And then, you know, further using other data
15 to do liking and recommendations sort of quality.

16 BY MR. OLSZEWSKI-JUBELIRER:

17 Q For example, Meta uses user data of teen users to
18 train its machine learning models, right?

19 MR. CHAPUT: Object to the form.

20 THE WITNESS: Yes. The team experience is
21 personalized similarly to other users.

22 BY MR. OLSZEWSKI-JUBELIRER:

23 Q And there was a period of time before which
24 Instagram did not require users to enter a birthdate; is
25 that right?

CONFIDENTIAL

Page 26

1 MR. CHAPUT: Objection. Scope.

2 THE WITNESS: I'm not aware of the sort of
3 specific history of when Instagram required sort of
4 birthday entering.

5 BY MR. OLSZEWSKI-JUBELIRER:

6 Q And you're aware that there was a time when
7 Instagram did not have stated ages for its users, right?

8 MR. CHAPUT: Objection. Scope.

9 THE WITNESS: I'm not specifically aware of
10 that. My career has always been on the Facebook app
11 within Meta.

12 BY MR. OLSZEWSKI-JUBELIRER:

13 Q Meta uses -- in the period of time in which Meta
14 did not have stated ages for its users, Meta still used
15 their user data to train its machine learning models, right?

16 MR. CHAPUT: Objection. Foundation. Scope.

17 THE WITNESS: I think you said the time when
18 Meta did not have stated age. I don't know what that
19 means.

20 And I think it was -- would have been
21 individual apps, not Meta collecting stated age.

22 Maybe you can clarify the question for me.

23 BY MR. OLSZEWSKI-JUBELIRER:

24 Q Yes. I'll represent to you that Meta's other --
25 another of Meta's designees has testified that before

CONFIDENTIAL

Page 27

1 December of 2019, Meta did not collect birthdates or stated
2 ages from users of Instagram.

3 In that period of time before Meta started
4 requiring new users of Instagram to state an age, Meta still
5 used those users' data to train its machine learning models,
6 right?

7 MR. CHAPUT: Objection. Form.

8 Foundation. Scope.

9 THE WITNESS: I don't have any reason to
10 believe that they wouldn't have used it.

11 BY MR. OLSZEWSKI-JUBELIRER:

12 Q How many users' data are typically included in
13 Meta's training datasets for its machine learning models?

14 MR. CHAPUT: Objection to form.

15 THE WITNESS: In some sense, like with all of
16 the users' data in that everybody -- not all users --
17 not everybody sort of is actively using the platform on
18 any given day.

19 But for people who are using it, we are
20 generally using their personal data to rank their
21 personal feeds and give them personal recommendations.

22 BY MR. OLSZEWSKI-JUBELIRER:

23 Q And in terms of the training data that Meta uses
24 to train its models and algorithms, Meta uses data from all
25 users on a given day that are active on the platform to

CONFIDENTIAL

Page 30

1 another and sort of delays in or removal propagation
2 that may not happen immediately, but that flag is
3 intended and sort of propagates within a matter of some
4 days to lead to the removal of all suspected under 13
5 data.

6 BY MR. OLSZEWSKI-JUBELIRER:

7 Q So in the period of time before that flag has
8 propagated after Meta has determined a user is likely under
9 13 and placed them in an age checkpoint, Meta would use that
10 user's data to train its machine learning models?

11 MR. CHAPUT: Object to form.

12 Characterization.

13 THE WITNESS: Yes. There is a short period
14 of time during which some of that data may appear in
15 the training systems before the deletion has
16 propagated.

17 That said, the primary impact of that data
18 would be for the sort of personalization and quality of
19 that specific individual's experience. And so as long
20 as that user were check-pointed and unable to consume
21 content, it would have a very small sort of impact on
22 the overall model.

23 BY MR. OLSZEWSKI-JUBELIRER:

24 Q But it would still have an impact on the overall
25 model even if small, right?

CONFIDENTIAL

Page 36

1 response to text prompts called Emu?

