

AMENDED Exhibit 403

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

Underage Users: Detection, Safety, Privacy, and Age-Appropriate Design

Areas of Focus

The Minor Safety team has launched four programs to ensure the safety and protection of our minor users, including the removal of underage users from the platform or age-gated features.

These working groups, and the projects within them, have been scoped and prioritized working in tandem with the head of Minor Safety Public Policy in the EU, EU legal groups, and Global Hub leaders. We've specifically worked with our EU partners as legislation out of that region is the most progressive and is often followed by the US. Because of the importance of these markets, we believe in doubling down on our efforts to get ahead of the curve on incoming regulations and to set TikTok apart as an eager player in the Minor Safety space.

This document will seek to answer the following questions:

How do we proactively prevent children (<13) from accessing the app? How do we proactively prevent underage minors (<16) from accessing the age-gated features?

How do we reactively detect children who are accessing the app? How do we reactively detect underage minors (<16) who are accessing the age-gated features?

Can we push children (<13) to a more appropriate experience without banning their accounts?

How do we stop minors from behaving inappropriately on the app?

How do we prevent harm to underage users in the app?

How do we support the mental wellbeing and healthy development of minors?

How can we improve on all of the above?

Our roadmap and prioritization reflects where we can address the questions above while elevating the integrity of TikTok for minors across the most critical areas of work. We've appointed program managers to head up our core working groups, and they will be exclusively focused on answering these questions within their working groups:

Age-Assurance

Focused on developing mechanisms, processes, and product solutions to better understand the age of our users globally, and processes in place to remove underage users. Creating an appeals mechanism is a critical aspect of this work.

Minor's Privacy and Safety by Design

Ensures TikTok creates a safe experience for users of different ages across the design and product process, while balancing the right of minors to express themselves. Ensures that privacy settings and data usage are legally compliant for minors.

Age-appropriate Content Classification & Curation

Ensure that content is appropriately restricted for different age-groups, that it is easily reportable, and that user-reporting in general reflects our commitment to Age-Appropriate design and Safety by Design.

Predator Detection

Builds out best in class Minor Safety predator detection, including external partnerships with key stakeholders.

In this review, we'll take a look at our proposed approach as well as review additional proposals to elevate our approach to keeping minors safe.

Age-Assurance

Focused on developing mechanisms, processes, and product solutions to better understand the age of our users globally and processes in place to remove underage users.

The following section will address these areas of concern:

How do we proactively prevent children (<13) from accessing the app? How do we proactively prevent underage minors (<16) from accessing the age-gated features?

We rely on an age-gate (self declaration) in all markets

Not every user has undergone an age-gate (some were acquired from Musically).
Everyone in the EU, US, and KR has undergone an age-gate.

How do we reactively detect children who are accessing the app? How do we reactively detect underage minors (<16) who are accessing the age-gated features?

We currently have an underage-model to detect underage users in White T markets and KR

This model pushes users into underage account queue that is reviewed by moderators

How can we improve on this?

Launching a global appeals process for users banned for being underage (or if they are banned from using a certain age-gated feature) and aligning the policy to be more strict on when we remove an underage account:

This is a critical piece of our age-assurance strategy and will have huge impact on our business as it may result in the removal millions if not tens of millions of underage users yearly.

Our policies are far too lenient on how we define underage users and we are currently allowing very obvious children on the platform. We need to take a more 'common sense' approach to our policies, meaning we will see a large increase in total banned accounts. While an appeals process using IDs isn't perfect, it's better than nothing at all.

Giving users the ability to appeal is equitable and fair.

Ensure our process complies with PAS standard 1296 on online age checking

We can reduce over-moderation risks by giving users a chance to prove their age.

We already have a global underage appeals process in place for Livestream.

We already have an underage appeals process in place in the US, KR, and Thailand.

Deciding whether to use in-house or third-party verification is still open for discussion, but preliminary review from the Minor Safety team indicates a preference for in-house.

There are many privacy concerns regarding the collection and storage of identification documents.

