

Message

From: Guy Rosen [/O=THEFACEBOOK/OU=EXTERNAL (FYDIBOHF25SPDLT)/CN=RECIPIENTS/CN=]
Sent: 11/5/2018 4:30:14 AM
To: Andrea Saul []@fb.com]; Vishwanath [] []@fb.com]
Subject: Re: CSER / CEI / Messenger
Attachments: image001.png

For those others, we don't proactively but we may well do so reactively if reported. I expect the number will be pretty small there.

From: Vishwanath []
Sent: Sunday, November 4, 2018 8:28:54 PM
To: Guy Rosen; Andrea Saul
Subject: Re: CSER / CEI / Messenger

I believe there are smaller numbers for terrorism, but I can find exact counts tomorrow. We probably have link-share deletes in messenger for Spam as well, which maybe a large number.

Afaik, we don't delete content in messenger for hate, graphic violence, nudity, bullying and harassment.

	A	B	C	D	E
1	How do we measure?		How do we enforce?		Remarks
2	Prevalence	Content Actioned/ Proactive Rate	Where do we delete content?		
3	Nudity	fb only	fb only	fb only	We measure prevalence for the same things we enforce on
4	Graphic Violence	fb only	fb only	fb only	We measure prevalence for the same things we enforce on
5	GTO	-	fb + msgr	fb + msgr	We dont measure prevalence
6	Hate Speech	[fb only]	fb only	fb only	We measure prevalence for the same things we enforce on
7	Spam	[fb only]	fb + msgr	fb + msgr	Prevalence and enforcement numbers are different, but not external
8	CN&SEC	-	fb + msgr	fb + msgr	We dont measure prevalence
9	B&H	-	fb only	fb only	We dont measure prevalence
10	Fake Accounts	fb only	fb + msgr	fb + msgr	Different

From: Guy Rosen <[]@fb.com>
Date: Sunday, November 4, 2018 at 8:19 PM
To: Andrea Saul <[]@fb.com>
Cc: Vishwanath [] <[]@fb.com>
Subject: Re: CSER / CEI / Messenger

On CSER: we should actually ask the question of whether we include any Messenger enforcement numbers at all. I'm only just internalizing that we do that at all, given how we talk about it is very FB centric, and we tend to minimize our discussion of message scanning anyway. Do we know how big it is of the other violation types?

Of the hard choices, I think excluding Messenger in CSER starting now (across the board) might be the less bad option. It seems defensible: yes, we published the CEI number including Messenger one-off, but generally the on-going report is FB only right now.



On Nov 4, 2018, at 8:16 PM, Guy Rosen <[REDACTED]@fb.com> wrote:

Early in the year we explored the idea of client-side hashing, which is hard but theoretically doable, with the WhatsApp team. It goes against e2e encryption principles as laid out today, because it makes us a mediator and could be abused in bad ways (imagine a govt ordering us to add a hash of an "illegal" photo to the bank). But, as we explore where Messenger goes, we're going to have some hard discussions on these principles. A wonderful outcome from my perspective would be to get comfortable with this not only on Messenger but also on WhatsApp - and get a chance to reduce a ton of harm.

On Nov 4, 2018, at 7:53 PM, Andrea Saul <[REDACTED]@fb.com> wrote:

Question: is there any way to pick up on hashed images even if things are encrypted? (Not trying to be dense, but curious if there's a way)

On Nov 4, 2018, at 7:30 PM, Vishwanath [REDACTED] <[REDACTED]@fb.com> wrote:

This is based on an old snapshot of only 2 months of data, but going by the same rate, we'd have a smaller number around 3M.
Just looking at egregious CEI content almost 90% of it is in messenger. 😊

I need to think more about this, but given we already shared it in the newsroom post, I think sharing a different number in CSER may be risky. There is very little chance of us not getting asked about it and then we basically out how much is deleted on messenger versus fb and that may just cause push back ("Why is Facebook now allowing millions of pieces of CEI to be sent on messenger?") later when we want to actually encrypt. It boils down to do we defend this now or do we defend this later.

Additionally, I think we delete terrorist propaganda also (the numbers there are smaller), and now co-ordinating harm (re: Myanmar) as well.

On safety carveouts: I think most of this (probably all) is images/videos (need to check) and if we can carve those out from encryption due to safety concerns, I think we could manage in the long run. Between privacy and safety, I think this could be a balancing act, if possible.

<image001.png>

From: Guy Rosen <[REDACTED]@fb.com>

Date: Sunday, November 4, 2018 at 6:57 PM

To: Vishwanath [REDACTED] <[REDACTED]@fb.com>, Andrea Saul <[REDACTED]@fb.com>

Subject: CSER / CEI / Messenger

Give where discussions are trending, at some point in the next couple of years we'll do more encryption on Messenger. I'm going to write up

more for the CI team later this week, but ahead of that wanted to get your thoughts given Vish I think you flagged a potential issue to me. The exact scope, safety carveouts, client-side detection tech etc are still under discussion and will likely take months to work out and maybe a couple of years to build. But the direction is fairly consistent at this point.

Should we consider limiting the CEI number we put in CSER to FB only, to avoid a step-change some time in the future when we lose access to private messaging? I know we've already released the top line number in the newsroom post -- so maybe it is already too late. But putting it in the CSER, where in the future we'll have a chart and everything, feels more irreversible.

Pros and cons here, and we don't know exactly what the future will look like.

What do you think?

<image001.png>