2 A Yes.

3 Q Emu is a generative AI model, right?

4 A Yes.

5 Q The generative AI models that are used to power
6 Meta AI and Emu were trained on a large amount of data,
7 right?

8 MR. CHAPUT: Object to form. Scope.

9 THE WITNESS: Yes. It depends a little bit
10 exactly on what you mean by large, but I think many
11 terabytes of data.

12 BY MR. OLSZEWSKI-JUBELIRER:

13 Q For example, Emu is trained on billions of images;
14 is that right?

15 MR. CHAPUT: Object to form. Scope.

16 THE WITNESS: I don't know exactly how many,
17 but I would not be surprised if it was in the billions.

18 BY MR. OLSZEWSKI-JUBELIRER:

19 Q Meta has used content posted by users of Facebook
20 and Instagram to train its generative AI models, correct?

21 MR. CHAPUT: Object to the form.

22 THE WITNESS: Yes.

23 BY MR. OLSZEWSKI-JUBELIRER:

24 Q And that includes the models under -- that
25 underlie Meta AI and Emu, right?

CONFIDENTIAL

Page 37

1 MR. CHAPUT: Object to form. Scope.

2 THE WITNESS: I believe Emu is primarily
3 trained on Instagram sort of data as opposed to
4 Instagram and Facebook.

5 BY MR. OLSZEWSKI-JUBELIRER:

6 Q Okay. And Emu was trained on photos posted on
7 Instagram by Instagram users, right?

8 A I believe it used a combination of publicly
9 available images, as well as publicly posted photos on
10 Instagram.

11 Q Did Meta use all public photos on Instagram to
12 train Emu?

13 MR. CHAPUT: Object to form. Counsel, sorry.
14 Do you mean that -- exclusively public photos
15 or every single public photo? The question -- I'm just
16 not clear on what the question is.

17 MR. OLSZEWSKI-JUBELIRER: I meant exclusive,
18 all of the public photos.

19 MR. CHAPUT: Object to the form.

20 THE WITNESS: I don't know if they used every
21 public photo or some narrower subset.

22 BY MR. OLSZEWSKI-JUBELIRER:

23 Q People post photos of themselves on Instagram,
24 right?

25 A People are permitted to post any community

CONFIDENTIAL

Page 38

1 standards compliant photos on Instagram which would include
2 photos of themselves.

3 Q And so Meta used people's selfies that they posted
4 on Instagram to train Emu, right?

5 MR. CHAPUT: Object to the form.

6 THE WITNESS: I believe that people's selfies
7 would have been contained within the Emu training
8 dataset at some level.

9 Again, I don't know the specifics of exactly
10 which public photos were selected and deemed as sort of
11 most useful for the Emu training process.

12 BY MR. OLSZEWSKI-JUBELIRER:

13 Q Meta's used videos posted by Facebook and
14 Instagram users to train its generative AI models?

15 MR. CHAPUT: Object to the form.

16 THE WITNESS: I'm not sure we've used videos.
17 Yeah, I don't know the answer to that.

18 I don't believe that videos were in the
19 initial sort of Emu training, but I'd have to follow-up
20 for a more definitive answer.

21 MR. OLSZEWSKI-JUBELIRER: Let's do A1,
22 please. That's -- this is a document with Bates Number
23 META3047MDL-044-00025293.

24 MR. CHAPUT: Tab 13 in your binder.

25 MR. OLSZEWSKI-JUBELIRER: And can you turn to

CONFIDENTIAL

Page 39

1 the page ending with a Bates number of 25320?

2 (Meta-Backstrom-2 marked for identification.)

3 BY MR. OLSZEWSKI-JUBELIRER:

4 Q Okay. Do you see where it says: "What user data
5 did Meta use to train its AIs?"

6 That's the last question.

7 A Yes, I see that.

8 Q Do you see the second bullet point where it says:
9 "Publicly shared posts on Instagram and Facebook, including
10 photos, videos, and text, were part of the data used to
11 train the generative AI models underlying the features we're
12 announcing at Connect"?

13 A Yes.

14 Q So Meta used publicly shared videos from Instagram
15 and Facebook to train generative AI models, right?

16 MR. CHAPUT: Object to form.

17 THE WITNESS: I think it's clear that we used
18 the posts, and less clear -- and this is sort of just
19 my technical understanding of how much we use sort of
20 the videos versus the metadata and sort of extracted
21 from them or things like that.

22 But yeah, it says that we used photos,
23 videos, and text. And I don't see a date on this to
24 know whether this is sort of referring to Emu or the
25 sort of Llama family of models.