

AMENDED Exhibit 207

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: Vaishnavi J [/O=THEFACEBOOK/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN= [REDACTED]]
Sent: 7/27/2020 9:28:24 PM
To: Vaishnavi J [REDACTED]
BCC: Antigone Davis [REDACTED]; Karina Newton [REDACTED]

Subject: Safety on Instagram: Biweekly update - July 27, 2020

Hi all,

(Recipients in bcc)

There has been a lot of great work happening across our teams and the broader company in the last two weeks, so this is a pretty hefty update. Before diving in, a quick plug for the H2 IG safety policy plan session on Wednesday. It's entirely optional, but if you're interested in what's top of mind in H2 for IG from a safety perspective, this should be helpful.

Thank you also for all the feedback you've shared over the past few iterations. Two recurring pieces of feedback from you that I've incorporated into this edition:

- A calendar of what launched in the last two weeks and what's coming soon, to provide as much transparency as possible. Some of these launch dates are tentative subject to the projects being approved, but it's helpful to get a sense of what might be coming up soon, even if not confirmed. Projects with tentative launch dates are marked with (tbd).
- Removing the SP / IG points of contact, as they give the impression that the named folks are driving the whole project, when the work is often a much broader cross-functional initiative than Safety Policy or IG Policy. If there are specific projects you'd like to know more about but aren't sure who to reach out to, please let me know and I'm happy to connect you with the right folks across our teams.

As always, if you missed any earlier issues, you'll find all the previous biweekly updates [here](#).

cheers,
Vaishnavi

JULY 27, 2020

Launched:

- Jul 21: Personal Fundraising on IG
- Jul 22: Hiding profile photos / links in message preview on MSGR for IGD parity
- Jul 23: Stricter enforcement on rape threats across IG / FB
- Jul 23: Public test for updated Live comment filter

Coming soon:

- Jul 27: IGD updates - Feature limits, Hide reported messages, Thread-level reporting, Increased reporting coverage
- Aug 4/5: Spark AR cosmetic filters reversal announcement
- Aug 5: Reels launches globally
- Aug 6/7: Recommendations guidelines go public
- Aug 8 (tbd): Updated Live comment filter launch
- Aug 10: Community Standards Enforcement Report goes live
- Aug 10: FB/IG Election Hub goes live
- Aug 27 (tbd): Video captioning on IG

Content

- **Implicit sexualisation of minors:** Following review with OCP leadership last week, OCP is now working on drafting the operational guidelines and will share with Safety Policy and IG Policy for feedback when ready. If we can secure the necessary operational and IG review resources, the team is (tentatively) working towards a September 2020 launch.
- **Public figure harassment:** Following a number of high-profile cases of harassment and escalations on IG, we are working on a proposal to improve the way we offer protections to public figures on Instagram. We will meet with OCP in the coming weeks to present this proposal and get feedback; in the meanwhile, please share with us any cases of harassment of public figures that have impacted you or your team.
- **Stricter enforcement against rape threats across IG / FB:** On July 23, we announced stricter penalties for rape threats, following an Instagram escalation in India where a man made rape threats towards a woman comedian. Going forward, accounts found to be making a severe rape threat will be disabled. We are working with OCP to define severity and lower the overall account disable threshold for rape threats. This is an important step towards treating high-severity violations with more severe penalties than the current strike system allows, and will go a long way in supporting our external arguments that we prioritise safety of users on our platform.
- **Recommendations Guidelines going public:** We will be publishing our Recommendations Guidelines across FB and IG on Aug 6 / 7. This will be the first time that we will be making our criteria for recommending content public and is an important step in transparency and accountability. The guidelines will live on the FB and IG Help Center pages respectively, tailored to each platform, but with the same underlying policy language.
- **Community Standards Enforcement Report:** We are reviewing the upcoming CSER on Aug 10 to understand which IG-specific numbers are most likely to invite additional scrutiny. Across the board, content actioned has decreased primarily due to a reduced and remote workforce as a result of COVID-19.

Regulatory

- **UK Samaritans industry guidance:** We went back to the Samaritans with comments on the second draft of the SSI industry guidance that they are developing. Notably, they have removed their calls to halt encryption and live-streaming services following our strong feedback to them as a unified industry voice. They are finalising the guidance this week and will be launching this publicly on September 10 to coincide with World Suicide Prevention Day.

Product

- **Enabling Room Creation on IG and WA without FB Account:** Messenger leadership has requested that policy leadership provide an answer on whether Policy would support enabling IG and WA users to create Rooms without requiring them to have a FB account. At the moment, it looks like we will be able to maintain all the necessary integrity mitigations to do this, with the exception of calibrating the IIC classifier to perform as well on

IG as it does on FB. Calibrating the classifier in time will depend on operational capacity from CI and CO, and we have escalated this ask to product teams.