Not all minors have standard IDs that can be used. We need to deeply think about what kinds of verification we will accept (ex: school IDs, allowing parents to verify the age of their child through Family Pairing) and what different IDs are acceptable in different regions.

Continue to limit use of hard identifiers and develop a clear guidance on when it is and is not appropriate to require ID documents for U18s.

Use Family Pairing for account authentication. Opportunities for enhancing our age assurance ecosystem through Family Pairing include:

Asking parents to confirm the age of a connected teen.

Using Family Pairing to allow parents to confirm their age and their relationship to a blocked user as part of the underage appeals process further minimising reliance on hard identifiers.

Allow parents to set the date of birth for connected teens - even if this overrides their self-declared age.

Lowering the risk profile of the platform by giving parents greater ownership over their children's access to high risk features. This could lower the overall risk profile of the platform and, in turn, lower the threshold for 'sufficient certainty' of the age of users.

Detecting misuse and training our teams to recognize regional IDs will be difficult. We need to decide what level of fraud we are comfortable with.

Revise our training programs and resources to ensure they provide meaningful, best practice guidance on assessing the age minors

Creating new underage models to reflect our new, more strict, policy

This is a critical piece of our age-assurance strategy.

Our current underage model can be easily adapted to reflect our new policy stance.

Increase use of proactive searching using keyword indicators to find underage users accounts and direct the account to the underage moderation queue

The Level Model is not always the most accurate and we will need to assess the accuracy across different markets.

Creating a new underage reporting button in the app and web

This is another critical piece of our age-assurance strategy.

Allowing users to report underage users will help us in the detection of minors on the platform.

We must wait until we have the new policies in place because we are now disclosing to users whether we took action on the content they reported. As of now, the policies are too lenient and it will raise many red flags as to why we did not remove an obvious child.

In order to make our policies more strict, we need to launch an appeals process to give users a way to prove their age.

Age-gating 100% of existing users

This is a less critical component of our age assurance strategy because it has launched in all p0 markets.

We can definitively say that users have undergone an age-gate

We can use this data to restrict certain features from underage users

This will have a large business impact in regions where regulation isn't as scrutinized as areas that have undergone an age-gate

Minor's Privacy and Safety by Design

Ensure TikTok creates a safe experience for users of different ages across the design and product process, while balancing the right of minors to express themselves and ensure that privacy settings and data usage are legally compliant for minors.

The following section will address these areas of concern:

How do we stop minors (<18) from behaving inappropriately on the app?

Our policies exceed industry standard on the types of behavior and content we allow from minors.

Livestream and DMs are being accessed by underage users who circumvent the age-gate, but once we can tighten up our policies we are hopeful we can make big gains in removing underage users (or removing access to features).

How do we prevent harm to minors (<18) in the app? How do we support the mental wellbeing and healthy development of minors (<18)?

In the US, U13's can only use a Kid's Mode version of TikTok that offers very little product engagement and especially curated content.

TikTok restricts access to certain features to ensure minor users are not exposed to high risk areas of the platform including DMs, Livestreaming, and lack of group chat. Note: proposals to remove these restrictions must be met with increased safety measures (ex: ability to review DMs, image hashing, grooming models).

Access to these features is determined by the age-gate meaning that our Age-Assurance strategy is an important component of our overall safety strategy.

Underage users in EEA are automatically prompted to make their accounts private

High risk ad content is restricted for underage users

We reduce the visibility of minors on the platform (for example U16s don't appear on the FYF and appear at the bottom of list for co-host opportunities on Live) and we mitigate risk in product design (for example by not facilitating inappropriate connections between minors and adults).

Minors also benefit from our more stringent safety measures overall. For example, we do not allow images to be sent in DMs and DMs are not encrypted. Note: proposals to remove these restrictions must be met with increased safety measures (ex: ability to review DMs, image hashing, grooming models).

How can we improve on this?

These are all critical pieces of work in our Minor's Safety and Privacy by Design program that will result in us fundamentally shifting the way we work with Feature teams and will result in a Minor-first safety strategy that can position TikTok as a leader in Minor Safety.

Revamping our default settings for minor users.

