Exhibit 18

UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA

IN RE: SOCIAL MEDIA ADOLESCENT ADDICTION/PERSONAL INJURY PRODUCTS LIABILITY LITIGATION	MDL No. 4:22-md-3047-YGR
THIS DOCUMENT RELATES ALL ACTIONS	

-- -- --

Wednesday, December 10, 2025

-- -- --

CONFIDENTIAL - PURSUANT TO PROTECTIVE ORDER

-- -- --

Video-Recorded Oral Deposition of LARS BACKSTROM, PhD, held at the offices of Covington & Burling LLP, 3000 El Camino Real, 10th Floor, Palo Alto, California commencing at 9:09 am PST on the above date, before Michael E. Miller, Fellow of the Academy of Professional Reporters, Certified Court Reporter, Registered Diplomate Reporter, Certified Realtime Reporter and California CSR #13649.

-- -- --

GOLKOW - VERITEXT
877.370.DEPS | fax 917.591.5672
deps@golkow.com

AI models in ranking and recommendation?

A. Not directly. I believe we use, you know, some of the content understanding capabilities as one of the signals that feeds into the feed ranking algorithm, but that is not -- it is not directly using generative AI, nor are we training the generative AI models as part of the Feed.

Q. I want to go back to a question about machine learning in a previous iteration of Feed when the machine learning model was trained on batches.

Do you recall -- I think you may have mentioned a previous version that was trained on batches of information?

A. Yes.

Q. Okay. How frequently would those batches of information be updated?

A. That varied over the course of the, you know, system. Over time, we've moved to more and more frequent.

You know, when we switched from batch training to more continuous training, I think we were generally doing daily, and we

would train daily, but using a batch of, you know, many days of data. I think at the time -- I don't remember if it was 30 or 90 days, and I think there's some trade-off there in terms of the system sort of compute time required versus the model quality.

But we would train on a block of, say, 30 days, and we would do that every day, so it would be this sort of sliding window of 30 days. Or maybe 90 days. I don't remember the exact sort of number. And it was probably different for different models as well.

Q. Do you recall when, what year that would have been?

A. Well, I think we still have a mixed system where some of the models -- you know, some of the less critical models are trained using that methodology.

I think the main Feed models -- I would be speculating a little bit on the exact timeline, but I would guess in the sort of 2019-2020 time frame, somewhere around there.

Q. Meta also uses embeddings in

Q. And you described for us generally what PYMK is.

Do you recall that?

A. Yes.

Q. What was the extent of your involvement with People You May Know?

And, I guess, to make it easy -- to try to make it easy, let me go back to the same -- as a software engineer, what was your role with PYMK?

A. The product existed when I joined Facebook, but the, like, implementation of it was, I think, had significant opportunities to improve from an infrastructure and machine learning perspective.

And so my role was to develop both the infrastructure, which would significantly increase the number of candidates considered. So whenever we are showing suggestions, there is sort of an initial phase to winnow down from, you know, every single user on Facebook to a smaller set.

Previously, that had been done

with a relatively naive sampling strategy, so I built the structure to expand that, and then in addition to expanding the retrieval set, also implemented the machine learning algorithms for making the predictions.

In this case, the primary prediction task would have been if we make a suggestion, will you send a friend request, and then will that friend request be accepted.

Q. Would that role have been the same when you were a tech lead?

A. So I switched off of PYMK in, I think, like, early 2011. I don't think I would have described myself as a tech lead prior to, like, late 2011.

Q. Okay. In your later roles as either engineering manager, director of engineering or VP of engineering, did any of those roles include overseeing People You May Know's work by others on the team?

A. No.

I guess I could expand on that a little bit in that the People You May Know is primarily owned by a separate organization

outside of Facebook, but there were certainly collaborations that occurred over the course of that period.

And PYMK is shown within the Facebook app, and so naturally, there would have been some collaboration between the team that was primarily responsible for that product and the rest of the Facebook app engineering team.

So while the primary ownership was never under me after, you know, 2011, that's not to say that there weren't some collaborations between my team and the People You May Know team.

