

AMENDED Exhibit 555

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

[P&C TnS] Feature - Minor Safety

2022 H2 Roadmap

pushpin

This roadmap aligns actionable and measurable product features to the Issue team @Christina Crimmins's initiatives for Minor Safety [(HYPERLINK

Feel free to comment or reach [REDACTED] for any feedback, thank you!

Updated on 7/5/22

Vision - TikTok Minor Safety

Be the best in the industry for a safe and enriching experience for minors (<A18 users)

- ◆ Build and deploy industry leading solutions for minor safety.
- ◆ Demonstrate and help the industry to build best-in-class protections
- ◆ Educate the users (parents, minors), regulators and other key partners about the state of the art features/protections
- ◆ Keep children away from harmful content proactively
- ◆ Actively identify and thwart any predator/grooming efforts
- ◆ Identify, remove and report the most egregious content
- ◆ Provide age appropriate experiences to children
- ◆ Ensure the digital well-being of children
- ◆ Educate minor users and their guardians to effectively use TikTok to find joy

"Be the best in the industry for a **safe** and **enriching** experience for minors"

While our focus is on minors, we also provide tools to their guardians/parents to ensure minors' wellbeing

Mission Statement - TnS PnP Feature

Provide a portfolio of effective product features to accomplish our vision for Minor Safety.

2022 H1 Summary / Recap

TikTok-Furlong

69

4/12/2025 - MOP

██████ to add research findings on July 8th

In the first half of 2022 feature work focused on: (1) age assurance, (2) keep harms away from minors, (3) content classification, and (4) empowering the user with tools. We've had a number of high profile launches including underage appeals, new screen time management tools, and updated model launches for badness detection. Through this we've achieved positive movement in 3 of our top 4 metrics (minor vv badness, minor vv extreme badness, # NCMEC reports).

tada

Highlights

Age assurance

- * Age appeals were launched globally for underage account bans
 - LIVE team has integrated with the process (50% rollout)
- * Underage detection strategies have improved, we're now banning significantly more minors than last year

Keep harms away from minors

- * Gandalf and other model improvements have led to a significant drop in minor vv badness (.93% -> .78%) and extreme minor vv badness (.33% -> .28%)
- * Launched NCMEC multi-video selection in CSAM TCS queue (collaboration with Account team), significantly reduced reporting failure rate from 84.24% US, 7.49% Global to -> 1.33% US, .05% Global

Empowering users with tools

- * Screentime management upsell aimed at teens has received positive press already
- * Family pairing upsell for likely parents has passed all internal reviews and is mid development, research findings are already validating the want

Content classification

- * Despite hurdles we've aligned on the V1A launch of content classification
 - Additional features will be available to experiment with right after

Team

- * Have created a very tight integration with the issue team
- * Close alignment with other T&S teams (digital wellbeing, LIVE, content classification)

exclamation

Lowlights

- * QA challenges in the age assurance projects led to long delays and internal confusion
- * Lack of documentation on proper integrations with TTP led to long delays on US launches for age verification
- * Despite our efforts, there is a slight increase in minor safety issue violation badness
 - We are researching the underlying cause

High-Level Data Summary: H1 2022

Below is a lookback against our H1 goals. Note that many of the goals were unfortunately not operationalized, thus not possible to measure. These have been hidden from view, or marked as undefined.

For our measurable goals we've made good progress, improving 3 out of the 4. Moving forward we will have a new set of goals that better aligns to our themes and are easily measurable.

More data available in the appendix

2022 H2 Key Focus Areas

Feature focus areas

The below list encapsulates the themes that the feature team will be developing products against in the upcoming half. This list aligns with the projects that we're carrying over from H1, and with the issue team's priority pillars as defined in their H2 roadmap: [HYPERLINK

bulb

Note that due to the feature development lifecycle, our pillar of "Keep harms away from minor" combines the issue team's two pillars of Stomp Out Child Sexual Exploitation and Keep Broader Harms Away from Minors

1. Foundation: Age assurance

In order to protect minors, we need to know who on our platform *is* a minor. This pillar is focused entirely on validating users' ages and giving them the tools to verify their age when we get it wrong, or require it for areas with higher risk (ex: LIVE).