Making a recommendation to Product on how/if we recommend minor users to others.

Continuing to build strong partnerships with parents and fold in their feedback/concerns into our product design process.

Ensuring Minors can make well-informed decision about their data and privacy.

Incorporate Age-Appropriate Design Framework and clearly understand what is and isn't.

developmentally appropriate and design features accordingly.

Ensure we are considering the impact on younger users in a systematic way across every product PRD.

Working together with EU teens on product and safety feature recommendations.

Review of how U18 users data is shared and proposals for minimisation.

Physically separate kids mode account data from current normal mode users and limit data sharing for Kids Mode data.

Adding a new section to the Youth Portal that most simplifies our Privacy Policy for younger users

Redesign/update of the Safety Center.

Recommendation on whether to turn off notifications for U16s at certain times, including 10pm/bedtime and schooltime.

Audit of how we use data of minor users for nudges and notifications in an effort to extend their engagement for a longer duration.

Research on active/passive technology and psychological well-being.

Research on Profiling compliance for minors across regions.

Clarifying Issue leadership in Minor Safety issues across Feature. We need to be included as sign-off partners on any Feature work related to our Minor Safety strategies, and we must be able to drive strategy across Features such as Search, Hashtags, Comments, Livestream, and Account. As of now, these workstreams are too disparate and we're having trouble making changes to Features (ex: inability to add words to keyword lists).

We have a number of recommendations in the Predator Detection program to tackle mismoderation and over-moderation that will have an impact on this program as well (ex: building an investigations team, hiring a Policy Operations Manager for Minor Safety, building a team of specialized moderators, and building Minor Safety specific machine learning models. See below under Predator Detection for additional information.

Age-appropriate Content Classification & Curation

Ensure that content is appropriately restricted for different age-groups, that it is easily reportable, and that user-reporting in general reflects our commitment to Age-Appropriate design and Safety by Design.

The following section will address these areas of concern:

How do we differentiate types of harm that may affect users of varying developmental age differently (i.e., 13- vs. 16- vs. 18- years old).

How do we prevent inappropriate content from feeding to minors (<18) in the app?

How do we support the mental wellbeing and healthy development of minors (<18) by promoting age-appropriate content and suppressing inappropriate content?

How can we improve on this?

We recommend creating a TikTok Content Classification system that provides backend levers for TikTok to catalogue content according to thematic maturity (i.e., age-appropriateness) to ensure we offer minors a safe and comfortable app experience.

Our Age-appropriate Content Classification & Curation strategy has the potential to fundamentally improve the product as we know it, yet would likely require an entirely new moderation strategy or an entirely new department to lead its efforts. An advanced Content Classification system, if completed to a "gold-standard" degree has the potential to completely replace T&S policies with a more granular and sophisticated backend system that identifies harm, risk, and opportunity. It can inform what content is "safe" for a Kids Mode, Restricted Mode, L1-L3 user experiences, and other experience innovation. Content Classification and Advisories are not just mechanisms for restricting content; when leveraged well, they empower users of all ages to opt-in to an informed product experience by offering self-curation mechanisms at sign on. This type of system has the potential offer levers to support the following innovative product functionality:

Content curation that feeds user's content that is deemed developmentally appropriate for users across L1, L2, L3, Restricted Mode, and a potential global Kids Mode.

Provides users with autonomy over what content to "opt-in" to at sign up by allowing users to select a "content maturity" option that may be more conservative than his or her age (i.e., enhanced Family Pairing options).

Empower users to self-curate their own "safe space" by offering granular product experience at the account level (i.e., opt out of content w/ substances, violence, profanity, etc.).

Enable our global product to offer regionally relevant user experiences by filtering content based on regional compliance and cultural comfort (e.g., suppresses smoking from all L1 & L2 in Turkey, but from only L1s in France, etc.).

For regions that request (and eventually require) displayable maturity ratings (i.e., 13+, 16+, etc.) and/or content advisories (i.e., alcohol use, violence, etc.), we will have the quick backend ease to curate and allocate maturity displayables.

Offers nuanced backend signals for when content warrants a trigger warning, content blurring, "opt-in" message, help resources, or other overlays that provide users with additional information on potentially troubling OR uplifting content, etc.

Backend curation of content to be socialized or "resurfaced" during important cultural events, holidays, or days of national awareness.

Content Classification efforts require the work and coordination across many teams: Product Policy, Public Policy, PnP, Engineering, Product Feature Innovation (is this a team? It needs to be.), Legal, Ops, and Moderation, Content Creative, and more. The concept moves away from a punitive Trust & Safety "Is it bad? Take it down." approach to an empowering "What audience is this best suited for?" approach. Product workflows to include when considering this effort:

What teams within our org could benefit from content curation and what are the different ways in which they would leverage this functionality? (PnP?)

How do we catalogue the content? What are the classification parameters and how do we apply them? (Product Policy)

Are the classification parameters globally applicable, or is there regional nuance? (Product Policy, Legal)

How do we potentially translate current systems to adopt or inform the new system? (Engineering)

How can we first teach an army of moderators, and eventually machines to apply these classification "tags" with accuracy? (Product Policy *Tracy, QA / Ops)

In what regions are we required (or encouraged) to leverage this system? (Public Policy, Legal)

In what regions do we need product displayables on the content (ratings, advisories, trigger warnings, overlays, etc.)? (Public Policy, Legal)

How do displayables affect the product experience (i.e., when needed, how can we avoid obstructing content with the displayable) and how may regionally-specific characters, font, etc differ? (Product Feature Innovation)

What holidays or special events are happening globally, and do we want to "bubble up" content that honors these themes? (Outreach Partnership Management, Public Relations)

From a Minor Safety and EU legal perspective, building out a mechanism to hide or limit access to offensive content is critical for the safety of Minor users (and can be leverage by any user or other projects like TikTok TV). This is an enormous overhaul of our current policies and moderation process. It will have a large impact on headcount and capacity planning, given the effort affect many teams and given we may need to hire a specialized team of analysts to tag content with the maturity framework.

Predator Detection and Workflow Optimization

Build out best in class Minor Safety predator detection, including external partnerships with key stakeholders. Additionally, build out clear processes for remediation and optimize workflows. Evaluate product launches to ensure our products cannot be abused by predators, or expose minors to harm.

The following section will address these areas of concern:

How do we prevent harm to minors (<18) in the app? How do we support the mental wellbeing and healthy development of minors (<18)?

Automated reported to NCMEC

Better Image Hashing software (NCMEC and YouTube CSAI Match)

Predator models (US Only)

How can we improve on this?

These are critical pieces of our Predator Detection and Workflow Optimization strategies that will push TikTok to meet, and then exceed, industry standard across detection and moderation in our highest risk areas.

Creating grooming models to detect predatory behavior in DMs and push those users to review

This is a critical piece of work for our Predator Detection workstream

We'd like to explore a partnership with an industry database managed by Thorn and Microsoft to help us detect grooming in DMs.

We know that most predatory behavior occurs in DMs and that most predators and victims mutually follow one-another on the platform, meaning they can gain access to DMs if they have circumvented the age-gate.

Accessing DMs is a critical strategy for predator detection and we will work with LERT and Legal to clearly define when it is appropriate to access this content.

Further developing our relationship with Cairo to help us understand high risk trends outside of only video (ex: Livestream, predatory comments, and compilation accounts)

This is a critical piece of work for our Predator Detection workstream

We have only gotten to the tip of the iceberg in terms of the types of behavior Cairo can help us identify as the trial was only focused on video.

Minor Safety has unearthed many high risk trends in Livestream, comments, and at the Account level that we would like to continue to explore.

Investigations teams

This is a critical piece of work for our Predator Detection workstream

We need Minor Safety specific investigations teams to help us review findings from Cairo, review escalations from moderation, and to assist with escalations.

We're overloaded with manual review at this stage and I am currently manually reviewing accounts with Moe or trying to get QA and LERT resources to help ban large swaths of accounts.

We are heavily focused on rule-based detection strategies (because the volume of content for our highest risks areas is too low for a machine learning model). We need a team of people to help us review accounts for patterns and work closely with our Data Analyst to convert these findings into queries to identify more bad accounts.

We can choose how much latitude to give these teams, but for instance, on the more extreme side, if account A is suspected of predatory DMs against account B, an investigation team could look at all interactions from account A, rather than just those that had been reported.

Specialized Moderation teams

This is a critical piece of work for our Predator Detection workstream

Part of the reason Cairo has been successful at finding Minor Safety violations is that it can require specialized training to recognize signs of Minor Safety violations. While we can do more work to train our moderation teams in these Minor Safety signals, it's hard for moderators to be experts in everything and if they aren't specialized, they rarely see the same type of content more than once. One possibility to help catch more Minor Safety issues, especially more sophisticated attacks, is minor-safety specific moderation. User reports of Minor Safety issues would be directed to specialized queues. Machine learning systems would be trained to separate out Minor Safety issues. This becomes easier in Policy 4.0, when moderators start labeling issues by policy violation, which then allows us to get training data for specific policies, which then allows us to build models for specific policies or groups of policies.

Policy Operations Managers

This is a critical piece of our Predator Detection program.

Over-moderation and mismoderation are large concerns for the Minor Safety team. Having a Policy Operations Manager who's KPIs directly tie back to improving moderation quality specifically for our Issue Vertical will help bridge the gap between moderation and Policy.

Honey Pots

This is a less critical piece of our Predator Detection program as we are awaiting the onboarding of our Program Manager to fully assess this strategy.

We can consider creating "honeypots." We can hire adult actors who appear to be minors, send their videos to suspected pedophiles, and then have specialized investigators who engage with the suspected predators.

Increasing PR around Predator Detection

This is a less critical piece of our Predator Detection program as we are awaiting the onboarding of our Program Manager to fully assess this strategy.

We can consider publicizing some of the efforts in this area. For instance, after catching a certain number of pedophiles with honeypots, even if the number is small, we can publicize those efforts. This could backfire (increasing the perception that Tiktok is a place that pedophiles are attracted to) or could demonstrate our willingness to engage.

In addition, this will hopefully scare predators into avoiding TikTok, especially if we are public about our use of honeypots.

Global TikTok Kid's Mode

A great Kids Mode is one way to make minors safer. If we can produce a truly compelling experience, we can attract children to this experience, and keep them safe through feature restrictions and robust moderation.

The following section will address these areas of concern:

Can we push children to a more appropriate experience without banning their accounts?

We have TikTok Kid's Mode in the US where children are routed if they are underage

We're launching a partnership with Common Sense Media and they will vet 200,000 videos from TikTok to specifically tag whether they are appropriate for

What could we improve?

We are not yet able to make a recommendation on whether this program would significantly increase Minor protection at Tiktok for the following reasons:

Assuming we could make a very compelling Kid's Mode, we assume that children would be less inclined to lie about their age in order to access TikTok, but this assumption has not been tested. What we do know is that our own (underdeveloped) Kid's Mode has limited success and that Youtube Kids has also limited success; their app has about 1 million downloads worldwide in September and is ranked around 275, compared to Youtube with a ranking of around 8 ([source](#)).

We've had multiple discussions about expanding Kid's Mode globally, but this has not been raised as a recommendation by any regional markets, Legal, or Public Policy; rather the preference has been to simply remove underage users entirely and to focus our resources on enhancing the safety of our key platform (as opposed to diverting resources to a new platform). It has been communicated to Minor Safety that establishing a Kids Mode is too high risk considering PR and legal concerns as well.

We are happy to reevaluate based on new business decisions, but will need additional time to fully understand the risks and benefits. Our recommendation is to design a research initiative for the next bi-monthly to explore this further.

Below are some preliminary considerations:

A kids app would promote communication between people who already know each other in real life. Perhaps, users need an email address or phone number to add someone as a contact. Note: many pedophiles bulk download minor's phone numbers specifically to try and exploit these types of 'real world' connections.

This experience must have features that are safer by design. For instance, we might allow Live streaming of video games, but not the ability to Livestream using a front-facing or selfie-camera.

This experience must ensure that commenting is safe and positive by design. We could potentially have pre-set templates with positive emojis, canned phrases, stickers, or gifs.

This experience could have much simpler editing experience, but with more “fun” effects.

This experience will likely include content that is specifically designed for children, and may need to be acquired by paid creators or other third-party providers.

We don’t know enough about how children access TikTok. They could be accessing this from their parent's phones, their own phones, or the web. If they are accessing TikTok on their own device (without parental controls) or on the web, we will continue to have issues with underage users on the platform.

Getting our app to be a compelling experience would require a lot of investment and must not divert resources from our safety strategies on TikTok overall.

How are we helping minors spend time more wisely and meaningfully with guardians and educators?

We believe that focusing on wellbeing overall is a more effective approach to enhancing the health and happiness of minors on TikTok. As such, we propose that we continue our focus on limiting time spent in the app, creating safety videos that speak to minors, and in working more closely with parents and teens in the design of the app.

We sincerely believe that Trust & Safety does not have to be all about harm; we can surface the joy and knowledge on our platform to support the healthy development of youth growing up in a digital-media world.

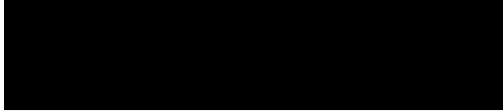
TikTok education initiatives are being piloted in the United States, and global should be considering complementary innovation approaches. For example, there is a TikTok Creative Learning Fund that is spearheading initiatives to consider how to leverage our content for educational purposes. Some initiatives are more “soft skills” (i.e., cooking, gardening, etc.) aimed at all audiences, yet others are leaning into our population of minor users to determine how TikTok itself can be leveraged as an online learning tool for educators. In the time of COVID19, educators, parents, and other youth advocates are desperate for creative ways to teach and support youth development. Consider the following innovation ideas:

Beyond Family Pairing, what if TikTok had Classroom Pairing capabilities for teachers or schools to leverage the platform for education?

Perhaps School Guidance Counselors and Psychologists could host accounts that teach how to cope with anxiety during covid, body positivity, etc. Pulling on the concept of content classification above, we can leverage identified risk themes into pro-social or civic awareness themes.

Beyond TikTok's content, which can be incredibly healthy and informative, the platform itself is one that captures students' attention. Consider the ways in which our features can be leveraged for academic development:; duets can be leveraged for "healthy debates," videos can be platforms for mini-speeches or classroom presentations, comments are digital discussion among learners, etc.

Probably lots more.



DOCUMENT METADATA

Bates Number: TIKTOK3047MDL-099-LARK-04568048-TIKTOK3047MDL-099-LARK-04568060

BEGATTACH: TIKTOK3047MDL-090-LARK-03501325

ENDATTACH: TIKTOK3047MDL-176-LARK-07538075

Attachment Count: 0

PRODVOL: MDL-099

Custodian: [REDACTED] FU, YUYI; MEDOFF, NATALIE

File Path: /DOCCNQK0ZRFS9Q85QT7GQVH7NLD.DOCX

Confidentiality Designation: CONFIDENTIAL

HASHVALUE: 3ED2402EABA6FF762A5494551B3CEA61

DocType: URL LINK DOCUMENT

Author:

Create Date: 10/10/2020 12:00 AM

Last Modified Date: 10/10/2020 12:00 AM

LAST MODIFIED BY:

TRACK CHANGES:

COMMENTS:

HASHIDDENATA:

Filename: DOCCNQK0ZRFS9Q85QT7GQVH7NLD.DOCX

Title: UNDERAGE USERS: DETECTION, SAFETY, PRIVACY, AND AGE-APPROPRIATE DESIGN

DOCEXT: DOCX

Email From:

Email To:

Email CC:

Email BCC:

Received Date:

Sent Date:

TIMEZONE:

Subject:

Case 4:22-md-03047-YGR

Document 2875-4

Filed 03/25/26

Page 16 of 16

THREADID:

REDACTIONS:

REDACTION TYPE: