

FILED UNDER SEAL

EXHIBIT 27

Message

From: Rory [REDACTED] [/O=THEFACEBOOK/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=[REDACTED]]
Sent: 1/19/2022 3:36:58 PM
To: Christopher [REDACTED] [REDACTED]@fb.com]; Rory [REDACTED] [REDACTED]@fb.com]; Remy [REDACTED] [REDACTED]@fb.com]; Ovidiu [REDACTED] [REDACTED]@fb.com]; Chandan [REDACTED] [REDACTED]@fb.com]; Aditya [REDACTED] [REDACTED]@fb.com]; Paniz Geder [REDACTED] [REDACTED]@fb.com]; Zoe [REDACTED] [REDACTED]@fb.com]; Wendy [REDACTED] [REDACTED]@fb.com]; Felix [REDACTED] [REDACTED]@fb.com]; Irakli [REDACTED] [REDACTED]@fb.com]; Chris [REDACTED] (CI) [REDACTED] [REDACTED]@fb.com]; Malia [REDACTED] [REDACTED]@fb.com]; Shilpa [REDACTED] [REDACTED]@fb.com]; Dipanjan [REDACTED] [REDACTED]@fb.com]; Scott [REDACTED] [REDACTED]@fb.com]; Baptiste [REDACTED] [REDACTED]@fb.com]; David [REDACTED] [REDACTED]ies@fb.com]; Mark A [REDACTED] [REDACTED]@fb.com]
Subject: Message summary [{"otherUserFbld":null,"threadFbld":4108330035936499}]

Jennifer Malia [REDACTED] (1/19/2022 05:34:00 PST):
>Hey--can we get some other background information on this [REDACTED] SEV?

Jennifer Malia [REDACTED] (1/19/2022 05:34:10 PST):
>(in a slightly safer forum)

Jennifer Malia [REDACTED] (1/19/2022 05:34:19 PST):
>I'm curious how it was found

Irakli [REDACTED] (1/19/2022 05:34:39 PST):
>Hey!

Irakli [REDACTED] (1/19/2022 05:35:09 PST):
>so it's a long story, our CO who was reviewing matches from StopNCII, was saying that there were no video matches

Jennifer Malia [REDACTED] (1/19/2022 05:35:15 PST):
>Oh boy.

Irakli [REDACTED] (1/19/2022 05:35:27 PST):
>but we thought it was because of [REDACTED] having low recall, basically no recall :D

Jennifer Malia [REDACTED] (1/19/2022 05:35:33 PST):
>Yeah, it always starts from "I was looking at this other thing and I noticed ..."

Irakli [REDACTED] (1/19/2022 05:35:45 PST):
>After that we started some investigation and Aditya found out that IG, and Messenger do not do any VMMD5 matching

Jennifer Malia [REDACTED] (1/19/2022 05:36:21 PST):
>Ok. So I'm trying to remember about [REDACTED] I could swear they are trying to move to [REDACTED] overall instead of vDNA. Do we know if that is happening?

Jennifer Malia [REDACTED] (1/19/2022 05:36:31 PST):
>And if so is this limited to just the external ingest hashes?

Irakli [REDACTED] (1/19/2022 05:37:19 PST):
>You mean for outside world? I was talking to David from ISE and he was saying [REDACTED] is the only viable option for external/open source use

Jennifer Malia [REDACTED] (1/19/2022 05:37:34 PST):
>Yes, so I think it's actually internal, too.

Irakli [REDACTED] (1/19/2022 05:37:37 PST):
>yes, because internally we have more powerful hashes which we can use

Jennifer Malia [REDACTED] (1/19/2022 05:37:41 PST):
>Like they are trying to deprecate vDNA

Irakli [REDACTED] (1/19/2022 05:37:57 PST):
>yeah, but they have some SSN, and Videntifier and some fancy things

Jennifer Malia [REDACTED] (1/19/2022 05:38:25 PST):
>I'm checking with our eng that worked this out with them.

Jennifer Malia [REDACTED] (1/19/2022 05:39:08 PST):
>Because I remember this was a slight concern.

Irakli Popkhadze (1/19/2022 05:39:27 PST):
>you mean for NCMEC, or generally ?

Jennifer Malia [REDACTED] (1/19/2022 05:41:01 PST):
>Overall. We were concerned that we'd lose older hashes that ae still being traded if we simply dropped
VDNA

Jennifer Malia [REDACTED] (1/19/2022 05:41:11 PST):
>It was a whole thing with Legal

Christopher Williams (1/19/2022 05:43:28 PST):

PRIV

Jennifer Malia [REDACTED] (1/19/2022 05:44:28 PST):
>Right. If we aren't matching, we don't know that the content is CSAM, so no legal risk.

Jennifer Malia [REDACTED] (1/19/2022 05:44:47 PST):
>There is prob legitimacy risk though

Christopher William [REDACTED] (1/19/2022 05:44:48 PST):

PRIV

Jennifer Malia [REDACTED] (1/19/2022 05:45:37 PST):
>Exactly, I don't think we could be taken to court for this, but obvi, that's why we have @Scott to
advise us, since I am not a lawyer :)

Christopher [REDACTED] (1/19/2022 05:45:44 PST):
>as such are we happy this is a sev 3 only? feels like we're falling a long way short of expectations

Christopher William [REDACTED] (1/19/2022 05:46:04 PST):
>I don't think the recall of these external banks is high

Jennifer Malia [REDACTED] (1/19/2022 05:46:06 PST):
>That all depends on whether you think this is a major legitimacy issue

Jennifer Malia [REDACTED] (1/19/2022 05:46:14 PST):
>We should get numbers here.

Jennifer Malia [REDACTED] (1/19/2022 05:46:24 PST):
>That's prob an @Felix ask.

Christopher [REDACTED] (1/19/2022 05:46:50 PST):
>i think it is - we're not enforcing on the surface where we actually know the harm is. Makes it look
like the ingestion is just a token gesture.

Scott David [REDACTED] (1/19/2022 05:47:04 PST):

PRIV

Christopher [REDACTED] (1/19/2022 05:47:36 PST):
PRIV

Christopher [REDACTED] (1/19/2022 05:47:45 PST):
>we're affected customers :)

Christopher William [REDACTED] (1/19/2022 05:47:47 PST):
>Uh what's broken?

Christopher [REDACTED] (1/19/2022 05:47:58 PST):
>a feature that was never built, it seems

Christopher William [REDACTED] (1/19/2022 05:47:59 PST):
>Is it ISE that broke? I suspect it's ISE

Paniz [REDACTED] (1/19/2022 05:48:10 PST):
>For stopNCII this is certainly a legitimacy risk since we've committed publicly to removing matches on Instagram

Jennifer Malia [REDACTED] (1/19/2022 05:48:13 PST):
>This is probably a SEV along the lines of the one in June, where we think the results of the fix may cause legitimacy issues.

Christopher William [REDACTED] (1/19/2022 05:48:16 PST):
>SEV->wontfix

Christopher William [REDACTED] (1/19/2022 05:48:35 PST):
>Is stopNCII [REDACTED] for video hashing?

Paniz [REDACTED] (1/19/2022 05:48:41 PST):
>yeah

Christopher [REDACTED] (1/19/2022 05:49:04 PST):
>sorry i'm now in mashups for the next 3 hours :(

Jennifer Malia [REDACTED] (1/19/2022 05:49:06 PST):
>Maybe, but I think we need to let SD and ISE hash it out (see what I did there?)

Paniz [REDACTED] (1/19/2022 05:49:27 PST):
>Yeah agreed with this

Christopher William [REDACTED] (1/19/2022 05:49:39 PST):
>Definitely had an extended sigh ☹️

Jennifer Malia [REDACTED] (1/19/2022 05:49:40 PST):
>This might be quite a longstanding issue if [REDACTED] isn't scanned.

Paniz [REDACTED] (1/19/2022 05:50:32 PST):
>Also the same with GIFCT I believe

Christopher William [REDACTED] (1/19/2022 05:50:32 PST):
>So [REDACTED] is oddly an ok algorithm for video matching because of the lack of widespread transcode today, however, I believe it was only ever seen as an interface outside of Meta since we always had better than [REDACTED] internal video matching algorithms

Jennifer Malia [REDACTED] (1/19/2022 05:52:23 PST):
>We likely need some info from SD eng here. There was a big discussion with Legal about deprecating VDNA, and some choices were made about what we would and would not support and I cannot remember the specifics. But we should get them.

Jennifer Malia [REDACTED] (1/19/2022 05:52:43 PST):
>(including all of their embedding/hash types)

Remy [REDACTED] (1/19/2022 05:53:16 PST):
>someone wants to take a stab at giving a TL;DR:?

Paniz [REDACTED] (1/19/2022 05:53:24 PST):
>Sure

Paniz [REDACTED] (1/19/2022 05:53:39 PST):
>We realised 2 days ago that [REDACTED] hashes can't be matched on any surfaces other than public FB

Paniz [REDACTED] (1/19/2022 05:54:16 PST):
>Which is a problem since [REDACTED] is what we receive from NCMEC, StopNCII and GIFCT.
>
>There's varying degrees of legal / legitimacy risk here, at least with StopNCII we've publicly committed to removing matches on public IG

Paniz [REDACTED] (1/19/2022 05:54:51 PST):
>Context: <https://fb.workplace.com/groups/photodna/permalink/1360672227700258/>

Remy [REDACTED] (1/19/2022 05:55:52 PST):
>fantastic, thanks Paniz

Remy [REDACTED] (1/19/2022 05:56:56 PST):
>and I suppose we have no easy way to pivot StopNCII and NCMEC to something like VPDQ (which is on for IG at least)?

Christopher William [REDACTED] (1/19/2022 05:59:48 PST):

>NCMEC has been exploring alternative hashes for a long time without much forward progress. @David [REDACTED] is much closer to this.

>StopNCII would require another algorithm capable of running client side ideally in javascript.

>GIFCT probably isn't possible given the fragmented and varying technical capabilities of the members.

David Thomas [REDACTED] (1/19/2022 05:59:53 PST):
>Hey, sorry all, I haven't started working yet

David Thomas [REDACTED] (1/19/2022 06:00:08 PST):
>I'll be in the gym for maybe another hour and then can engage

Christopher William [REDACTED] (1/19/2022 06:00:22 PST):
>Has anyone checked the SEV guidelines? Is missing expected functionality really a SEV 3?

Christopher William [REDACTED] (1/19/2022 06:00:38 PST):
>There's an Integrity SEV wiki.

David Thomas [REDACTED] (1/19/2022 06:01:43 PST):
>My understanding is this is the state that has been in place for at least a couple of years, I can pull in ISE oncall if we should triage now

Christopher William [REDACTED] (1/19/2022 06:02:54 PST):
>I think we should switch to SDI as the primary for this rather than ISE

Christopher William [REDACTED] (1/19/2022 06:03:12 PST):
>Looks like Zoe is out sick

Charles Rory [REDACTED] (1/19/2022 06:03:25 PST):
>I would normally advocate for missing functionality not being a SEV but in this case we have made a public commitment to be doing this for StopNCII at least.

>
>>help prevent the image or video from being shared anywhere on Facebook and Instagram
>
>It feels like that meets the bar for "An issue that is high enough priority to focus on, but with more limited immediate product impact or less urgency for resolution" in terms of urgency but I don't feel super strongly.

Christopher [REDACTED] (1/19/2022 06:03:35 PST):
>agree ^

Remy [REDACTED] (1/19/2022 06:05:20 PST):
>I'm happy with sev ownership switching to SDI (@Chandan [REDACTED] since it's a video coverage issue). I guess the next step would be to define the steps (and weight the tradeoffs) of adding [REDACTED] coverage for IG

Jennifer Malia [REDACTED] (1/19/2022 06:05:59 PST):
>Legitimacy issues can also be an Integrity SEV rather than an engineering SEV.

Christopher William [REDACTED] (1/19/2022 06:06:03 PST):
>If we're scanning FB especially retrospectively and converting those hashes into the hashes of other algorithms (need confirmation from SDI or potentially ISE who I'd expect to build such functionality on this) I could argue we are meeting this commitment.

Jennifer Malia [REDACTED] (1/19/2022 06:06:04 PST):
>Basically

Jennifer Malia [REDACTED] (1/19/2022 06:06:26 PST):
>I don't think we are though. Scanning ... how?

Christopher William [REDACTED] (1/19/2022 06:06:44 PST):
>Md5? The little gun if nothing else

Christopher William [REDACTED] (1/19/2022 06:07:00 PST):
>Then once we have the video we have the other hashes

Jennifer Malia [REDACTED] (1/19/2022 06:07:58 PST):
>Maybe I'm really confused, but I think I need to understand what the [REDACTED] issue is first. Big gun and little gun are just retriggering matches using existing, working algos, right?

Remy [REDACTED] (1/19/2022 06:08:32 PST):
>I'm not aware of such behaviour but that may just be my ignorance (behaviour: ensure a UGC matching a seed gets fingerprinted for other embeddings)

Irakli [REDACTED] (1/19/2022 06:08:49 PST):
>If video is only shared in Messenger, we won't have any other hashes, right?

Christopher William [REDACTED] (1/19/2022 06:08:53 PST):

>SDI's opinion should weigh more than mine but...

>

>I think the realistic future for SDI is only two video matching algorithms running globally because of computational issues: (1) the production version, today this is the Videntifier (2) a beta algorithm or iteration on the production algorithm potentially even sampled.

>

>We need a translation layer on the video hashing side for sure.

Jennifer Malia [REDACTED] (1/19/2022 06:09:15 PST):

>> UGC matching a seed

>

>So my concern is the original match to the seed, right?

Christopher William [REDACTED] (1/19/2022 06:09:34 PST):

>FB public is a no trivial percentage of recall

Christopher William [REDACTED] (1/19/2022 06:09:42 PST):

>Is a non trivial percentage of recall

Jennifer Malia [REDACTED] (1/19/2022 06:10:00 PST):

>My experience is that it takes quite a while for something viral on chat to get onto public

Jennifer Malia [REDACTED] (1/19/2022 06:10:07 PST):

>Someone typically messes up.

Paniz [REDACTED] (1/19/2022 06:10:33 PST):

>I mean, even messenger aside, public IG?! Surely that's not defensible

Christopher William [REDACTED] (1/19/2022 06:10:39 PST):

>Look I'm not advocating for not addressing this: (1) is this a SEV? What's the scale of this issue? (2) what's a good short term mitigation and the long term plan

Remy [REDACTED] (1/19/2022 06:10:41 PST):

>taking stopncii as an example. [REDACTED]

Christopher William [REDACTED] (1/19/2022 06:10:57 PST):

>yes

Remy [REDACTED] (1/19/2022 06:12:33 PST):

>ok, I'll talk to SDI oncall or Chandan. Let's see how we can mitigate this for IG public that would be a start. Long term I really hope we can rally the industry and the non-profit org toward a model where we provide a service for them to fingerprint using whatever tech we provide. (a bit a-la-SK TBA)

Jennifer Malia [REDACTED] (1/19/2022 06:13:05 PST):

>Ok, so let me see if I have the landscape correct here:

>Legal risk --> not for CSAM or DOI, maybe for some countries for NCII

>Legitimacy risk --> yes, from medium to quite a bit, needs a SEV

>Enforcement volume risk --> likely not enough to open a traditional SEV or make it higher than SEV3

Jennifer Malia [REDACTED] (1/19/2022 06:13:59 PST):

>Yes?

Paniz [REDACTED] (1/19/2022 06:14:10 PST):

>>Legal risk --> not for CSAM or DOI, maybe for some countries for NCII

>

>I don't think there's any legal risk for NCII, unless there's legal risk associated with not meeting our public commitment made on StopNCII

ovidiu [REDACTED] (1/19/2022 06:14:28 PST):

>took me a while to go through this. I think there is another aspect that might be relevant here about IG and [REDACTED] hashing.

>

>A couple of years ago, we had a conversation with some folks from IG that were working on video. They told us that the encoding of the videos for IG on the client side is lossy. This means that the same video encoded different times by different client will have a different [REDACTED] So [REDACTED] matching on the server side was not useful for IG.

Christopher William [REDACTED] (1/19/2022 06:14:30 PST):

>Doesn't the Netherlands have a law on the books already?

Christopher William [REDACTED] (1/19/2022 06:14:34 PST):

>Some Benelux country

Jennifer Malia [REDACTED] (1/19/2022 06:14:59 PST):

>(or India? But I refuse @Paniz knows the laws)

Paniz [REDACTED] (1/19/2022 06:15:05 PST):

>Yeah, but they send us content and we bank it. StopNCII is the only place where we only receive hashes.

Ovidiu (1/19/2022 06:15:32 PST):

>We actually tried to do client side computation before any encoding/filters being applied

Ovidiu (1/19/2022 06:15:54 PST):

>but that did not work out on the client side app... so we fully abandoned that idea.

Jennifer Malia (1/19/2022 06:15:58 PST):

>For IG because of the heavy emphasis on filters and cropping, right?

Christopher William (1/19/2022 06:16:07 PST):

>This really speaks to the only shot term solution to be to use FB Messenger or IGDirect to find the matching video file and then grab its hashes for other algorithms and finally bank those hashes

Christopher William (1/19/2022 06:16:29 PST):

>Any idea why?

Christopher William (1/19/2022 06:16:47 PST):

>PDQ client side in IG had some success on photos if I remember what you and others have told me Ovidiu

Ovidiu (1/19/2022 06:17:13 PST):

>I think the library that was computing the video required the video to be fully loaded in memory on the phone

Ovidiu (1/19/2022 06:17:23 PST):

>and it created performance issues

Ovidiu (1/19/2022 06:17:30 PST):

>this is a very long time ago :(

Scott David (1/19/2022 06:17:37 PST):

PRIV

Ovidiu (1/19/2022 06:17:38 PST):

>so memory is very fuzzy

Jennifer Malia (1/19/2022 06:21:39 PST):

>Who was working on this, do you remember @Ovidiu. I also remember this work happening, but I wasn't following closely.

Ovidiu (1/19/2022 06:31:34 PST):

>So here is what I've been able to dig from 2019 when we tried to build PDQ client side hashing - <https://fb.quip.com/6lF0AxbNZGbZ> - not a very relevant quip and some eng POC's from the IG side - Chris Walker

Christopher William (1/19/2022 06:33:04 PST):

>Should this be an ISE effort or a SDI effort?

>

>I feels like a feature primarily for ISE

Christopher William (1/19/2022 06:33:19 PST):

>Let me flip flop a couple more times so I'm ready to go into politics

Aditya (1/19/2022 06:42:41 PST):

>For the impact, we discovered this when we were trying to figure out why we don't have any incoming videos in StopNCII review queues - (as a prior 15-20% of matches from the earlier ncii pilot are videos). we still don't have any violating videos from stop ncii flow even though we are matching in facebook

Jennifer Malia (1/19/2022 06:44:40 PST):

>signing you up to PMQs next week

Charles Rory (1/19/2022 06:45:08 PST):

>For content reported to NCMEC it looks like about 13% of FB public matches using MMS are using as the embedding: https://fburl.com/scuba/pms_actions/s45xjdq3 . That doesn't mean it'll be the same on private surfaces obviously but it's some kind of data point at least

Paniz (1/19/2022 06:45:16 PST):

>Can you invite me to your parties @Christopher

Christopher William (1/19/2022 06:45:34 PST):

>Work parties aren't parties. They're work

Christopher William (1/19/2022 06:45:47 PST):

>Essential work

David Thomas [REDACTED] (1/19/2022 06:55:08 PST):

>Hey all, coming online now. Just trying to get my head on the convo so far:

- >
- >1. [REDACTED] matching is not occurring on many surfaces, and that has been the case for a long time
- >2. We think that we want to fix this, because it undermines (at least) our commitments to StopNCII
- >3. There have been previous experiments on original video md5 matching, but we aren't sure it's practical
- >4. We don't know what a solution might look like, or who should tackle it.
- >5. While [REDACTED] has its problems, and proprietary hashing on external content is something we're interested in, that's not an option for StopNCII right now

Paniz [REDACTED] (1/19/2022 06:58:07 PST):

>I'm also still a little confused by the NCMEC situation. Yeah we might not be crossing a legal line but we're still not matching any incoming NCMEC video hashes that exist anywhere outside of public FB

Jennifer Malia [REDACTED] (1/19/2022 06:59:18 PST):

>Yeah, so this is a major thing for our relationship with them

Christopher William [REDACTED] (1/19/2022 06:59:28 PST):

>Lgtm

>On 4 I do think using FB to find other instances, grabbing their hashes, and banking those is a viable strategy

Jennifer Malia [REDACTED] (1/19/2022 06:59:31 PST):

>So while not illegal, also really not good

Irakli [REDACTED] (1/19/2022 07:00:49 PST):

>StopNCII is already doing that, if we find a content we are grabbing better hashes. But problem is if content is only circulated using Messenger or never leaves IG

David Thomas [REDACTED] (1/19/2022 07:01:11 PST):

>This is my first time diving into a potentially very deep topic that crosses a lot of team lines in SDI - I'd like to create a doc or something and start scribing. Is there a particular model that folks prefer?

Christopher William [REDACTED] (1/19/2022 07:02:28 PST):

>@Ovidiu [REDACTED] gave the history of [REDACTED] on IG (and decent chance IGDirect). Does anyone know about the possibility of getting the original [REDACTED] on messenger?

Aditya [REDACTED] (1/19/2022 07:02:41 PST):

>we can, we already get sha256

Christopher William [REDACTED] (1/19/2022 07:03:07 PST):

>This is great news. I agree that this is still a problem but it's now an even further bounded risk.

Christopher William [REDACTED] (1/19/2022 07:03:37 PST):

>From the client or from the server upload? Do any of the clients transcode?

Scott David [REDACTED] (1/19/2022 07:04:20 PST):

PRIV

Aditya [REDACTED] (1/19/2022 07:04:21 PST):

>not sure whether the client does any modifications, from what I saw we do this in the server

Christopher William [REDACTED] (1/19/2022 07:04:52 PST):

>So covering the [REDACTED] could be a very quick solution

Irakli [REDACTED] (1/19/2022 07:09:36 PST):

>fwiw, I tried to test Messenger uploads yesterday.

>

- >1. Filmed a test video
- >2. Uploaded on Messenger
- >3. Downloaded from Messenger
- >4. [REDACTED] hashes for original video and result from step 3 had different hashes
- >5. Uploaded video from step 3 again to Messenger
- >6. Downloaded from Messenger again
- >7. Hashes for those two videos were still different

Ovidiu [REDACTED] (1/19/2022 07:22:40 PST):

>it's from the server side

Jennifer Malia [REDACTED] (1/19/2022 07:22:59 PST):

>Question: If we bank an vmd5 hash from a collaborator and then find a better hash, we bank that hash automatically in the same bank?

David Thomas [REDACTED] (1/19/2022 07:23:01 PST):

>Wrote some stuff here: <https://fburl.com/s258090>

>

>I'll keep combing the thread and adding context, taking a short break

Jennifer Malia [REDACTED] (1/19/2022 07:23:19 PST):

>(that's likely an SD question)

Chandan Kumar (1/19/2022 07:33:11 PST):

>yes we have a pipeline that uses md5 matches to backfill other hashes: `backfill_exact_md5_matches.py`

Jennifer Malia [REDACTED] (1/19/2022 07:35:27 PST):

>And then those hashes remain in the original bank (e.g. NCMEC_HASH_INGESTION or whatever?)

Chandan Kumar (1/19/2022 07:36:01 PST):

>yes

Jennifer Malia [REDACTED] (1/19/2022 07:36:58 PST):

>thank you, that helps my understanding