Q. And is People You May Know a function that shows up in a user's feed?

A. A function? I guess I would describe it as a product rather than a function, but yes, it shows up within the Feed.

Q. All right. We talked earlier about your familiarity with Meta's machine learning models.

Do you recall talking about that earlier?

some sliding window of data, you know, would include data from recent days where the user was active, even if they were not active today.

BY MR. COCANOUGHER:

Q. And does Meta still use that batch training in their models?

MR. CHAPUT: Asked and answered.

A. In some cases.

BY MR. COCANOUGHER:

Q. For the models using the batch data, are those models trained on photos?

MR. CHAPUT: Objection, vague.

A. Which models in particular? Do you mean the feed ranking models when you say those models?

BY MR. COCANOUGHER:

Q. Sorry. Let's do that.

For the Feed ranking models that use the batch process, are those models trained on the user's photos?

MR. CHAPUT: Objection, form.

A. They would be trained on data from Feed, which would include, you know, a

variety of posts, including photos.

BY MR. COCANOUGHER:

Q. Would that data from Feed also include videos or audio files?

A. I don't think we really have audio, but would include video.

I should clarify that we're training mostly on the metadata of the posts, and so, you know, while it may be sort of active research of how we would use the raw pixels of the photos or raw content of the videos, the training is generally happening on -- you know, not on those raw pixels, but on various pieces of metadata about the posts which might include whether or not it's a photo or a video or the video length.

And then the way that we, you know, maybe use aspects of the raw pixels are through, you know, some of the other models that feed into ranking.

So I described a little bit previously, we might have a separate model which maps from a piece of content to a topic, for instance. And that would be sort of outside the scope of the feed ranking

models but would be a model that takes a -- you know, the pixels of a photo and then tries to predict which -- what the topic of that photo is, and then the feed ranking algorithm might consume that topic as an input.

Q. What about for the feed ranking model, would the metadata from a user name be included in that?

MR. CHAPUT: Objection, form.

A. I don't recall us ever using user name. I think, you know, in principle, it might be, but I don't think that that would be a very useful feature.

And I don't recall that we ever tried to -- you know, it's possible that we did, but I would be a little bit surprised if that was one of the features that we used.

BY MR. COCANOUGHER:

Q. What about user IDs?

MR. CHAPUT: Objection, form.

A. So the main way that user IDs would be used is less as a raw feature, because, ultimately, it's sort of a random number that doesn't have very much meaning,

and so it would be hard for the model to, you know, learn anything from the ID directly.

But, you know, to -- the ID would be used to retrieve some of my, like, past, you know, historical interactions, where one of the features might be how many, you know, likes have I given to this particular user, and then there would be some lookup where you would map from -- say, okay, for this specific user, you would look up how many times have I liked that person's content in the past.

BY MR. COCANOUGHER:

Q. What about cookie IDs?

A. I don't think that we have ever looked at that, and I don't see how it would be useful.

Yeah, from all the work I can recall of adding the features and looking at the various feature importance, I don't think cookie ID has ever appeared.

Q. IP addresses?

A. Yeah, in some ways, similar to ID, like, the IP address is not going to be a very useful signal, and I'm pretty sure we've

never used that as a feature directly.

I think it is probably one of the inputs into, you know, location, and so at a very, like, course granularity, I think it's sort of, like, quite imperfect, but we can estimate what country a person is in based on their IP address.

And while that is, you know, not a definitive sort of answer to what country they're in for a variety of reasons, that's the type of thing where IP addresses maybe had some -- it has not been directly used, but has translated into some feature that then has been used.

Q. What about users' precise locations?

MR. CHAPUT: Objection, form, vague.

A. I don't think that we've ever -- I'm pretty confident we've never used the precise location. I think the variance of location that I've seen, like the most granular, are more at the sort of, like, scale of a, like, city, or I guess for large cities, maybe somewhat more granular. Not --

and I am kind of coming back to what's going to be useful for the algorithms.

What's useful to a consumer is generally not going to depend on their precise location. Even if, you know, we wanted to use it, it's something that isn't available much of the time and is not going to be as useful as the sort of -- if I live in the Bay Area, I'm interested in Bay Area news sort of more broadly, or maybe like I live in Atherton, I'm interested in Atherton news, as opposed to the exact location.

So I don't think we've ever pursued that.

BY MR. COCANOUGHER:

Q. Just two more. Information about users such as age or gender?

MR. CHAPUT: Objection, form.

A. We have used age and gender in our ranking algorithms as features.

BY MR. COCANOUGHER:

Q. Data about how a particular user interacted with content on the platform.

MR. CHAPUT: Objection, form.

A. I guess maybe I'm not

understanding, because to a certain extent, much of my -- as I described earlier, we are trying to ascertain a specific user's interests based on, you know, who they've interacted with, for instance, or what content they've interacted with or, you know, what pages they follow or a number of other kind of behaviors on the platform. And so we are using those things as inputs into the ranking algorithm.

BY MR. COCANOUGHER:

Q. All right. I want to talk now, move to teen engagement.

MR. COCANOUGHER: Can we pull up Tab 5, and we'll mark this as Exhibit 4.

(Whereupon, Backstrom-4, E-mail(s) re: Teens, META3047MDL-040-00240836 - META3047MDL-040-00240837, was marked for identification.)

BY MR. COCANOUGHER:

Q. All right. Mr. Backstrom, this is an e-mail sent to you, correct?

A. No.

Q. If that user attempts to complete the checkpoint by providing information confirming that they are under 13 years old, at that point, Meta will disable the account; is that fair?

A. I believe, yes.

Q. If the user does not respond at all within 30 days of being checkpointed, Meta will disable the account; is that fair?

A. Yes.

Q. And when we -- when I'm saying disable, can you describe what happens when Meta disables an account?

MR. CHAPUT: Object to form.

A. They can no longer log in and they can't use any of the services -- well, yeah, I guess, sorry. That's -- when it's checkpointed, you can't use any of the services.

Once it's disabled, I think at that point, it's not -- your sort of time is already expired, so beyond just not being able to log in, your account is sort of on a path towards deletion.

And then on the other side, the

account is disabled, and so nobody can see the content that was, you know, previously produced by that account.

If you navigate to the profile, it would not exist. It would be as if you went to a nonexistent profile, et cetera.

BY MR. COCANOUGHER:

Q. And just so I'm clear, the scenario when if you went to the account there would be no account and it looked like it didn't exist, that would be once the user is checkpointed? The first step, I guess?

A. That's my understanding, though I guess I would say it's like a little bit outside of my area of sort of expertise.

But I believe once it's checkpointed, until that checkpoint is cleared, the account is in a disabled state, which would mean in the disabled state, you can no longer see the content.

Q. During the period of time after a user has been checkpointed but before Meta disables the account, Meta uses the user's data to train their machine learning models, correct?

MR. CHAPUT: Object to form, incomplete hypothetical.

A. During that period, the account is only suspected, and so, you know, I think a number of those users eventually clear those checkpoints. So while it's in the checkpoint state, it is sort of uncertain as to whether or not the account is under 13.

BY MR. COCANOUGHER:

Q. And so in that scenario, Meta is still using that user's data to train their machine learning models, right?

MR. CHAPUT: Object to the form, incomplete hypothetical.

A. My understanding is that the kind of like hard guarantees of not using data are not in effect during that period.

You know, so there may be -- you know, it's sort of flagged as being checkpointed, and so I'm not sure exactly, you know, all of the implications of that flag.

But the sort of much more affirmative guarantees about not using data do not come into play until either the 30

days expired or the user is confirmed to be under 13.

BY MR. COCANOUGHER:

Q. During that same period of time, do you know if Meta uses the same -- the checkpointed user's data to train their generative AI models?

MR. CHAPUT: Object to the form, incomplete hypothetical, vague as to what period of time is being discussed.

A. I'm not familiar with the exact sort of data filtering of the gen AI team. I think the accounts are flagged as being checkpointed, but the gen AI team is separate from the feed ranking team, and so not aware of the sort of nuances of exactly how those flags are used.

BY MR. COCANOUGHER:

Q. Once the user's account is disabled, Meta begins the process to begin deleting that user's data, correct?

MR. CHAPUT: Object to form, incomplete hypothetical, vague.

A. I think for -- in many senses,

the account is disabled while they're in the checkpoint in that they can't use it and you can't see that user's data.

The process of deletion begins when the 30 days expires or when the user is confirmed to be under 13.

Q. So when that happens, so when the user's account is disabled and Meta begins the process to begin deleting the user's data, the first step is to flag that that user's data is not intended to be used for business purposes, correct?

MR. CHAPUT: Object to form, incomplete hypothetical, asked and answered in Mr. Backstrom's prior deposition in this matter.

A. Yes. Once the 30 days is expired or the user is confirmed, then the data is flagged for deletion and is not intended for business purposes.

BY MR. COCANOUGHER:

Q. But it is not the case that immediately after an account is disabled that user's data is removed from business use, correct?

MR. CHAPUT: Object to form, vague, incomplete hypothetical.

A. I think it is -- so I think it is a 45-day sort of period after the 30 days has expired for things like legal review and to ensure that that is not needed for ongoing investigations and things of that nature.

And so, you know, during that period, the data has not been fully deleted, and while the -- it is not intended for business purposes, the sort of more affirmative deletion does not happen immediately.

BY MR. COCANOUGHER:

Q. And so there's a 45-day period in which, after a user's account is disabled, there's then a 45-day period in which, as you described, certain things happen before the data is actually deleted?

MR. CHAPUT: Object to form, asked and answered, incomplete hypothetical, vague.

A. There is a 45-day review for those -- for, like, I think primarily, like, to understand if it is necessary for legal

retention. And then at the expiration of those 45 days, the deletion process sort of fully begins.

BY MR. COCANOUGHER:

Q. Okay. So is there a period of time during which some of the user's data is used to train Meta's machine learning models before the not-for-business-use flag propagates to remove all of the suspected under-13 user's data from Meta's training systems?

MR. CHAPUT: Objection to the form, vague, incomplete hypothetical, asked and answered.

A. I am not familiar with the specifics of exactly sort of the actual propagation delays.

I think the thing that we are, you know, very sort of diligent about enforcing are the guaranteed delays. I think in some cases, it will happen faster, and some of this depends on the sort of implementation details of the system, exactly how long these various flags propagate through the sort of various aspects of the

answer your question on the exact timelines that it undergoes.

BY MR. COCANOUGHER:

Q. Let's say the implementation took the full 45 days.

Do you know if, for that particular situation, if that data would be used in the model?

MR. CHAPUT: Objection, vague, incomplete hypothetical.

A. I guess, if the implementation were to only remove the data after 45 days, then it would be used for 45 days.

I guess, in your hypothetical, it seems like somewhat tautological that you're saying if it were there for 45 days, would it be there for 45 days? So it's yes.

BY MR. COCANOUGHER:

Q. So we don't know -- in that 45 days, we don't know whether it's deleted on day 1 or day 45, but for whatever period -- practically, for whatever period of time in that 45 days -- strike that. I confused myself.

And so the guarantee, when you

say hard guarantee, is that Meta doesn't use the data for model training after 45 days if an account fails the checkpoint?

MR. CHAPUT: Object to the form, vague, incomplete hypothetical, lacks foundation.

A. After 45 days, the deletion process begins, if the data was not flagged for retention.

BY MR. COCANOUGHER:

Q. So for the second guarantee, which is the 90-day period, is there any amount of time in that 90-day period in which the data still goes into the models?

MR. CHAPUT: Object to the form, vague, incomplete hypothetical, lacks foundation.

A. Similar answer in that the guarantees that we kind of, like, rigorously abide by are the 90-day deletion window. The exact implementation time, I'm not specifically aware of.

BY MR. COCANOUGHER:

Q. And are you aware of whether the implementation time varies -- can vary?