Sub projects include (not limited to):

- Age detection capability enhancement
 - Underage Account Ban Strategy
 - Underage Account Removal Requests
- Age appeals / verification
 - Vertical expansion (LIVE, DM, New age verification use cases, etc..)
 - Enhancements (Removing login step, notifications)
 - Add additional "medium certainty" age appeal methods (ex: Instagram's 3 account verification tool)

How do we measure success?

- (TBD) Underage Accounts success metric
 - Appeal: Drop out rate
 - Underage ban: Accuracy

2. **Keep harms away from minors**

Core to Minor Safety is protecting the platform, our community at large, from minor safety violative content and behaviors. This is a wide ranging pillar that includes *protection from inappropriate content* to *interactions between minors and potential predators* in areas such as DM and LIVE.

The feature work involved here is primarily based around creating new comments that detect minor safety issue violating behavior, and creating new ones.

Sub projects include (not limited to):

- Minor safety violating **video content** model improvements
- Minor safety violating **comment** model improvements
- Comment filtering
- DM predator detection

How do we measure success?

- MS P00 Issue Badness < 0.XX%
- MS Issue Badness < 0.11%
- Minor Violation Badness < 0.70%
- Minor High Harm Badness < 0.25%

3. **Empower minors with tailored tools and safety concepts**

TikTok has a suite of features that provide a safe and positive experience for younger users. This pillar supports the development of new features and the enhancement of existing ones,

but also the promotion of the existing features that disproportionately impact younger users (ex: screen time management upsell for minor).

Sub projects include (not limited to):

- Family pairing - [HYPERLINK

]

- Digital wellbeing - [HYPERLINK

- Leveraging content classification for new tools - [HYPERLINK

- Restricted mode enhancements to align with content classification levels

How do we measure success?

- * User Safety Perception Score ("TikTok is a Safe Place for Teens") > 4.9 out of 7 (2021 score)

- Note we will analyze this for both teens and adult respondents

4. **Reactive tooling capability**

Despite our best efforts to proactively protect minors, things will slip through and we need proper tools in place to react quickly. Our focus is on in app tooling, such as increasing awareness of user level and content reporting, and NCMEC reporting to get highly violative content to law enforcement agencies as quickly as possible.

For NCMEC reporting please refer to the much more robust roadmap from the PM lead,



[HYPERLINK

Sub projects include (not limited to):

- NCMEC improvements
- Reporting flow enhancements

How do we measure success:

- Reports Prioritized by NCMEC (100%+ increase compared to [x] in 2022 H1, as provided by NCMEC)

bulb

Note on content classification

Content classification is being separated from the Minor Safety roadmap, into it's own pillar. Please see [HYPERLINK

for their

bookmark

Blue = User safety tools @ [REDACTED] @Matthew Tenenbaum; Red = Model team @ [REDACTED] Yellow = Account team @ [REDACTED] Green = LIVE team @ [REDACTED] Comment team @ [REDACTED] DM team @ [REDACTED] Search team @ [REDACTED] NCMEC @ [REDACTED] Digital Wellbeing @Jordan Furlong

Prioritized List

straight_ruler

Prioritization Framework

Below are criteria and rationale we use to prioritize features. We will focus on the most impactful and high compliance projects first and work through the backlog progressively.