- **Instagram Reels launching globally on August 5:** Ahead of the launch, we have been reviewing curation guidelines, ensuring integrity measures are in place, and tracking cases of under / over-enforcement on Reels. Reels have also been integrated into the current IG HERO monitoring program which proactively reviews the most viral Reels for any harmful content. This program is currently operational in US, and expanding to India, Brazil, and Indonesia this week. This should mitigate the fact that our integrity classifiers are less effective on video and are not calibrated for short-form video in particular (i.e. all Reels content). One continued exposure for us is that we still do not have a sense of how minors' follow or DM graphs are changing after they go viral and the way in which this risk will be amplified when we launch globally. We had asked for this mitigation to be in place following the v2 launches but have hit a wall with the Data Science team working on this due to limited bandwidth. We are now reaching out to Messenger to see if they have bandwidth to help us measure and track this.
- **IGD updates around reporting and moderation - July 27:** IGD will globally launch
 - **Feature limits**, which prohibit a user from being able to send messages, make calls, or react to messages on IGD after incurring a strike for a violation. The feature limit can be applied for different lengths of time depending on requirements. At present, there is no disable threshold for repeated violations in IGD or Messenger that do not result in an immediate account disable. An enforcement threshold for repeat violators in IGD and Messenger will be established in H2 2020.
 - **Thread-level reporting**, which will enable people to report the last 30 messages in a thread instead of only having the option to report an individual message at a time.
 - **Contextual review**, which will allow reviewers to see up to 15 messages before and after a reported message instead of a single reported message in isolation
 - **Increased reporting coverage**, which will allow people to report stickers and other previously unreportable message types
- **IGD updates around reporting and moderation - Aug 4:** IGD will, for cross-app threads, launch
 - **Hide on Report**, which will hide reported messages or threads from the reporting individual(s), which will avoid re-victimising them.
 - **Thread Warning**, which will warn individuals when they are entering a thread that they had previously reported.
- **Private by default launch pushed back:** The private by default launch will be pushed back from August 13 by at least two weeks, exact date TBD. We are holding on safety partner briefings for now until it has gone through leads review and been approved.
- **Closed captioning coming to IG videos:** In a promising win for accessibility, video captions are being planned for IG and IGTV videos. We're reaching out to understand what safety mitigations are in place for bullying, SSI, or child safety content that's detected through our captions, and syncing with the FB accessibility team to understand what mitigations they have in place.
- **Redesign of IGTV standalone app:** The team has launched a redesign with a focus on discoverability and recommendations. They are also working with the IG Live team to incorporate Live video into IGTV discovery units in H2.
- **Live comment filters:** Live comments are those showed on users' live video streams with Instagram Live. The filter that the Bullying team currently uses is more than two years old so the team is updating model with more training data accumulated. Public tests began this week and the initial launch will be in English only.
- **Continuing to promote Wellness Guides:** The distribution efforts to amplify Wellness Guides on IG are set to expire on August 11. However, given the influx of upcoming mental health moments in H2 for IG, and the IG social comparison team moving onto voter empowerment work in H2, we feel the Guides could continue to be valuable for our users. We are putting together a proposal to extend these distribution efforts or continue to amplify Wellness Guides on IG in other ways.

Outreach

- **Instagram ambassadors safety briefing:** We briefed young people in the US from the IG Ambassadors scheme on our safety tools, key safety tips, and asked for feedback on what is and is not working for them. This is part of our ongoing engagement with the program, to ensure our safety messages are aligned, and that we are using opportunities to engage with young people in a valuable way.
- **Web Foundation consultation on online gender-based violence:** We joined the Web Foundation's consultation on advancing the human rights of women activists online. We shared the greatest challenges around women's safety and online gender-based violence on Instagram, how we have addressed online gender-based violence as it relates to content moderation and speech, and what has been most effective in mitigating harms. We put together some talking points for the session that might be helpful for this broader group if you're in similar conversations with your safety or policy stakeholders.
- **Project deSHAME Youth Advisory Board:** Project deSHAME II is a collaboration between Childnet (UK), Save the Children (Denmark) and UCLan (UK), co-financed by the EU. The group aims to tackle peer-based online

sexual harassment by empowering local communities to work together to increase reporting among young people. We joined their youth advisory board meeting to hear feedback around the types of sexual harassment young people are experiencing on Instagram, and what policy and product interventions would be most effective for them to feel safer on our platform.

- **Get Digital on Instagram:** In August 2020, Get Digital will relaunch on FB. The Get Digital program promotes digital citizenship and well-being and provides children with the 21st century skills they need to navigate and succeed in the digital world. We will be amplifying the launch of Get Digital through Instagram, sharing IG-friendly visual content with safety and policy partners to amplify through their respective IG handles.

Great finds on / about Instagram

- **Suicide prevention content in multiple languages on major platforms:** This study from Cyber Policy Center at Stanford looks at whether suicide prevention resources across the industry are available to a broader audience. Notably, IG has the most aggressive approach among all the platforms towards addressing potential SSI content and offering help. The post also suggests that platforms find ways to provide support resources in countries where platforms may not have clear civil society partners, such as inorganically serving up supportive content.
- **Which topics are associated with changes in social comparison on Instagram?** This excellent note from the IGWB research team looks at which kinds of posts on Instagram inspire the most positive social comparison, and which elicit negative comparison. The team surveyed approximately 15,000 people in 9 countries over two waves in March and April 2020. Particularly interesting results to me were what affected teens and women differently from the general population, and why they tended to feel more positively or negatively about content.

DOCUMENT SLIPSHEET

Bates Number: META3047MDL-034-00329532-META3047MDL-034-00329535

BEGATTACH: META3047MDL-034-00329532

ENDATTACH: META3047MDL-034-00329535

TNAG BEGBATES:

TNAG ENDBATES:

Create Date: 7/27/2020 2:28 PM

Last Modified Date: 7/27/2020 2:28 PM

Custodian: [REDACTED] CRENSHAW_LIZA; DAVIS_ANTIGONE; GELDWERT_DAYNA;
JAYAKUMAR_VAISHNAVI; [REDACTED] [REDACTED] NEWTON_KARINA; [REDACTED]
SINHA_RAVI; [REDACTED]

Title:

Author:

Confidentiality Designation: HIGHLY CONFIDENTIAL (COMPETITOR)

META Pointer:

META Pointee: