TikTok response to Ofcom's call for evidence on video-sharing platform regulation

TikTok's approach to safety

At TikTok, our mission is to inspire creativity and bring joy.

- We do this by building a safe space for positive, creative expression
- Over a short space of time we have built a diverse and highly engaged community
- This has resulted in a unique space for creativity in many different forms – from music, humour and sport to beauty, comedy and magic.
- TikTok is a place where the emphasis is on ideas, rather than image.

TikTok's top priority is to promote a safe and positive experience so that everyone - from emerging artists to dancing grandparents - can feel free and empowered to express their creativity.

There are three main ways we do this:

1. Our Terms of Service and Community Guidelines reflect our values and establish the kind of behaviour we expect from our community. We enforce these rules using a combination of cutting-edge technology and thousands of safety experts based around the world.

2. We develop robust safety policies and features, including restricting direct messaging to over 16s and allowing parents to manage their teen's TikTok experience through Family Pairing. We actively promote these features to our users to ensure they have a genuine impact.

3. We collaborate with industry partners to make the digital world safer for everyone. We work with safety organisations such as Internet Matters and our signatories to industry-wide initiatives such as the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse, the Code of Practice on Disinformation and the Code of Conduct on Countering Illegal Hate Speech Online.

As a newer platform we are uniquely positioned to set the bar high for safety. We will never stop working to make TikTok a safe platform for our community, so that our users feel truly free to express themselves.

Part A: Questions for industry

Services

Question 1: Are you providing a UK-established service that is likely to meet the AVMSD definition of a VSP?

TikTok is a global, short-form video platform that provides its users with a vibrant, creative experience in a fun and safe environment. Our mission is to inspire creativity and bring joy. Millions of people come to the platform every day to express themselves creatively, share their talent, enjoy entertaining content, and engage with a diverse and global community.

TikTok appears to meet the definition of a VSP in the AVMSD as one in which "the principal purpose of the service, or of a dissociable section thereof, or an essential functionality of the service is devoted to providing programmes and/or user-generated videos, or both, to the general public, for which the video-sharing platform provider does not have editorial responsibility".

TikTok is provided to residents of the EEA and Switzerland by TikTok Technology Limited, which is registered in Ireland. We consider that Ireland is our centre of operations, main establishment and country of origin for the EEA for AVMSD purposes. TikTok is provided to residents of the UK by TikTok Information Technologies UK Limited, which is registered in England. We await the statutory instrument and guidance that we understand the UK Government will produce in order to give clarity on the basis on which VSPs are deemed to be established in the UK after the Brexit transition period.

Question 2: Is your service able to identify users based in specific countries and do you provide customised User Interfaces (UI), User Experience (UX) functionality or interaction based on perceived age and location of users?

TikTok provides a range of functionality, which allows us to customise the platform to our users based on their location and self-declared age.

Location

When users use TikTok on a mobile device, we collect information about their location in order to customise their experience. We infer their approximate location from their IP address. In certain jurisdictions, we may also collect Global Positioning System data. This information will sometimes be used to customise the user experience, since some features and settings are only available in certain regions. Users can set a choice of language when they create a TikTok account, and this will change the user interface accordingly.

In addition, the "For You" feed reflects preferences unique to each user. The system recommends content by ranking videos based on a combination of factors, including their country setting.

Age

When users open a TikTok account, they are required to enter their date of birth. Only users with a date of birth that indicates they are 13 years and older may create an account. Based on the self-declared age of users, some features will not be available. For example in the UK:

- Under 18s cannot purchase virtual coins, or send and receive gifts (gifts are used to show appreciation for content shared on TikTok);

- Only over 16s with over 1,000 followers can host a live-stream

- Under 16s cannot use direct messaging (all users are prevented from sending unsolicited messages and can't send images or videos as attachments);

- Under 18s are automatically prompted to make their account private when creating an account.

Terms and conditions (Ts&Cs)

Question 3: How does your service develop and enforce policies for what is and is not acceptable on your service?

Ensuring TikTok remains a safe and positive environment for our users is our top priority.

Through our Terms of Service and Community Guidelines (as defined below), we develop and enforce policies for what is and is not acceptable on Tiktok.

Clear Terms of Service

Our Terms of Service, which apply to all users in the EEA, include and go beyond the categories of harm listed in the AVMSD, stating that users may not:

"intimidate or harass another, or promote sexually explicit material, violence or discrimination based on race, sex, religion, nationality, disability, sexual orientation or age"

"use the Services to either intentionally, recklessly or negligently upload, transmit, distribute, store or otherwise make available:

- *any material which does or may infringe applicable laws or which infringes someone else's rights;*

- *any material which infringes privacy or personality rights of any other person or deceased person;*

- *any material which is defamatory of any person, obscene, offensive, pornographic, hateful or inflammatory;*

- *any material that would constitute, encourage or provide instructions for a criminal offence, dangerous activities or self-harm;*

- *any material that is deliberately designed to provoke or antagonise people, especially trolling and bullying, or is intended to harass, harm, hurt, scare, distress, embarrass or upset people;*

- *any material that contains a threat of any kind, including threats of physical violence;*
- *any material that is racist or discriminatory, including discrimination on the basis of someone's race, religion, age, gender, disability or sexuality*
- *any material that, in the sole judgment of TikTok, is objectionable or which restricts or inhibits any other person from using the Services, or which may expose TikTok, the Services or its users to any harm or liability of any type.*

Users must agree to our Terms of Service when they register for a TikTok account. Users are also required by our Terms of Service to comply with our Community Guidelines.

Strong Community Guidelines

Our [Community Guidelines](#) reflect our values and establish the standards we expect our community to adhere to. Our Community Guidelines were last updated on August 5th and describe in detail the categories of content that are prohibited on TikTok including:

- Dangerous individuals and organisations, including terrorist and terrorist organizations;
- Illegal activities and regulated goods;
- Violent and graphic content;
- Suicide, self-harm, and dangerous acts;
- Hate speech, including hateful ideology;
- Harassment and bullying;
- Adult nudity and sexual activities, including grooming behaviour;
- Minor safety;
- Integrity and authenticity, including misleading information; and
- Threats to platform security.

We are constantly reviewing our Community Guidelines to ensure they evolve with changing behaviour, mitigate emerging risks, and keep TikTok a safe place for our users.

Our Trust & Safety teams have developed detailed policies on each of these categories of content. We carefully review feedback from users and from our moderators, as well as quantitative information related to the application of our policies, in order to continuously improve these policies. We have teams of experts whose job it is to craft new content policies to make TikTok a safe and engaging platform globally. We frequently review our policies and enforcement criteria to check their effectiveness, to test moderators' understanding and assess fairness to users.

We invest significant resources into ensuring our Terms of Service and Community Guidelines are upheld, and are constantly reviewing and refining how they are applied to ensure they are fully achieving their goals. For example, at the start of the Covid-19

pandemic we introduced new policies banning misinformation likely to cause societal panic and real-world harm, this covered new types of conspiracy theories linking Covid-19 to 5G.

We enforce these rules using a combination of cutting-edge technologies and thousands of moderators around the world (as discussed further in our answer to Question 4).

Regional hubs for local policy development

Content moderation decisions can depend on context and require a sophisticated understanding of cultural nuances. To reflect this we employ local teams to design and enforce our policies.

We have set up an EMEA Trust and Safety hub in Dublin which enables us to work closely with communities in the UK to keep users safe. The regional hub focuses on strengthening our policies, technologies and moderation strategies and ensuring that they complement both local culture and context. It enables us to collaborate closely with regional regulators, policymakers, government and law enforcement agencies where required in the continued pursuit of promoting the highest possible standard of user safety.

Ensuring users understand our policies

We take a number of steps to ensure users are made aware of our trust and safety policies. The following resources are available in-app (via the Settings and Privacy section) and on our website:

- Our Terms of Service and Community Guidelines.

- Our educational video series, "You're in Control": the series presents TikTok's safety and privacy controls in an accessible and easy to understand fashion. We created this short-form video series, and involved several of our most popular creators, to educate users about safety in the TikTok format they're most accustomed to viewing. The videos can also be accessed directly in-app @TikTokTips.

- Our Youth Portal: teens and their families can learn about internet safety and the tools and controls built into TikTok in a user-friendly format. In the "TikTok community" section of the Youth Portal, users can find a simplified version of the principles that ensure TikTok remains a safe and positive environment, as well as a link to our more detailed Community Guidelines.

- Our Safety Centre: the centre contains resources to help users further understand our policies on topics including anti-bullying and impersonation. The Safety Centre also contains our Transparency Report, which provides insight into the volume and nature of content removed for violating our Community Guidelines or Terms of Service.

When people report potentially abusive content to our teams, or when we remove inappropriate content from our platform, we seek to communicate clearly to users to help them understand the rationale behind a removal to enhance understanding about our policies.

Drawing on external and user expertise and experience

As well embedding subject experts across our business, we also work with external advisors including NGOs and academics to develop and refine our policies and practices. For example, in the UK, we seek out the views of NGOs such as NSPCC, Internet Watch Foundation, Holocaust Educational Trust, Tell MAMA and Beat.

We also align our policies and practices with best practice principles on illegal and harmful content. For example, we are signatories to the European Commission's Code of Practice on Disinformation and Code of Conduct on Countering Hate Speech Online, and have endorsed the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse.

We also conduct research with users themselves. In the UK, we commissioned a series of focus groups with parents and children to seek their views on our safety strategy, including their perception of risk and harm on our platform.

Question 4: How are your Ts&Cs (or community standards/ acceptable use policies) implemented?

Users are required to adhere to the Terms of Service and Community Guidelines which set out the types of content and behaviour that is prohibited on the platform.

Our approach to moderation

Content that violates our Terms of Service and/or Community Guidelines is quickly removed, and accounts involved in severe or repeated violations are suspended or banned.

We have a range of systems in place to identify content that may violate our Terms of Service and Community Guidelines. We do this using a combination of human and automated processes. Through our innovative technology and expert review teams, we moderate content at scale and detect content and accounts that violate our Terms of Service and Community Guidelines. TikTok is deeply invested in making our moderation best-in-business. We will continue to improve our technology, but we also know that nuanced and scalable moderation cannot be effective without a continuous investment in teams of highly skilled moderation experts.

We also encourage users to take an active role in keeping the platform safe through simple and clear in-app reporting (for more information, see our answer to question 8).

Investing in enhanced detection of harmful content and accounts

After being uploaded, user-generated content (UGC) including videos and comments goes through a technology-based automated review process which is intended to flag content which may violate the Terms of Service and Community Guidelines, including illegal content. This is necessary given the volume of UGC posted to the platform every minute. The review process is trained to detect harmful content based on specific criteria involving a range of different automated systems.

These systems also take into account patterns or behavioral signals. We regularly study evolving trends, academic learnings, and industry best practices to enhance our systems continually.

Human moderation

Context and understanding of local cultural nuances is vital, so in addition to our technologies we are continuing to rapidly grow our expert team to cross-review and remove content and accounts that violate our Terms of Service and Community Guidelines. We have over 10,000 moderators globally, including thousands in Europe and over 360 in London.

Where appropriate, content is sent to trained human moderators for manual review. Trained moderators determine the appropriate remedial action to take, including in some cases removing it from the platform. We have moderation teams that speak 52 languages and are based all around the world. Having country-specific moderators ensures we have the necessary cultural context to review user reports and, where necessary, remove inappropriate content, regardless of where and in which language it is posted.

In appropriate circumstances, content may be referred to our Law Enforcement Response Team, which will consider whether the content gives rise to a law enforcement reporting obligation under applicable law. In the case of emergency situations, where the content indicates a credible and imminent risk of harm, the content will be escalated to our Emergency Response Team for appropriate action.

(See detail on sanctions and referrals in our answer to question 8)

Quality assurance

Human moderation is subject to quality assurance to ensure the accuracy and quality of the moderation decisions. We are actively monitoring and seeking to improve our accuracy to ensure that we are catching content that violates our Terms of Service and Community Guidelines. This includes partnership work with trusted organisations. Our systems are designed to ensure that the highest risk content, in terms of content type and its popularity, will be reviewed as quickly as possible.

Easy reporting

We have a reporting feature that enables users to report inappropriate content or behaviour they might be concerned about. Users can report content (such as videos, comments, audio, direct messages or a user's profile) if they believe it violates our Community Guidelines, including content believed to be illegal. Users can report content in our app and through our website by choosing a reason why they think the content might violate our Community Guidelines. Human moderators review every user report.

For inappropriate content that clearly violates our Terms of Service or Community Guidelines, we will take appropriate action, including removing or disabling access to content and/or accounts. We encourage our users to report any content that they find inappropriate. (see more detail in response to Question 8 below).

Non-TikTok users such as NGOs, government bodies and law enforcement agencies may also report content to us (for example by emailing legal@tiktok.com), after which the content will be reviewed by a human moderator.

Effectiveness and transparency

In our latest Transparency Report, we provide detail on the effectiveness and impact of our moderation against nine of our content categories. In the first half of this year (1 January 2020 to 30 June 2020), we removed 104,543,719 videos globally, which is less than 1% of all the videos our users created, for violating our Community Guidelines or Terms of Service. Of those videos, we found and removed 96.4% of videos before a user reported them, and 90.3% were removed before they received any views.

Advertising

Question 5: Does your service have advertising rules?

Advertising on TikTok is governed by a strict set of rules. Our policies for advertisers in the EU can be found [here](#) and [here](#) ("Advertising Policies"). In addition, advertisers on TikTok must comply with our Community Guidelines.

Our Advertising Policies are split into two different sections. The first is the industry entry section, which provides information about the products and services that can be advertised on TikTok. The second section provides guidance on the format of advertising allowed. When advertisers place an order, each advertisement is reviewed against these policies.

We constantly review our Advertising Policies in response to new trends and developments. For example, Since the outbreak of Covid-19, we've seen a rise in the number of problematic claims appearing in ads for diet products such as unhealthy or unattainable weight-loss slogans and references to body images. This led us to change our policies to put additional restrictions on weight loss and implied weight loss claims, including banning irresponsible claims and weight loss/management products/services that promote an unhealthy body image or relationship with food. In addition, weight management products can only reach an 18+ audience.

We have taken swift action to update our Advertising Policies to prevent misinformation about Covid-19 and conspiracy theories on the platform.

Our Advertising Policies cover content, products and services banned in all jurisdictions as well as jurisdiction-specific (including the UK) restrictions.

The restrictions include prohibitions on the advertising of:

- Tobacco products including e-cigarettes;
- Drug-related products and services;
- Weapons;
- Police and military equipment;
- Gambling;
- Sexual products and services;
- Political ads;
- Counterfeit goods;

- Products and services that enable dishonest behaviour;

- Illegal services and activity;

- Discriminatory content;

- Sensational and shocking content;

- Content that violates minor safety/protection; and

- Any other content that violates our Community Guidelines.

UK specific industry restrictions include:

Prohibited Products/Services

- Ads promoting alcoholic beverages (wine, beer, spirits, etc.), alcohol clubs/subscription services, alcohol making kits, or alcohol sponsored events;

- Ads for products or services that are specifically intended for or appeal to children including toys, games, apps, and clothing. Advertisements marketed specifically toward children, even if the product may be for a general audience, are also prohibited;

- Ads promoting single securities or other tradable financial assets, bail bonds, debt assistance programs, get rich quick schemes, pyramid schemes (including non-financial pyramid schemes) and multi-level marketing, penny auctions, binary options, cryptocurrency wallets, unaccredited digital banks that perform any traditional bank-like function, cryptocurrency credit or debit cards, initial coin offerings, token sales, or other means of promotion or advertisement of individual digital currencies or tokens;

- Ads promoting fantasy sports, bingo, lotteries, or other gambling related content;

- Ads promoting prescription drugs;

- Ad promoting weight loss/management fasting products or services; and

- Ad promoting weight loss/management supplements (including fat burning pills, appetite suppressants, weight loss/detox teas or lollipops).

Restricted Products/Services

- **Financial Services -** Ads must comply with applicable laws, regulations, licensing obligations, and industry requirements. For all ads within this category, we may ask for disclosures and evidence from appropriate regulators and/or third-party accreditations or certifications. Ads may only target users aged 18+ and must include proper disclaimers;

- **Food and Drinks** - Ads promoting foods or drinks which are classed as high in fat, salt or sugar (HFSS), as identified using nutrient profiling, may be subject to restrictions. Ads may only target users aged 18+ and must include proper disclaimers;

- **Pharmaceuticals, Healthcare, Medicines** - Ads for pharmaceuticals, healthcare, and medicines must target users aged 18+ and comply with local laws of the target country. Ads promoting over-the-counter drugs must be approved by the regional regulating authority and may not contain misleading or inaccurate claims. We do allow advertisements for brick and mortar pharmacies and online pharmacies that are certified by a third-party licensing organization such as the GPhC or regional equivalent. Proof of licensing may be necessary for approval;

- **Dating Applications or Services -** Ads that promote dating applications or services must target users aged 18+. Such ads must not be overtly sexual or reference transactional companionship or cheating. Dating ads should not be sexually suggestive in nature (e.g. focusing on bodies, sexual language or innuendo, etc.); and

- **Media and Entertainment** - Ads that promote films, TV shows, or games must not contain profanity and crude language, sexually suggestive language or images, graphic violence, or drug use, and must be targeted at the appropriate age group.

How we enforce our Advertising Policies

When advertisers place an order, each advertisement is reviewed against our Advertising Policies. The ad review process involves vetting the products/services promoted, ad caption, text, images, audio, visuals, age/region targeting and landing pages. The turnaround time for ad review is typically less than 24 hours, although it may take longer in some cases.

Advertiser accounts and ad content are held to our Advertising Policies. They must follow our Community Guidelines and Terms of Service. If a proposed ad does not comply it will not be allowed to appear on TikTok. If an advertiser violates our policies, TikTok might ban the advertisers from our platform.

Question 6: How far is advertising that appears on your service under your direct control, i.e. marketed, sold or arranged by the platform?

There are two main ways an advertiser can conduct a paid-for advertising campaign on TikTok, and some advertisers may use both routes simultaneously:

1. An advertiser can register and create an account on the TikTok ads tool. Advertisers can then utilise the self-serve ad platform to bid for and purchase in-feed advertising and manage their own campaign. This method of advertising is referred to as "Auction Ads".

2. An advertiser can recruit TikTok to manage an advertising campaign on its behalf by contacting the TikTok brand partnership team. TikTok has teams across Europe who partner with creative and media agencies, and directly with brands, in order to offer platform education, media and creative guidance, and account and campaign management. TikTok's brand and creative strategy teams respond to briefs and work in collaboration with partners to plan the best campaigns for any brand's marketing goals. This type of advertising is referred to as "Reservation Ads".

An advertiser or a media agency must [create an account](#) on the TikTok ads tool and sign a Data Sharing Agreement in order to purchase Auction Ads or Reservation Ads. In addition, the advertiser must sign a Framework Agreement if they will be purchasing Reservation Ads.

When using the TikTok ads tool to purchase in-feed ads through the Auction Ads route, the advertiser can determine their campaign by targeting a certain audience and the duration for which they want their campaign to run.

If an advertiser has chosen the Reservation Ads route, campaign management is carried out by a TikTok brand partnership team which manages the campaign according to the advertiser's budget and strategy. TikTok will reserve the advertising for the advertiser and undertake tracking and analysis for them.

As explained in our answer to Question 5, when advertisers place an order, each advertisement is reviewed against our Advertising Policies. We determine the size, placement, and positioning of ads. For Hashtag Challenges, we offer a premium service for advertisers, whereby the pinned videos on the Hashtag Challenges page are moderated by our ad reviewers according to the advertisers' Brand Safety Requirements. More broadly we are exploring contextual targeting capability for advertisers and looking to launch brand safety tools.

TikTok also operates the TikTok Creator Marketplace (TCM) which acts as an introduction platform for advertisers to contact creators to create content.

Question 7: What mechanisms do you have in place to establish whether videos uploaded by users contain advertising, and how are these mechanisms designed, enforced, and assessed for effectiveness?

Creators are under legal obligations to ensure that any advertising they have arranged is properly disclosed and ASA guidance requires creators and influencers to disclose paid partnerships. Creators on TikTok often do so by adding #ad on their videos. We do not have mechanisms in place to establish whether videos uploaded by users contain advertising. We understand that Ofcom will produce guidance on the appropriate measures VSPs are required to implement under the Audiovisual Media Services Regulations 2010.

Reporting and flagging

Question 8: Does your service have any reporting or flagging mechanisms in place (human or automated)?

Encouraging our community to use our reporting tools

We provide users with a reporting mechanism that enables them to report inappropriate content or behaviour they might be concerned about. TikTok users can report content (as detailed below) which they believe may violate our Terms of Service and/or Community Guidelines, or other content they are concerned about, both through our app and on our

website. When reporting, users can either select from a list of reasons why they think something might violate our guidelines (such as violence or harm, harassment, or hate speech) or select "other".

Content users may report includes:

- another's user's profile for reasons including pretending to be someone else, posting inappropriate content, having an inappropriate profile picture or any other reason;

- another user's videos for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- the audio used in another user's videos for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- comments on a user's video for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- a direct message from another user for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- a user during a livestream for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- Hashtags for suspected violations that align with the categories of prohibited content set out in our Community Guidelines;

- Ads for suspected violations that align with the categories of prohibited content set out in our Community Guidelines.

We have designed our reporting mechanisms to be easily found and user-friendly, and are continuously monitoring their effectiveness and exploring ways to improve them.

Awareness of our reporting mechanisms

We make our users aware of our reporting features through a number of channels. These include the TikTok app itself, an explainer in the "Safety Center" section of our website, our Youth Portal, the "For Parents" page and our "You're in Control" safety videos.

Reporting process for users

Users can report content via our website or in app. In app, users can access our report function to report a video, comment, hashtag or sound via our share icon. Videos and comments can also be reported by simply holding a finger down on the video or comment itself. Direct Messages and Accounts can be reported using the "..." icon at the top right of the screen. We take all reports from users seriously. Every report is reviewed by human moderators and we take action on reports quickly.

Reported content is placed in a queue for our moderators to review, and the moderator can see the reason that the user gave for reporting the content. The user is informed about the outcome of their report on a video through a message to their inbox.

Action taken when our Community Guidelines are violated

If content (video, comment, livestream, hashtag or sound) is found to violate our Terms of Service and/or Community Guidelines following a user report, we will remove it from our platform. If a user profile (such as their picture or biography) is found to violate our Terms of Service and/or Community Guidelines we will reset the user's profile.

Depending on the circumstances, we may also take appropriate action on the user's account. This can include disabling the user's access to certain features (such as sharing videos, using direct message, launching livestreams) for a certain period. Severe or repeated violations of our Terms of Service and/or Community Guidelines will lead to an account suspension or a permanent account ban.

Depending on the circumstances, we may also take appropriate action on the user's account. This can include disabling the user's access to certain features (such as posting videos, using direct message, posting comments) for a certain period. Severe or repeated violations of our Terms of Service and/or Community Guidelines will lead to an account suspension or a permanent account ban.

In appropriate circumstances content may be referred to our Law Enforcement Response Team which will consider whether the content gives rise to a law enforcement reporting obligation under applicable law in a particular jurisdiction. In the case of emergency situations, where the content indicates a credible and imminent risk of harm, the content will be escalated to our Emergency Response Team for appropriate action. For the most serious circumstances relating to minor safety, we also work with the National Center for Missing and Exploited Children (NCMEC) and law enforcement agencies.

Safeguarding freedom of expression

Balancing freedom of expression with safety is top of mind for all of the teams that create our content policies. We continuously review our existing guidelines, look for feedback from external experts as well as from our teams, and make adjustments to make sure we keep balancing these interests. In the past few months we have engaged with external NGOs and academics in the US, Europe, APAC etc. on a range of policies, to seek feedback on our approach and forthcoming policy proposals. We have also referred to documents such as the UN Convention on the Rights of the Child when it comes to balancing protection and participation in our approach to minor safety.

We recognise that content which may be controversial can have value to the public. Our Community Guidelines allow exceptions under certain circumstances for content that would otherwise be prohibited if it provides value to the public, such as educational, scientific, artistic, satirical and newsworthy content.

If users believe that we have made the wrong decision to remove content, they may appeal and ask us to reconsider.

Notifying users of removals and allowing for appeals

If a user's video is removed because we believe it violates our Terms of Service and/or Community Guidelines, we will inform the user that we have taken down their content for this reason. The user is given the in-app option to indicate that they believe we have made

a mistake and should reconsider. If the user makes such a request, a human moderator will consider the appeal and if it is determined that the content violates our Terms of Service and/or Community Guidelines, the content will not be reinstated. If the moderator determines that the content does not violate our Terms of Service and/or Community Guidelines, it will be reinstated to its original status on the platform.

If a user's video is removed because we believe it constitutes infringement of intellectual property rights, the user may submit a Counter Notification Form and their application will be handled by a specialist team.

If other action is taken on the user's account, for example a suspension or an account ban (for severe or persistent violations of our Terms of Service and/or Community Guidelines), the user may ask us to reconsider our decision. The matter will subsequently be considered by a human moderator or escalated as appropriate. If we decide that the action was not necessary, we may reinstate the user's account.

Quality assurance

We run regular quality checks and have a Quality Assurance team in place to check the accuracy of automated decisions. We also conduct sampling to check for false negatives that we may have missed. Detection teams also sample our automated processes to ensure these are subject to human scrutiny and remain as effective as they can be. As well as checking the accuracy of automated decisions, our Training and Quality Team ensure that our human moderators have the knowledge and support they need to fulfill their roles.

Content rating

Question 9: Does your service allow users to rate different types of content on your platform?

TikTok does not have a rating system. However there are a number of ways in which users can indicate they have enjoyed a piece of content. They can like it, share it, comment on it and save it to their favourites. Similarly, they can also tell TikTok when they are not interested in a certain type of content. Our recommendation system learns from all these signals and others what content a user is likely to enjoy.

Age assurance

Question 10: Does your service use any age assurance or age verification tools or related technologies to verify the age of users?

Our age-assurance policies

TikTok users have to be 13 years old or older to use our platform. By using TikTok, users confirm they are over the relevant age (our Terms of Service state that people "may not… access or use the Services if you are not 13 years or older"). We use an age gate on registration and moderation techniques to make sure TikTok is only used by those who meet our age requirements.

Users are required to enter their date of birth when creating an account, and we prevent registration if individuals do not meet the minimum age of 13 years. All of our users in EEA countries have passed through this age gate. A prospective user who says they are under 13 is sent from the enter your DOB page back to the registration page, with a 24-hour block in place to prevent the prospective user from immediately re-entering a different DOB.

If we become aware that someone under 13 is using the Services, we will terminate and permanently block that user's account.

We have also given TikTok a 12+ App Store rating and a Google Play Store rating of "Parental Guidance Recommended." Parental controls on each App Store allow parents to block apps with such ratings on their children's devices.

Our service is also tailored in a number of ways to meet age-appropriate needs, as set out below:

- Under 18s cannot purchase virtual coins, or send and receive gifts (used to show appreciation for content shared on TikTok);

- Under 16s cannot host a live-stream (all users must also have 1,000 followers);

- Under 16s cannot use direct messaging (all users are prevented from sending unsolicited messages and can't send images or videos as attachments) ;

- Under 18s are automatically prompted to make their account private when creating an account.

Next steps

During the transition period for the Age Appropriate Design Code, we are working towards a strategy for establishing the age range of our users with a level of certainty appropriate to the risks and taking into account a child's right to privacy.

This is an industry-wide challenge. We know there is more work to be done in the area of age-assurance and are committed to working collaboratively with industry peers, regulators, and key stakeholders to find a solution that ensures our platform is only used by those who meet our minimum age requirements.

Parental controls

Question 11: Does your service have any parental control mechanisms in place?

Our parental control tools and policies

To give parents' greater control over their kids' TikTok experience, we recently introduced Family Pairing which enables a parent or guardian to link their TikTok account to their teen's to provide greater visibility and oversight about the way their teen uses our platform.

Family Pairing allows a parent or guardian to link their TikTok account to their teen's and set controls, including:

- Screen Time Management: Control how long teens can spend on TikTok each day. In addition, TikTok has partnered with top creators to make short videos that appear right in the app and encourage users to keep tabs on their screen time. These fun videos use the upbeat tone our users love while offering suggestions to take a break and do something offline, like read a book;

- Restricted Mode: Limit the appearance of content that may not be appropriate for all audiences. Even without Family Pairing enabled, parents can help their teen use Screen Time Management and Restricted Mode by visiting the app's Digital Wellbeing controls at any time;

- Direct Messages: Restrict who can send messages to the connected account or turn off direct messaging completely (note that direct messaging is not available to under 16s);

- Search: Parents or guardians can choose to turn off the search functionality for their child.

Our "For Parents" page is our chief means of making parents aware of how TikTok works, how to keep their teens safe on the platform and the parental control tools available to them. However we know that parents also seek online safety information from trusted third parties. We therefore work with online safety charities such as Internet Matters to develop safety resources. Internet Matters' Parents Guide to TikTok has been viewed over half a million times and the charity's guide to our Privacy Settings has been viewed over 100,000 times.

Complaints handling

Question 12: Does your service have a complaints mechanism in place?

The functionality by which users can report another users' account, content shared by other users, or direct messages sent by other users, is described in our answer to Question 8 along with a description of the appeals process when users' content is removed.

If users wish to contact us or complain about any other issue (including problems with their account or system issues) they can "report a problem" in the "Support" section of the in-app settings. The user will then receive a reply in their in-app inbox. Users can also contact us by filling out this form.

Media literacy

Question 13: What media literacy tools and measures are available on your service?

We actively promote safety features to our users so they know how best to stay in control of their TikTok experience.

The Safety Centre

Our website features a range of media literacy tools and resources to keep our users safe and informed. These are brought together in our Safety Centre, which provides information and advice on our safety features and measures, including a page for parents who want to find out more about the platform. Key resources include:

- **Anti-bullying:** Tools, resources and information for users affected by or concerned about bullying;

- **Covid-19**: Information from trusted partners including the WHO on Covid-19 and the measures TikTok has taken as a platform during the pandemic;

- **Suicide hotlines:** International hotlines for users that are seeking support for themselves or someone they know in relation to suicidal thoughts and behaviour;

- **Videos**: The educational video series "You're in Control" and "Be Informed" (described in question 8 above) help us in our mission to foster a safe, positive environment, where users are in control and can express themselves creatively;

- **Partnerships:** Links to the websites of our safety partners, where users and parents can find tips and support for staying safe online.

The Youth Portal

The Youth Portal offers both in-app tools and educational content to provide users with all the information they need to have the best possible experience as part of the TikTok community, including age appropriate advice for teenagers on our key safety and privacy features and our Community Guidelines. Designed to help teens learn about digital safety and help parents build ongoing dialogue with their families, the Youth Portal covers the basics of digital literacy suitable for all ages, including:

- **Internet Security:**
    - Tips on what you should keep private and not share;
    - How to spot suspicious behaviour.

- **Personal Privacy:**
    - Privacy settings to define your public presence;
    - How to set comment, DM and Duet controls;
    - How to block users;
    - How to enable Family Pairing.

- **Community best practices, including:**
    - How to be kind and positive online;
    - Our code of conduct outlined in our Community Guidelines;
    - Tips on being a creator.

TikTok proactively promotes safety videos in users' For You feeds. These videos reinforce our Community Guidelines and offer users mini 'how to' tutorials including how to: block a

user; report inappropriate behaviour; filter comments; make an account private; and set screentime limits.

Reporting and evaluation

Question 14: Do you publish transparency reports with information about user safety metrics?

Sharing our progress and achievements in ensuring user safety with a wider audience is a major priority for TikTok. To this end, we publish Transparency Reports every half year. The reports set out our approach and policies to protect the safety of our community, including how we establish and enforce our Community Guidelines, how we empower our community with tools and education, and the volume of videos removed for violating our Community Guidelines

TikTok published its third Transparency Report on 22 September 2020 for the period 1 January 2020 to 30 June 2020. This included information on:

- Promoting safety and community on TikTok including information about our policies and public commitments;

- How we connected users to authoritative and educational content;

- How we have worked to enhance youth safety;

- The partnerships we have developed with safety organisations;

- The volume of videos removed for violating our Community Guidelines and percentage of total videos this represents, including what proportion were caught and removed before a user reported them;

- Content moderation metrics against nine of our content policy categories;

- Explanation of terminology and definitions used in our moderation strategy;

- Details of the countries/markets with the largest volume of removed videos;

- Details of the volume of legal requests we received for user information from law enforcement in the first half of 2020, broken down by country, and how we responded;

- Details of the volume of legal requests we received for user information from government agencies in the first half of 2020, broken down by country, and how we responded;

- Details of the volume of copyright takedown notices received and the percentage of requests where some content was removed.

Question 15: What processes and procedures do you have in place to measure the impact and effectiveness of safety tools or protection measures?

We refer to our earlier answers which describe our safety tools. We have metrics in place to understand, track and drive down instances of abusive content and behaviour as outlined above.

Question 16: How do you assess and mitigate the risk of inadvertent removal of legal or non-harmful content?

We recognise that content which may be controversial can also have value to the public. Our Community Guidelines allow exceptions under certain circumstances for content that would otherwise be prohibited if it provides value to the public, such as educational, scientific, artistic, satirical and newsworthy content.

Content moderation can in some instances lead to the inadvertent removal of legal or non-harmful content. Given the sheer volume of content, some degree of over-moderation is inevitable whether it is by human or automated review, but we work continuously to spot patterns of over-moderation in order to learn and adjust our training process or technology.

A user may ask us to reconsider where we have removed their content or taken other action in respect of their account. We have provided details about this process in the answer to Question 8.

Question 17: Have you previously implemented any measures which have fallen short of expectations and what was your response to this?

TikTok is a global community for creativity. Our platform empowers users to express themselves, inspire others, and have fun in a safe and welcoming environment. We are always looking for ways that we can enhance users' experience on the app including, crucially, their safety.

To cite just one example, at the end of last year we reviewed our policies and protections around virtual gifting. This feature allows users to send virtual gifts to creators during a live-stream, and it is one way users can show their support for and reward their favourite creators. Although the feature is overwhelmingly used positively, we know we have a responsibility to improve our features and policies to protect against misuse. In addition to the existing 16+ age limit for a user to host a live-stream, our updated policy restricts the ability to purchase, send and receive virtual gifts to users aged 18 and over. We made these changes to foster a safe environment where users who are over 16 years can enjoy a live-stream without encountering misuse, such as any pressure to send virtual gifts.

As with the example of virtual gifting, we are constantly reviewing the different components of the TikTok platform to ensure they are fit for purpose, and we would not hesitate to make changes should we see the need.

Question 18: How does your service develop expertise and train staff around different types of harm? (e.g. do you have any partnerships in place?)

At TikTok we place great emphasis on hiring staff with the latest and most relevant expertise, and on ensuring that this knowledge remains up to date with the latest trends.

We have hired staff with deep expertise in counter-terrorism, minor safety, tackling hate speech, and suicide/self-harm prevention from a range of backgrounds including academia, public policy and the technology industry. We strive to ensure that those responsible for specific policy areas have a deep understanding of the relevant harms both offline and online. These teams work in a collaborative way with our operations teams to ensure that there is a rapid feedback loop in relation to new trends and developing risks of harm.

In addition to this, our Trust and Safety teams have partnerships and work with a wide range of external stakeholders and initiatives to keep abreast of developments, develop expertise, and identify new potential risks and harms. These stakeholders and initiatives include, but are not limited to:

- The Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse, adopted by the Five Eyes alliance;
- The EU Code of Practices on disinformation and hate speech (respectively);
- WeProtect;
- Internet Matters;
- NSPCC.

We also learn from the TikTok community, including from the focus group sessions that we run with TikTok users.

Our training & learning system includes clear and frequent communication channels between policy team and moderators, with training and QA teams acting as the bridge.

The training team will work closely with the policy team to design, prepare and assess training lessons materials. Policy training to moderators, or other trainers, is conducted by our certified trainers, making sure the training is productive, effective and enjoyable. This includes regular testing of moderators during the learning process to ensure the quality of the knowledge delivery.

Moderators are encouraged to raise their concerns, doubts and advice in the learning cycle. Regular Q&A and alignment sessions allow a free flow of information between moderators and trainers & QAs. This helps ensure that policy enforcement is consistent across the teams and the policy could be enhanced by receiving last information from the frontline.