- * **P00: Compliance Risk:** A measure of how compliant we are with all the regulations across the world
- * **P00: Extreme harm:** A measure of how minors get impacted in the real world due to their experience on TikTok
 - Sexual exploitation (including revictimisation)
 - Life-changing/life-threatening physical or psychological harm
- * **P0: Harm:** A measure of the impact of the experiences on TikTok on minors' wellbeing
 - Impaired physical, emotional or intellectual development
 - Serious physical or psychological harm
- * **P0: Perception harm (Brand Risk):** A measure of how industry/regulators/media/policy-makers/NGOs/parents/general public perceive us
 - Our responsibility is to the minors; however, the ecosystem plays a significant role in how we operate. Building perception is key to maintaining the brand value for TikTok

exclamation

This living list only contains P0s. Please check the appendix for P1/P2 backlog items

Commented [1]: Will fill in following columns after getting more internal alignment.

Appendix Deep Dive

Below is a list of detailed projects. Almost all P0 projects are represented above, however more exploratory projects that are P1 / P2 will be included below along with more details about the pillar.

Additional focus area context / low priority backlog

1. Foundation: Age assurance

How can we identify what age our users are, and allow them to appeal decisions that we make regarding their age? At TikTok, the systems and processes for age-assurance (some of which are already built) will be industry first. This is a massive initiative that requires extensive coordination across Feature/R&D, Issue, Policy, Ops, Legal, Comms and Public Policy...

pushpin

Foundational Pillar - High quality age assurance is *key* to success for all other MS initiatives. If we can't determine age, we can't provide tailored experiences.

1.1 Age Determination (Account ban)

While users create accounts on TikTok, they are shown a screen to enter their date of birth (age gate). Using this data to personalize the user experience on TikTok is a default mode of operation till date. However, relying on age-gate turns out to be less effective because there is no way for us to ensure that users are accurately entering their true date of birth.

In response, we need a sophisticated user age determination process that involves using combination of machine learning models to predict the age of the users with high precision. The ultimate aim for this effort is to identify and remove all accounts of users that are <13 years. The account feature team is focused on:

- * Recall leakage - Model leakage has been increasing. To combat, we will integrate new signals (comments, Live, etc..) to trigger account recall
- * Enqueue logic - Ensuring that accounts are enqueued efficiently, so that high probability underage accounts are reviewed first
- * Moderation efficiency - Improving the accuracy and speed of account reviews with changes to ops tools

Much more details in the Account U13 team roadmap: [[HYPERLINK](#)]

1.2 Appeals Expansion and Enhancements

While the above initiative helps us in determining the age of the users, we will never be 100% accurate. To provide a viable recourse to the users, we have also invested in building an appeals process. Using this age appeals process, users will have the ability to prove that they are >13 year old using one of the 3 methods - ID based verification, verification with a trusted adult, credit card based verification.

bulb

In H2 we're focused on both **expansions and enhancements** to the underage appeal flow.

Appeal expansion into more verticals and spaces

More and more vertical areas are requiring age appeals. As we develop more products, additional risk is often introduced that we need to ensure minors are not exposed to. Now that we have our foundational work complete for age appeal flows we're able to integrate into these new and existing areas. Example project include

- * Integration with U16/18 LIVE age checks and bans
- * Integration with the LIVE KYC check
- * Integration with DM U16 removals

Enhance the age appeal process

Our initial launch was truly an MVP, focused on developing a creative solution to get as many users as possible a usable verification flow --**quickly**. Now that we've had our launch we can start looking towards improvements based on data validated pain points, user feedback, and industry trends.

1. Remove the login step from the underage appeal flow
2. Send push notifications to accounts when banned, or an underage appeal decision is reached
3. SOP updates
4. Appeal limits

2. Keep harms away from minors

How can we reduce/minimize minors 'exposure to age inappropriate content and lower the risk of leading to real world harm to minors.

To reduce CSAM and prevent predators from targeting and contacting minors users in TikTok, we are upleveling our predator detection capabilities to proactively detect/ban

predators, send the suspected users to CST queue for further review and set limits on their DM access.

For minors, we aggressively detect underage users and remove them from DM feature. We are building stricter privacy settings and content intervention features to enable minors to filter requests from high-risk senders and protect themselves from potential harm.

