

AMENDED Exhibit 663

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Case No.: 4:22-md-03047-YGR

MDL No. 3047

In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation

Message

From: [REDACTED] [/O=THEFACEBOOK/OU=EXTERNAL
(FYDIBOHF25SPDLT)/CN=RECIPIENTS/CN=4DED1DB347E54CB382A4FF4BCE42F960]
Sent: 4/20/2017 11:11:54 PM
To: Kevin Systrom [REDACTED]
Subject: Re: Integrity HC @ IG
Attachments: Site Integrity_ Key issues to resource.pdf

Hi Kevin,

Enclosed please find the document on site integrity, containing the problems we're trying to solve, and examples thereof. I've also pasted it below, for easy access.

Please note there is one outstanding statistic that the team still needs to track down (% WAP who sees suggestive sexual content). They're running the query for that and expect to get it late tonight. Wanted to get you this sooner, rather than later, but will send an update version with the stat, once we have it.

Please let me know if you have any feedback or need anything else.

Thanks!

Site Integrity: Key issues to resource

Executive Summary

There are 4 key areas that we believe will benefit from more intense integrity work.

- **Comments:** Prime area where bullying occurs for regular accounts, women and teens particularly impacted. In a recent study, 46% of those surveyed reported experiencing bullying/harassment on IG; 50% reported receiving unwanted or offensive comments. Teens delete twice the % of comments from unconnected commenters than connected commenters. Top Accounts also impacted.
- **Abuse and harm in Live:** Every week: 55M+ Live broadcasts, with 65K videos reviewed based on user and automated reporting. Weekly, 1.5K videos are removed from Top Live and 300 are deleted from the platform. On average, we escalate 3 Live videos a week to our internal LERT team which then contacts local law enforcement authorities because users are in imminent danger from issues ranging from self harm to child exploitation. Post-Live launch will make this even more of a vulnerability.
- **Suggestive sexual content:** 40% of nudity photo reports are on suggestive content (e.g. cleavage, bent over, etc.) and X% of WAP see these content. "Edge" stories will exacerbate the problem if there is no down-ranking algorithm in place for suggestive content. In addition, 12% of WAP see violating nudity content as defined by us but 22% reported seeing violating content from user research. This underscores the fact that what IG defines as violating nudity is different than what a global, diverse community considers to be objectionable.
- **Hashtag abuse:** Hashtags are used to congregate abuse because of their ease of discoverability. We have some detection systems for spam and nudity, but not for other abuses such as self-harm and bullying. Ex: recently, 'death group' hashtags in Russia encouraged teenagers to commit suicide. We don't have a systematic way to discover or monitor these hashtags except for manually restricting the hashtag to 'Top-Post Only'.

##

Problem: Comments

- **Regular accounts are often bullied via comments, particularly teens and women.**
 - In a recent study, 46% of the total surveyed reported experiencing bullying/harassment on IG; 50% reported receiving unwanted or offensive comments
 - Teens delete twice the % of comments from unconnected commenters than connected commenters
- **Teens and women suffer from repeat bullying by the same small set of individuals.**
 - 25% of all deleted comments come from repeat bullies (4+ deletes in a month)
 - In a recent study, 46% of those bullied or harassed were harassed 2-4 times on Instagram and 30% were harassed 4-10+ times

- Many of the profiles that are ultimately disabled have not posted anything and appear to be created expressly to troll comment threads
- **Top Accounts are consistently bullied and sometimes leave the platform as a result. Each incident creates a secondary effect of indirect bullying for teen fans who read the comments and experience the negativity while also creating a negative press cycle.**

Redacted for PII

Problem: Abuse and Harm in Live Videos

- **Every week, there are more than 55M live broadcasts.**
 - 65K videos are reviewed from user or automated reporting. 1500 videos are removed from Top Live and 300 are deleted from the platform.
 - On average, we escalate 3 live videos a week to our internal LERT team which then contacts local law enforcement authorities because users are in imminent danger from issues ranging from self harm to child exploitation. Some examples include:

Redacted for PII

concerning.

- **Verified + Shielded IG accounts going Live with Community Standards violations.**

Redacted for PII

- 8% of all Shielded accounts on IG are *not* Verified and don't have any partner managers. These accounts still receive the highest level of protection when it comes to CO review, which makes it very difficult to take their content down if its violating.
- 35% of all Verified accounts are *not* in the Media Ops book-of-business. In the event of violating content, it takes a long time to identify the accounts' partner.

Possible Solutions

- **Show self-harm checkpoint as soon as review team takes the action, instead of at the end of the video.** In the event of a credible suicide threat, have don't have the tools to provide immediate resources to the broadcaster. The self-harm checkpoint, which contains 24/7 hotline links and other well-being information, is only shown at the end of the broadcast.
- **Use FB's suicide comments classifier in IG to prioritize credible self-harm reports.** Through manual labeling, Facebook has developed a classifier that is 80% accurate in using comments left on live videos to identify Livestreams that contain actionable suicide threats. These Livestreams can then be prioritized in CO queues for faster review, which in turn decreases exposure of this content within the broader community.
- **Implement tiered shielding, so that only the most valuable accounts receive the highest protection.**

Problem: Suggestive Sexual Content:

- **What IG defines as violating nudity is different than what our users consider as nudity because our community is diverse and has varying perspectives.**
 - 12% of WAP see violating nudity content as defined by us but 22% reported seeing violating content from user research.

- This discrepancy is even more prominent for some countries like Iran, where 52% of survey respondents said they saw violating nudity. Similarly, there is also a varying degree of how upset users get from seeing objectionable content.
- **Suggestive content is prevalent on Instagram.**
 - 40% of nudity photo reports are on suggestive content (e.g. cleavage, bent over, etc.) and X% of WAP see these content.
 - Edge stories will exacerbate the problem if there is no down-ranking algorithm in place for suggestive content.

Possible Solutions:

- Give user control in what they don't want to see and down-rank their feed accordingly.
- Quarantine potentially violating/suggestive content until it is reviewed.

Problem: Hashtag Abuse

Hashtags are a vector for abuse because of their ease of discoverability. We have some detection systems for spam and nudity but not for other abuses such as self-harm and bullying.

Redacted for PII

Possible Solutions:

- Implement a system that can restrict content to 'Top-Post' only in real time based on report or action ratio of the hashtag.

From: Kevin Systrom [REDACTED]

Date: Wednesday, April 19, 2017 at 3:15 PM

To: [REDACTED] [REDACTED] [REDACTED]

Cc: [REDACTED] [REDACTED]

Subject: FW: Integrity HC @ IG

[REDACTED] please see below.

I'd like to go back to Mark with one short term thing and one long term thing:

1) I don't think Mark understands the urgency of working on integrity related issues at IG. Do you guys have anecdotes and links I can send? I'm assuming the child killing himself on live is an important one, but I think there were others (a suicide?). My point to him will be that it's happening and he's just not as close to it, but my goal is to build empathy. Anything related to how bad comments have gotten and the many issues we've faced with public figures leaving would be good too. I basically need a list with supporting evidence.

[REDACTED] if you want to start running with this and gather things from [REDACTED]/team today that'd be great. Please do not share this note though.

2) Long term, I don't think that Mark thinks about IG needing similar things to FB because he's not as close to it. How do we close that gap? I'd like to talk through this with you guys and figure out how I can get ahead of these issues in the future.

From: Mark Zuckerberg [REDACTED]

Date: Wednesday, April 19, 2017 at 2:30 PM

To: Kevin Systrom [REDACTED]

Subject: Re: Integrity HC @ IG

Thanks for the heads up. These both seem like important initiatives and I'm glad you're focused on them.

Here's what I'll do: I mentioned in small group a few weeks ago that of the ~50 unallocated engineers, I'm going to allocate ~25 to integrity initiatives, ~12 to video and ~12 to groups / communities. Since I had not

been tracking that IG had additional integrity needs, I asked the PAC, feed integrity and ads integrity teams to propose how to allocate the ~25 people. I'll now make sure we include IG in this mix.

I should call out though that we're facing more extreme issues on FB right now with the murder, bad activity in private groups, etc. So I do view funding that as more urgent in the near term. I may also shift some of the video and groups headcount to integrity initiatives to make sure they're funded. This is just meant to say that I probably can't get you 13 engineers in the near term, even though I'm supportive of these projects and agree we should allocate more to them as more engineers become available.

From: Kevin Systrom

Sent: Friday, April 14, 2017 1:33:33 PM

To: Mark Zuckerberg

Subject: Integrity HC @ IG

Mark,

I wanted to give you some context on our Integrity-focused headcount request that will come with the rest of the integrity asks. There are two initiatives we need to staff up: comments integrity (8 eng HC) and business integrity (5 eng HC).

Comments integrity: I made a public commitment to making Instagram a place where people feel safe to be themselves, without criticism or harassment. Because of the public nature of Instagram, comments are the most sensitive aspect of this, the place where the majority of abuse happens. We have lost public figures over this, come under criticism from policy groups, and generally taken heat in the media.

We need to focus a team around comment safety and integrity (not simply adding features), both giving control to post authors and automatically identifying and stopping abuse. Right now we are faced with staffing down growth work on stories or feed ranking to fund this, and we will do one of those in H2 without further headcount. This HC is in NYC.

Business Integrity: Businesses on Instagram are set to double to over 12.5M just since January. We need to build a business integrity team to ensure people don't continue to encounter ig-specific negative experiences (fraud, for example). In the beginning the focus will be integrating with Business and Pages integrity systems at FB, and over time there will be more Instagram-specific integrity work (as we have found with PAC).

Out of context these asks may seem strange (eg. 'Why comments? That doesn't sound like integrity' or 'Why can't the business team just plug into FB's work?'). I wanted to make sure you understood that they're both critical areas for safety on IG and we're doing our best to solve them.

Thanks,

Kevin