Backlog

3. Empower minors with tailored tools and safety concepts

How can we promote a healthy experience on TikTok while ensuring that minors feel safe? The digital wellbeing team is making a lot of improvements to the features (including the same offerings in Family Pairing), Minor Safety has an opportunity to tailor this for families and upsell them.

4. Reactive tooling capability

Despite our best efforts to proactively protect minors, things will slip through and we need proper tools in place to react quickly. Our focus is on in app tooling, such as increasing awareness of user level and content reporting, and NCMEC reporting to get highly violative content to law enforcement agencies as quickly as possible.

Subcategory: Reporting awareness and enhancements

* How can we increase reporting/blocking volume and accuracy for minors? Note this is a subcategory of "Tooling Awareness", however reporting has been identified multiple times as a key feature, thus the separate callout.

H1 Review

H1 Prioritized project list

In H1 we delivered on **15 of our original 27 P0 projects**.

H1 Data deep dive

Minor Safety Issue Health (MS Violation Badness)

Minor Protection (Minor Viewed Badness)

Refer to [REDACTED]
[REDACTED] for more details.

H1 Additional key data observations

Outside of our goals, data has generally trended positively for the Minor Safety ecosystem.
Below are some key callouts and H2 related outcomes.

Key Contacts

Cross feature collaborator POCs

Feature Module	Description	Contacts
User	User Safety tools/ User Safety Perception team will lead features to reduce underage users and provide age appropriate content for minors.	[REDACTED] [REDACTED]@Matthew Tenenbaum [REDACTED] [REDACTED]
Content classification		[REDACTED]
Content	Video backend safety strategy, hashtag safety, etc	[REDACTED]
LIVE	Integrating Minor Safety features with the LIVE team - ensure that our principles flow through to new and existing LIVE experiences	[REDACTED]
Search	Integrating Minor Safety features with the Search team - ensure that our principles flow through to new and existing search experiences	[REDACTED]
Direct Message	Integrating Minor Safety features with the DM team - ensure that our principles flow through to new and existing DM experiences	[REDACTED]

Comment	Filter comments that are not suitable for minor users, design comment safety tools.	
Account	The account team will enable the detection of underage accounts, and help us accurately determine the age of minors (13-18). Their work is integral in ensuring we serve the right experiences to users	
Model	The model PM build minor safety related models to aggressively filter inappropriate content for minors	

[该类型的内容暂不支持下载]

DOCUMENT METADATA

Bates Number: TIKTOK3047MDL-111-LARK-05914855-TIKTOK3047MDL-111-LARK-05914866

BEGATTACH: TIKTOK3047MDL-072-LARK-01131507

ENDATTACH: TIKTOK3047MDL-111-LARK-05914868

Attachment Count: 0

PRODVOL: MDL-111

Custodian: FURLONG, JORDAN

File Path: /DOXUSB4YA6YFI9OUYVFSRWAQ3LD.DOCX

Confidentiality Designation: CONFIDENTIAL

HASHVALUE: A529C79A9D7517FB016C64370E0AA807

DocType: URL LINK DOCUMENT

Author:

Create Date: 7/1/2022 12:00 AM

Last Modified Date: 7/1/2022 12:00 AM

LAST MODIFIED BY:

TRACK CHANGES:

COMMENTS:

HASHIDDENDATA:

Filename: DOXUSB4YA6YFI9OUYVFSRWAQ3LD.DOCX

Title: [P&C TNS] FEATURE - MINOR SAFETY 2022 H2 ROADMAP

DOCEXT: DOCX

Email From:

Email To:

Email CC:

Email BCC:

Received Date: 12/31/9999 7:00 PM

Sent Date: 12/31/9999 7:00 PM

TIMEZONE:

Subject:

Case 4:22-md-03047-YGR

Document 2877-9

Filed 03/25/26

Page 15 of 15

THREADID:

REDACTIONS:

REDACTION TYPE: