



## DSA RISK ASSESSMENT REPORT 2025

Confidential

28 August 2025

---

### CONTENTS

1. FOREWORD	2
2. EXECUTIVE SUMMARY	4
3. INTRODUCTION	5
4. ABOUT TIKTOK	7
5. RISK GOVERNANCE	11
6. AIGC AND RECOMMENDER SYSTEMS	17
7. ILLEGAL CONTENT: RISKS OF ILLEGAL HATE SPEECH CONTENT	22
8. ILLEGAL CONTENT: RISKS OF TERRORIST CONTENT	27
9. ILLEGAL CONTENT: RISKS OF INTELLECTUAL PROPERTY INFRINGING CONTENT	32
10. ILLEGAL CONTENT: RISKS OF CHILD SEXUAL ABUSE MATERIAL AND CHILD SEXUAL EXPLOITATION AND ABUSE	37
11. ILLEGAL CONTENT: RISKS OF GENDER-BASED VIOLENCE CONTENT	42
12. ILLEGAL CONTENT: RISKS OF ILLEGAL AND UNSAFE PRODUCTS	47
13. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO AGE-INAPPROPRIATE CONTENT	51
14. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO AGE-ASSURANCE	55
15. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO ONLINE ENGAGEMENT	59
16. CIVIC INTEGRITY AND MISINFORMATION: RISKS TO ELECTIONS AND CIVIC INTEGRITY	65
17. CIVIC INTEGRITY AND MISINFORMATION: HARMFUL MISINFORMATION	71
18. FUNDAMENTAL RIGHTS: RISKS TO FUNDAMENTAL RIGHTS	78

---

## 1. FOREWORD

TikTok presents its third annual Systemic Risk Assessment Report pursuant to the framework of the EU Digital Services Act (“**DSA**”). The DSA has dual aims, which we are proud to contribute to: the protection of a safe, predictable and trusted online environment, and the protection of fundamental rights, including freedom of expression. We uphold freedom of expression by providing a space for our users to create, find community and be entertained, while at the same time striving to make our Platform safe for all of our users. This is reflected in our approach to safety, which spans policies, product, practices and partners which together create a space where creativity and free expression can thrive.

To protect the integrity of our Platform and to keep users safe, we invest in advanced moderation technologies as well as dedicated safety personnel. These efforts ensure that we continue to manage risks and keep our users safe as our community grows. We use many technologies (for example, vision-based, audio-based, text-based and LLM-based technologies) to detect violations of our rules and to moderate content on the Platform. We employ thousands of safety professionals globally who help build our technologies, identify risks, develop our policies and design new features to keep our community safe. We also work closely with external stakeholders to inform our approach to risk identification, management and mitigation. This includes our ongoing engagement with our European Safety Advisory Council, which is made up of diverse experts on systemic risks, as well as our partners and other subject-matter experts. We also continue to invest in our Youth Council, which we have almost doubled in size in 2025, and which provides our Younger Users with a voice in the decisions that affect them. Our work with the Youth Council is just one of the ways we continue to learn, adapt and strengthen our approach to youth safety.

We continue to innovate and find new ways for our community to connect and engage with each other. Late last year, we launched TikTok Shop in a number of EU member states, which provides new opportunities for businesses and independent sellers to market and sell their products on TikTok. On Shop, Adult Users can discover and purchase physical products directly on our Platform. We assessed how the introduction of TikTok Shop could impact the Article 34 systemic risks prior to launch and have reflected our ongoing assessment of such risks in this Report, including through our new dedicated Risk Assessment module related to the promotion and sale of illegal or unsafe products.

We recognise that our Platform is one of a number of online platforms where users may be exposed to different types of risks, including from bad actors who seek to circumvent our policies and moderation measures. TikTok acknowledges its position of responsibility to protect users against those risks and we continue every day to build on our extensive efforts to keep our users safe and to ensure that our Platform has a positive impact on our users and society.

One challenge being faced across the online world relates to the growing sophistication and prevalence of content created using Artificial Intelligence (“**AIGC**”), which is posing challenges for governments, policy makers, online platforms and members of the public.

## 1. FOREWORD

We recognised this risk in our Year 2 Risk Assessment Report, where we detailed some of the policies and practices that we have introduced with a view to addressing it, and have continued to give AIGC further consideration as part of our Year 3 Risk Assessment process.

We have also retained our focus on the potential impact of our recommender systems on the DSA systemic risks and have carefully appraised the cross-risk mitigations that we have in place to address those risks. Since our Year 2 Report, we have prioritised user empowerment, by augmenting the tools available to our users to customise content they see in their For You Feed. This has included the launch of our industry-leading Manage Topics Tool and the enhancement of our keyword filtering measures, which give users greater control over content recommendations.

Against the background of numerous elections across the EU over the last 12 months, TikTok has continued to invest in and enhance its measures to address elections and civic integrity risks that can arise where bad actors attempt to misuse our Platform. Our Elections Integrity Programme helps us to detect, manage and mitigate risks relating to misinformation during elections. Since August 2023, TikTok has worked with electoral commissions and civic society organisations (including fact-checkers) to introduce in-app Elections Centres which connect users with authoritative voting information, including when, where and how to vote. We have strengthened our transparency reporting efforts by regularly updating our community on our efforts to ensure that our Platform is not used to disrupt the integrity of elections in the lead up to and during elections. Since January 2025, we have published information relating to our efforts to protect the integrity of elections in Poland, Portugal, Germany, Romania and Croatia on our Global Elections Integrity Hub. In the Hub, we detail our efforts to work with local authorities, NGOs and fact-checking partners, our extensive efforts to monitor for inauthentic and deceptive behaviour, including covert influence operations, and to remove content that violates our Community Guidelines.

We continue to enhance the protections we put in place to protect our Younger Users across different risk areas and to promote mindful use of the Platform. Over the last 12 months, we have improved the information and resources we provide to Younger Users and their parents and guardians to promote safety and digital wellbeing on the Platform. For example, we have updated our online Teen Safety Centre, our Wellbeing Guide and our Guardian's Guide and have launched new tools for parents and guardians to manage the time their teenagers spend on TikTok, including enabling parents to schedule times when the app is not available for their teens (for example, during school hours). Additionally, we have added features that are automatically applied to accounts for Younger Users, like new features that help Younger Users wind down and switch off in the evenings.

Our efforts to protect our Platform do not stop at the conclusion of this year's annual Risk Assessment process. For example, in July 2025, we have continued these efforts by launching new trust and safety tools for creators, families and the TikTok Community. We

## 1. FOREWORD

will continue our extensive efforts to keep our Platform safe and look forward to further engagement with policy makers, regulatory bodies and civil society in due course.

**Adam Presser**  
**Director, TikTok Technology Limited**

## 2. EXECUTIVE SUMMARY

### Introduction

This DSA Risk Assessment and Mitigation Report (the “**Report**”) summarises the results of TikTok’s third annual Systemic Risk Assessment within the framework of the DSA. This Report uses the terms “**Year 1**” to refer to the 2023 risk assessment and report, and “**Year 2**” to refer to the 2024 risk assessment and report.

### Methodology

Consistent with its approach in Years 1 and 2, TikTok has classified the systemic risks arising from the design or functioning of TikTok and its related systems into 12 “**Risk Modules**.” TikTok applied a risk assessment framework to discern the “inherent risk” (i.e. the risk level, based on likelihood and severity, before taking into account TikTok’s existing controls) for each Risk Module; and to identify the “residual risk” (i.e. the level of risk remaining after taking into account TikTok’s existing controls), based on an assessment of TikTok’s measures to mitigate the inherent risk. Building on learnings from Year 2, TikTok has refined its risk scoring methodology, as detailed in section 5 below.

### AIGC and Recommender Systems

TikTok has conducted a thorough Risk Assessment, which, consistent with Year 2, identified both AIGC and the impact of recommender systems as factors continuing to pose potential risks across the systemic risks. They are therefore factored in where relevant in the Risk Modules.

### Transparency reporting

TikTok publishes a number of transparency reports that are relevant to the matters covered in this Report. Where relevant, data in those reports has been analysed as part of TikTok’s Year 3 Risk Assessment and has been included in this Report. TikTok’s transparency reporting initiatives include:

1. DSA transparency reports, found [here](#);
2. EU Monthly Active Recipients reports, found [here](#);

3. Community Guidelines Enforcement reports, found [here](#);
4. Covert Influence Operations reports, found [here](#);
5. Government Removal Requests Reports, found [here](#);
6. User Information Requests Reports, found [here](#);
7. Terrorist Content Online Regulation reports, found [here](#);
8. TikTok Elections Integrity Hub, found [here](#);
9. Community Engagement Reports, found [here](#); and
10. Code of Practice on Disinformation reports, found [here](#).

### **Risk Assessment results**

Each of the Risk Modules below, which start from section 7 of this Report, sets out both the Inherent Risk and Residual Risk levels as assessed by TikTok. Where relevant, those sections include contextual detail that explains where there has been a change in risk level from the Year 2 Risk Assessment.

### **Geographic scope and purpose of this Report**

This Report has been prepared in compliance with Articles 34 and 35 of the DSA and is focused on systemic risks in the EU. As such, it focuses on TikTok's risk profile in the EU and its mitigation measures that apply to the use of the Platform by EU users. Therefore, the contents of this Report should not be relied upon as representative of TikTok's position outside of the EU. Furthermore, the Report has been prepared for the sole purpose of compliance with Articles 34, 35 and 42 of the DSA. It is not intended to be a definitive statement of TikTok's position on the matters covered as they may relate to other laws and regulations in the EU and should not be relied upon for any other purpose.

## **3. INTRODUCTION**

This Report has been prepared in compliance with Article 42(4)(a), (b) and (e) of the DSA. The Report summarises the results of TikTok's third Systemic Risk Assessment under Article 34 and the mitigation measures that have been put in place in line with Article 35 of the DSA. It has been prepared by TikTok Technology Limited ("**TikTok Ireland**") in relation to the operation in the European Union ("**Europe**" or "**EU**") of its online platform named TikTok (referred to as either "**TikTok**" or the "**Platform**" within this Report), which has been designated as a Very Large Online Platform ("**VLOP**") under the DSA.

Consistent with its efforts to continue to enhance its risk assessment processes, TikTok has enhanced the structure and contents of the Report since its Year 2 Report. This Report:

- Consolidates the risks identified in the Year 1 and Year 2 reports into a single overview of the risks TikTok has identified, relevant to each systemic risk.
- Provides a consolidated list of mitigation measures for each systemic risk.

- Includes a dedicated Illegal Content Risk Module relating to the Sale and Promotion of Illegal and Unsafe products, which has been included following the introduction of TikTok Shop in a number of EU member states since Year 2 (see section 12).
- Consolidates two Risk Modules from Year 2 that addressed risks to public health from medical misinformation and risks to public security from harmful misinformation content into a single harmful misinformation Risk Module (see section 17).

As reflected in TikTok's Year 1 and Year 2 reports, TikTok has examined the "potential for serious negative consequences to users' physical and mental well-being" (both of adults and youth), which is specified in Art. 34(1)(d), in the process of risk identification in all Risk Modules. This year, TikTok further builds on the assessment of Online Engagement (including mental health and wellbeing) of both adults and youth who use TikTok's service.

Consistent with the approach in Year 2, this Year 3 Report sets out the results of each Risk Module, grouped into topic-based risk categories (in the following order: (i) illegal content, (ii) youth safety and online engagement, (iii) misinformation and civic integrity and (iv) fundamental rights).

As explained in TikTok's Year 2 report, seeking advice and input across a wide array of stakeholders has been a cornerstone of the development of TikTok's safety efforts prior to the DSA and continues to be a central pillar of TikTok's risk management efforts. These engagements are integral to TikTok's process of developing and refining its safety policies and operations and continuously improving the Platform's capacity to reduce and mitigate the risk of harm. TikTok's stakeholder engagement strategy reflects a commitment to dynamically adapting its services based on continuous feedback and collaboration. Over the past year, TikTok has continued to proactively engage with stakeholders across various regions and contexts to ensure that it has a comprehensive understanding of the diverse challenges and opportunities that arise from its nature, scale and impact. In alignment with Articles 34, 35 and Recital 90 of the DSA, TikTok has and will continue to prioritise external stakeholder engagement in its risk assessment and mitigation strategies. In this Report, TikTok has listed relevant stakeholder engagement activities within each Risk Module.

### **TikTok Shop**

Starting in December 2024, i.e. after completion of its Year 2 Risk Assessment, TikTok has launched TikTok Shop ("**Shop**") in five EU markets: France, Germany, Ireland, Italy and Spain. Shop is an end-to-end e-commerce solution integrated within the Platform, available to "**Adult Users**" (i.e. users with a declared age of 18 years or older). It enables verified sellers and corporations to display and sell physical products to Adult Users, and permits Adult Users to buy products directly within the Platform.

As part of its Year 3 Risk Assessment, TikTok has considered the introduction of Shop as part of its assessment of the systemic risks. Relevant results are reported in this Report, including in the Risks of Illegal and Unsafe Products Risk Module and the Risks of Intellectual Property Infringing Content Risk Module.

## **European Commission Guidelines on the Protection of Minors (Article 28(4) of the DSA)**

As TikTok was finalising its Year 3 Risk Assessment, the European Commission published its Guidelines on measures to ensure a high level of privacy, safety and security for minors online, pursuant to Article 28(4) of the DSA (the “**Article 28 Guidelines**”). TikTok has made extensive efforts to develop its Platform in a way that provides a high level of privacy, safety and security, particularly for its “**Younger Users**” (those 13 to 17 years of age), and is committed to continually improving its practices to respond to emerging risks and new technologies. Where relevant, these measures are detailed in this Report.

The Guidelines state that the providers of VLOPs can conduct this risk review as part of its general systemic risk assessment pursuant to Article 34 of the DSA. Since the DSA came into force, TikTok has actively considered different types of risks to minors (for example, content and conduct risks) as part of its annual risk assessment efforts. This includes risks relating to age assurance, age-inappropriate content and online engagement and wellbeing and the appropriateness of the mitigation measures that TikTok’s implements. TikTok will continue to build on its approach to assessing risks to minors’ privacy, safety and security as part of its future risk assessment efforts.

TikTok is also actively assessing the Article 28 Guidelines and considering what additional recommended measures it may implement, beyond its current extensive suite of protections for Younger Users.

## **4. ABOUT TIKTOK**

This section of the Report sets out a high-level overview of key features of the Platform, how TikTok takes action against violative content and other considerations that have informed the Risk Assessment. Additional detail regarding features and key mitigations are set out in the Risk Modules that follow.

### **TikTok and its main features**

TikTok’s mission is to inspire creativity and bring joy. TikTok is available in many countries globally, including in the EU. The Platform had on average 159 million monthly active users in the EU between July 2024 and December 2024.

TikTok’s features include video, photo, livestream and comments. Users can share, skip, swipe, like, comment on or replay videos. Users can also Duet (side-by-side) with the video of another creator or Stitch another creator’s content into their video. Content in the For You Feed (see below) provides users with entertainment that is personal, and which connects people from all around the world through shared humour, interests and passions. Users in the EU have the option of making their feeds non-personalised. Users can send virtual gifts to creators whose content they like.



## **How the For You Feed works**

The For You Feed (“FYF”) is a TikTok feature that uses a personalised recommendation system to allow each community member the ability to discover a breadth of content, creators and topics. Content in the FYF is served based on: (i) User Interactions (such as content a user likes, shares or comments on); (ii) Content information (such as sounds, hashtags, number of views and the country in which the content was published); and (iii) User information (such as device settings, language preference and location).

TikTok maintains content [Eligibility Standards](#) for the FYF that prioritise safety and are informed by the diversity of TikTok’s community and cultural norms. TikTok makes certain content ineligible for the FYF, as it may not be appropriate for a broad audience, and may also make some of this content harder to search for. In order to mitigate the risk of harmful content appearing in the FYF, TikTok operates a range of processes and provides an additional layer of controls for Younger Users. The FYF is designed to ensure that users are given opportunities to discover new interests and are not presented with types of content that may be fine when viewed occasionally, but may be problematic if viewed repeatedly.

As explained further below, this year TikTok has implemented a new feature called “Manage Topics.” With Manage Topics, users can customise how often content related to over 10 topics is recommended in their FYF. TikTok has also introduced AI-powered “Smart Keyword Filters” to help users limit content they don’t want to see recommended to them at all.

## **Advertising on TikTok**

Businesses display ads on TikTok to reach audiences in a creative and meaningful way. TikTok publishes a Guide to Ads and Your Data and is transparent with its users about how it collects, uses and shares data for ads. In Europe, Younger Users do not see personalised advertising, based on their activities on or off the Platform, and will only see generic ads.

TikTok’s Advertising Policies determine the type of products and services that can be advertised on TikTok. Users will see different kinds of ads (marked as such) when they use TikTok, including Auction In-Feed Ads, TopView ads and Promote content. Users can interact with the ad in much the same way as content posted by other users. For example, users can share, skip, swipe, like or replay an ad. Users can also comment on an ad if the advertiser enables that feature for a particular ad. If users consider an ad to be violative of TikTok’s Advertising Policies or illegal, they can report it to TikTok.

## **How TikTok identifies and takes action on violative content**

TikTok operates proactive and systematic measures to identify, remove or restrict access to content and accounts that violate its Community Guidelines, Terms of Service, Advertising Policies or TikTok Shop Policies. TikTok’s content moderation is fundamental to its overarching risk management strategy as it underpins its ability to respond effectively to existing and emerging risks. As part of TikTok’s risk management strategy, it endeavours



to proactively identify and remove violative content before it is reported by users, non-users or third parties.

TikTok operates its content moderation processes using automated and manual (human) means in accordance with the following four key principles, which provide that TikTok will:

1. Remove violative content from the Platform that breaks its rules;
2. Age-restrict mature content (that does not violate its Community Guidelines, but which contains mature themes) so that it is only viewed by Adult Users;
3. Maintain FYF eligibility standards to help ensure that any content promoted by its recommendation system is appropriate for a broad audience; and
4. Empower its community with information, tools and resources.

TikTok operates automated and manual content moderation systems and processes, as well as a range of other safety features. These are developed, maintained and applied by a range of teams. All video, photo and text-based content uploaded to the Platform are subject to a real time, technology-based automated review. While a video is undergoing this review, it is visible only to the uploading user/creator. Users can report user-generated and ad content which they consider to be violative of the Community Guidelines or TikTok's Advertising Policies (as applicable).

Product listings on TikTok Shop are also reviewed through a combination of TikTok's automated and manual moderation measures. Where policy violations are established, product listings either do not go live or are removed in accordance with TikTok's enforcement framework.

TikTok's Trust and Safety teams action reports of violative content made by users, non-users and third parties. These teams perform a manual review against TikTok's Moderation Policy Framework, which provides moderators with necessary detail on how to apply TikTok's Community Guidelines. Additionally, users can report user generated and ad content that they consider to be illegal under European or member state law.

TikTok takes action in relation to content that it considers violative, which may include a review by its Trust and Safety team to determine whether the content should be removed or made ineligible for the FYF according to the Community Guidelines. Decisions can be appealed. TikTok also operates a "strikes" policy for accounts that repeatedly post violative content, which could result in various account level actions up to, and including, a permanent account ban. TikTok adopts a range of processes to review the accuracy of decisions when moderating potentially violative content, in order to ensure that content moderation measures are reasonable, proportionate and effective (and in particular, to ensure they do not disproportionately impact on users' rights to freedom of expression and information).

TikTok places considerable emphasis on proactive moderation of ads. Ads are reviewed against the Advertising Policies through a combination of automated and human moderation. Prior to being made available to users on the Platform, all ads are subject to technology-based automated review, designed to flag ads that may violate TikTok's Advertising Policies. This automated moderation of ads is supplemented by a number of

proactive and reactive detection measures, including risk monitoring, further review, user reports and regulatory or legal escalations.

### **TikTok's data-related practices**

TikTok offers a comprehensive suite of in-built privacy settings and controls that empower all users of the Platform to manage their experience on the Platform. For example, users can choose whether their account is public or private, decide who can view their content and control whether others can comment, send messages or interact through features like Duet and Stitch. Additional tools allow users to manage their discoverability, including whether their account is suggested to others or linked to their phone contacts or social profiles. Visibility settings such as activity status and profile view history can also be tailored to individual preferences. Together, these controls ensure that users have the flexibility to shape their TikTok experience in a way that aligns with their personal privacy expectations.

For Younger Users, there are supplemental privacy and safety controls, complementing those described above. For example, accounts for users with a declared age of under 16 are set to private by default, requiring approval for followers; and features like comments, Duets, Stitch, video downloads, message requests and account discovery are restricted or defaulted off. As users get closer to 18 years old, they gradually gain more flexible settings, but still with certain defaults (e.g. Duet/Stitch set to "friends" and downloads off unless enabled). For example, for users with a declared age of between 16 and 17 years old, when they sign up for TikTok, their account is pre-selected to private by default. Family Pairing allows parents/guardians to link and manage a teen's settings (including privacy settings), screentime limits, keyword filters, feed preferences, interactions and even time-away schedules, ensuring tailored oversight. Taken together, these options offer a layered, age-appropriate suite of privacy and safety settings, so users and families can choose the appropriate level of openness and protection.

In addition to user-facing controls, TikTok has adopted a sophisticated and integrated security infrastructure and information security management program, which includes a wide range of technical, contractual and organisational measures to ensure the integrity, confidentiality and security of all personal data and to prevent unauthorised access or disclosure. For example, TikTok implements a multi-tier system of technical access controls, such as system entry controls. TikTok has established and continues to enhance real-time cyber incident monitoring, response and investigative capabilities across hubs in Singapore, Dublin and Washington, D.C. This includes a robust incident management program and forensic tools to support timely recovery and data restoration, with infrastructure redundancy maintained through a comprehensive backup strategy. Operational and network security are strengthened by controls aligned with industry standards, including vulnerability scanning and continuous network monitoring. These capabilities ensure enterprise-level readiness, combining advanced threat detection with cybersecurity and incident response teams. TikTok's security team analyses data from internal traffic and external services to detect anomalies, such as unusual employee account activity or attempts to access user information. These efforts are supported by threat detection tools, enterprise-grade incident response capabilities and robust protections against malware, phishing and spoofing. TikTok additionally has contractual arrangements, with supplementary measures, in place

with group entities and its external service providers to ensure GDPR compliance. TikTok has implemented a number of internal organisational and policy measures to further control access to and use of personal data. Employees and contractors receive regular training to increase their awareness of TikTok's internal information security and data protection policies and best practices.

TikTok's Privacy Policy applicable to EU users of the Platform provides information and transparency regarding TikTok's personal data collection and processing activities, including data transfers and data retention. It also explains to users their rights regarding their personal data, including information on their rights of access, objection, restriction, deletion, rectification and portability of their personal data. This Privacy Policy was updated in November 2023 to address DSA requirements related to sharing data with researchers, and further updates were published in 2024.

### **TikTok's Fact-Checking Programme**

To deliver on its commitment to tackle harmful misinformation, including election misinformation, medical misinformation and other harmful misinformation (as defined in sections 16 and 17), TikTok operates a global Fact-Checking Programme. TikTok works with more than 20 IFCN-accredited fact-checking organizations who assess the accuracy of content on the Platform in over 60 languages and 130 markets across the world. The core objective and mission of the Fact-Checking Programme is to leverage the expertise of external fact-checking organisations to verify potential misinformation claims, and support TikTok's Trust & Safety teams in determining the appropriate enforcement action. TikTok considers that partnering with third-party fact-checking organisations is an important component of its strategy to effectively tackle harmful misinformation. More detail regarding the Fact-Checking Programme is available [here](#).

## **5. RISK GOVERNANCE**

### **A. INTERNAL RISK GOVERNANCE**

TikTok recognises that effective risk management is critical to providing a safe and welcoming online platform. TikTok operates a safety-first approach that integrates human rights standards and fosters cross-functional collaboration.

TikTok strives to employ effective, reasonable and proportionate risk mitigation measures tailored to identified systemic risks, whilst avoiding unnecessary restrictions to Platform use or fundamental rights.

TikTok balances user safety with other fundamental rights (as relevant) when assessing risks at the intersection of safety, privacy, protection of youth and freedom of expression and information. Acknowledging the complexity, TikTok is committed to achieving a fair balance for its users.

TikTok's policy-driven approach to risk management is reflected in its Community Guidelines, which establish a common code of conduct for all users and set out details of TikTok's approach to enforcement ("**Community Guidelines**"). Continuous improvement to risk management is fostered through internal training and awareness programmes, which drive well-informed decision-making and sound governance. TikTok also prioritises raising user awareness of risks on the Platform, through in-app and online resources and interventions. Proactive risk management and detection strategies are employed, even though general monitoring for illegal content is not required under the DSA.

## B. SYSTEMIC RISK ASSESSMENT METHODOLOGY

The methodology comprises six key phases, all of which have been reviewed and repeated for Year 3.

Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Systemic risk scoping and definition	Risk identification	Inherent risk analysis and assessment	Identification and assessment of mitigations	Calculation of residual risk and analysis of future improvements	Reporting, governance and review

### Phase 1: Systemic risk scoping and definition

TikTok has developed and maintained a systemic risk taxonomy that outlines the scope of the systemic risks that could stem from the design, functioning, use or mis-use of its Platform. In Year 1, TikTok developed a process for scoping and defining risks based on an analysis of three key elements:

- The four primary systemic risk categories, as described in the DSA;
- A legal and textual interpretation of key DSA terms, other relevant EU law and jurisprudence and relevant international legal instruments; and
- A functional interpretation made by TikTok's internal teams.

In Years 1 and 2, this process informed the development of a systemic risk taxonomy that included twelve distinct Risk Modules, grouped under four primary categories. Upon careful consideration of its prior approach, TikTok decided that it was appropriate to adhere to a similar taxonomy for Year 3, with some amendments to reflect TikTok's ongoing efforts to refine its approach to risk assessment. For Year 3, TikTok continues to group the Risk Modules under four primary categories: "**Illegal Content**," "**Youth Safety and Online Engagement**," "**Civic Integrity and Misinformation**" and "**Fundamental Rights**." While TikTok also continues to include twelve distinct Risk Modules, it has made the following changes for Year 3. We: (i) introduced a Risk Module on the promotion and sale of illegal and unsafe products; (ii) consolidated the modules on risks to public health from medical misinformation, and risks to public security from harmful misinformation, into a single module on risks arising from harmful misinformation; and (iii) expanded the risk module on online engagement (which addresses risks to physical and mental health) and now

separately report on risks related to online engagement and risks related to age-inappropriate content (which were summarised together in Year 2).

## **Phase 2: Risk identification**

Phase 2 involves identifying the specific ways in which each risk area may manifest on TikTok. For Year 3, TikTok reviewed the risks it identified in Years 1 and 2 and identified any new or emerging risks, or new ways the risks could manifest on the Platform.

This risk identification is informed by three key processes:

- Consideration of internal sources (e.g. data insights from content moderation practices);
- Consideration of external sources (e.g. stakeholder engagement, law enforcement reports and credible independent research); and
- An assessment of potential emerging risks (e.g. proactive risk assessment, internal research, stakeholder engagement).

In identifying how the systemic risk may manifest on the Platform, TikTok assesses how it is influenced by the factors laid out in Article 34(2) of the DSA and relevant recitals. Article 34(2) instructs platforms to assess *“the design of recommender systems and any other relevant algorithmic system; content moderation systems; the applicable terms and conditions and their enforcement; systems for selecting and presenting advertisements; [and] data related practices.”* Where the risks are localised or there are linguistic differences, TikTok takes these into consideration when assessing risk and corresponding mitigations.

In the Year 2 Risk Assessment, TikTok has specifically addressed two areas that are of a cross-cutting nature across several risk areas. These are:

- The impact of an increased prevalence of AIGC; and
- The impact of recommender systems on the content consumed by users.

As part of its Year 3 Risk Assessment, TikTok has continued to focus on these priority areas. Identification of these priority areas has been informed by internal risk detection analysis and an analysis of priorities expressed by stakeholders, including the European Commission.

## **Phase 3: Inherent risk analysis and assessment**

Within the assessment of inherent risk, TikTok has considered severity and likelihood for each Risk Module, prior to considering the impact of any implemented mitigation measures.

To assess inherent risk, severity and probability are assessed before being combined to determine the inherent risk level. To quantify each factor:

- **Severity** is assessed through a weighted measurement of the scope (the affected population) and seriousness (the degree and type of harm) for each Risk Module; and

- **Likelihood** is assessed based on whether (and how much) the risk has already been observed and whether it is expected to occur in the future. Given that controls are in place, it is difficult to measure the probability of events occurring without any controls in place.

The severity and likelihood outputs are then combined to produce an inherent risk score for each Risk Module, on the following scale: **Very Low, Low, Medium, Medium-high** and **High**.

#### **Phase 4: Identification and assessment of mitigations**

TikTok has identified the mitigation measures that it has implemented relevant to each Risk Module, with reference to the types of mitigations set out in Art. 35(1)(a)-(k) of the DSA, noting new mitigations that have been implemented in Year 3. Mitigation measures (i.e. controls) are evaluated to determine the extent to which they mitigate the inherent risk. This evaluation is then used to calculate residual risk.

#### **Phase 5: Assessment of residual risk and analysis of future improvements**

Residual risk represents the assessed level of risk that remains for each risk area after the impact of relevant controls/mitigations on reducing inherent risk is considered. It is calculated by evaluating severity and likelihood in light of the effectiveness of the mitigations for each Risk Module to produce a residual risk for each Risk Module on the following scale: **Very Low, Low, Medium, Medium-High** and **High**. This analysis informs TikTok's ongoing assessment of risks and risk mitigation efforts.

#### **Phase 6: Reporting, governance and review**

Each Risk Module is assessed and approved by subject matter experts and leaders to ensure robust governance and strong accountability, as follows:

1. The Risk Module is first submitted to a specialist Risk Assessment Review Group for analysis and assessment. Risk Assessment Review Groups are composed of senior internal stakeholders with expertise in the risk in question. The Risk Assessment Review Groups' process, including the Compliance function's input, is designed to ensure that all risks referred to in Article 34 are identified and properly reported on, and that relevant mitigations have been identified and properly assessed;
2. The results are then reported to the Online Safety Oversight Committee, TikTok's steering group of cross-functional leaders, which has been in place throughout its DSA programme; and
3. Finally, the results are submitted to the Board of Directors of TikTok Technology Ireland for review and approval.

### **C. RISK ENVIRONMENT**

TikTok is one of a number of online platforms whose users face similar, but not identical risks. These risks often transcend the boundaries of individual platforms and involve

complex interactions between online experiences and real-world events, facilitating personal expression and the receiving and imparting of information and ideas. This is a central value of the EU, a fundamental right and a core pillar of European democracy and social life. Each platform is expected to operate on its own analysis of the appropriate balance between safety and fundamental rights for its users. TikTok takes an approach that prioritises caution in areas where risks are more likely to impact real-world harms, or where they may be more likely to impact Younger Users.

TikTok strives to only host content within the bounds of relevant laws and the rules it imposes on itself and its users. TikTok therefore aims, through its use of automated and human moderation measures, to proactively identify and remove violative content before it is seen by any user. However, operating within such a complex environment, it is not possible to eliminate all risks, and moderation may itself have an impact on fundamental rights, such as freedom of expression. TikTok therefore aims for a balanced and proportionate approach to reducing risks continuously, in line with Recital 86 DSA.

While TikTok is not specifically aimed at minors or predominantly used by them, TikTok makes it a priority to address how risks may impact these users. Please see sections 13, 14 and 15 of this Report.

TikTok has a comprehensive crisis management plan to address unforeseen challenges. TikTok maintains a dedicated risk containment team to address urgent issues and to contain and minimise harm. Additionally, TikTok is committed to learning from past incidents, adapting its strategies and fortifying its Platform against future risks.

#### **D. WHAT TIKTOK DOES TO COMBAT INTENTIONAL MANIPULATION OF THE SERVICE**

TikTok aims to foster an environment where genuine interactions and content thrive. However, bad actors may attempt to manipulate the service through deceptive behaviours like account impersonation, spam activity and fake engagement. This can range from individual behaviour, such as creating a fake account, to sophisticated and coordinated influence operations.

To combat these issues, TikTok prohibits and works to disrupt intentional manipulation and inauthentic use of the Platform. TikTok's verified account badge helps protect users by providing a reliable indication of notable accounts.

Expert teams at TikTok focus on detecting, investigating and disrupting covert influence operations and other inauthentic use of the Platform. Suspicious accounts and behaviours are removed, and inauthentic engagement signals are ignored in estimating account or video popularity. TikTok reports on the number of accounts actioned under its policies against covert influence operations and inauthentic use of the Platform in its public transparency centre.

Covert influence operations ("**CIOs**") are coordinated, inauthentic behaviours where networks of accounts work together to attempt to manipulate or corrupt public debate while



also misleading TikTok systems or users about identity, origin, operating location, popularity or overall purpose. TikTok recognises that CIOs will continue to evolve and that networks may attempt to re-establish a presence on the Platform. As new deceptive behaviours emerge, TikTok is committed to evolving its response, strengthening enforcement capabilities and publishing its findings for scrutiny and more effective cross-Platform risk mitigation.

As detailed in section 16 below and in TikTok's transparency reports, one area where TikTok has identified issues regarding intentional manipulation of its service relates to elections and civic integrity risks. In TikTok's [Code of Practice on Disinformation \(COPD\) transparency reports](#), TikTok reports on its efforts to identify and disrupt CIO networks that seek to sway public opinion and mislead users of the Platform. In TikTok's [Elections Integrity Hub](#), it also reports on its efforts to identify and block the efforts of CIOs to disrupt elections that are taking place in the EU.

## **E. STAKEHOLDER ENGAGEMENTS**

TikTok has actively engaged with a broad spectrum of stakeholders to strengthen its risk assessment and mitigation strategies under the DSA. This proactive approach is central to TikTok's commitment to user safety and the protection of fundamental rights on its Platform. Over the past year, TikTok has consulted with a diverse range of experts, including members of its European Safety Advisory Council ("**ESAC**"), its Youth Council, policymakers, civil society organisations and content creators. Through close collaboration with stakeholders across member states, TikTok can better understand and respond to local contexts, ensuring that its safety initiatives are not only globally robust but also take into account specific regional considerations. These partnerships enable TikTok to combine its existing Trust and Safety teams' expertise with further local knowledge, creating more effective and targeted risk mitigations, including campaigns that support a safer online experience for different communities.

Developing Trust and Safety policies and processes to support the many communities that are celebrated on the Platform is not something that TikTok does alone, as TikTok purposely seeks to engage with as many different perspectives as possible at all steps of the development process. This dialogue is important for a number of reasons. While TikTok has an experienced and knowledgeable Trust & Safety team, external perspectives are important to help it more deeply understand and proactively address topics of concern. TikTok integrates independent third-party perspectives to support fact-based decision making aligned with international best practices. These partnerships come from a range of entities, such as NGOs, industry experts, academia, researchers, creators and the voice of lived experience. TikTok manages these engagements in a variety of ways, including via ESAC and its Youth Council, one-off consultations, long-term partnerships, roundtable discussions, focus groups, vulnerable groups and community engagements. Diversifying how TikTok gathers these external inputs allows it to support its teams in a range of different ways, appropriate to policy areas of focus.

Third-party engagement is also important in building trust across stakeholders and TikTok's user community. With expert input, TikTok can transparently demonstrate its efforts to

understand risks and develop mitigation strategies that are balanced and factual. This process provides reassurance in the content decisions that TikTok makes and provides additional safeguards that prevent harm before it happens, improving overall user experience. TikTok also provides post-implementation feedback to its external partners where possible so that those partners can understand and observe the impact of their engagement.

In this Year 3 report, TikTok has listed relevant stakeholder engagement activities that have taken place over the last 12 months within each Risk Module.

## **F. HOW TIKTOK COLLABORATES WITH DSA TRUSTED FLAGGERS**

Trusted Flaggers designated under Article 22 of the DSA play a role in risk identification. TikTok is engaged with existing Trusted Flaggers and continues to be ready to engage with additional partners as they are designated.

As noted in TikTok's fourth DSA Transparency Report, TikTok introduced an additional reporting channel for TikTok's EU community to "Report Illegal Content," which enables Trusted Flaggers to report content they consider to be illegal. TikTok publishes metrics regarding reports received from Trusted Flaggers in its DSA Transparency Reports.

The work of Trusted Flaggers supplements the efforts of TikTok's Community Partner Channel. This includes hundreds of safety and civil society organizations with a broad range of expertise across content safety and youth safety issues, which can report potentially violative content to us directly for review. These partners supplement TikTok's content moderation practices by identifying and sharing locally nuanced risks.

## **6. AIGC AND RECOMMENDER SYSTEMS**

### **Introduction**

Consistent with its approach in Year 2, TikTok has identified AIGC and recommender systems as specific risk areas that impact the systemic risks. TikTok has adapted and implemented a series of mitigation strategies to address these risks comprehensively across the Platform. These measures are designed to safeguard the user experience and ensure the Platform's resilience against emerging threats in these areas.

### **A. THE IMPACT OF INCREASED USE OF AIGC**

#### **Definition and cross-module risks**

TikTok continues to assess the potential risks that could arise through the sharing and dissemination of deceptive - but highly realistic - images, videos and audio, potentially making it more difficult to distinguish between fact and fiction on the Platform. AIGC creates new and evolving opportunities for bad actors to create and disseminate illegal content

(including image-based abuse). TikTok assesses the growing sophistication and capabilities of generative AI models, which could heighten the risks on the Platform.

As AIGC tools can be used to create content that is indistinguishable from other user-generated content (“**UGC**”), the use of such tools creates potential risks to the integrity and reliability of content hosted on TikTok. “**Deep fakes**” - extremely realistic photos, videos or voices of real people - are a particular area of concern, given their risk of harm to depicted individuals. Additionally, AIGC technologies can be used to facilitate the creation of illegal content, such as synthetic Child Sexual Abuse Material (“**CSAM**”), and image-based sexual abuse.

The rise of AIGC has also created new challenges for the industry in the context of elections.

AIGC tools are now widely available to create manipulated content at low cost and high speed, which magnifies the chance of harmful content being created and disseminated.

### **Cross-risk mitigations**

TikTok continues to implement measures across UGC, Shop and advertising features to mitigate potential risks associated with AIGC.

Regarding UGC, TikTok prohibits AIGC that violates its Community Guidelines to the same extent as any other UGC. For example, AIGC that portrays CSAM is prohibited to the same extent as non-edited/non-AI-generated content portraying CSAM.

TikTok’s Community Guidelines also contain an Edited Media and AIGC Policy, which:

- Requires users to disclose AIGC or edited media that shows realistic-appearing scenes or people;
- Prohibits AIGC that shows:
  - the likeness of private figures without their permission; or
  - certain depictions of minors;
- Prohibits misleading AIGC such as:
  - content made to seem as if it comes from a news organisation;
  - a crisis event, such as a conflict or natural disaster; or
  - content showing a public figure who is: (i) being degraded or harassed or engaging in criminal behaviour; (ii) taking a position on a political issue; or (iii) being politically endorsed or condemned.

TikTok’s Synthetic and Manipulated Media Advertising Policy addresses the risk of AIGC in ads on the Platform. In particular, it prohibits visual or audio material that has been edited, spliced or combined in a way that may mislead a viewer about real-world events. This policy also sets out the following requirements:

- Manipulated or synthetic media must include clear disclosures or labels in ad content.

- If ads with synthetic media contain the visual or audio likeness of a real person (whether a public or private figure), advertisers must ensure they have consent from the real person to be used in an ad in this way.
- Advertisers must ensure that any ad with synthetic media complies with copyright laws, data protection laws and privacy laws.

TikTok moderates content against these policies and has processes in place to remove content that is violative.

TikTok will further review and iterate on these policies as AIGC trends, risks and enforcement capabilities evolve.

For the purposes of detecting and enforcing against the use of AIGC to create violative content in either a UGC or ads context, TikTok has implemented tools for users and advertisers to voluntarily label their content as “AI-generated.” TikTok was the first platform to launch an AIGC labelling tool, which tens of millions of creators worldwide have used since September 2023. From Q3 2024 to Q1 2025, 4,690,837 unique users in the EU have labelled their content with the “Creator labelled as AI-generated” label, accounting for 10,680,682 unique videos. Additionally, any AIGC that is made with TikTok’s AI effects is automatically labelled. If users believe that content they encounter is unlabelled AIGC that violates TikTok’s Community Guidelines requirements for disclosure regarding AIGC, they can report it to TikTok, who may then remove the content. TikTok reports content removed under its Edited Media and AIGC Policy in its Community Guidelines Enforcement Reports.

TikTok also works with industry groups and civil society to detect undisclosed AIGC. TikTok participates in the Coalition for Content Provenance and Authenticity (“**C2PA**”), an open technical standard and content provenance solution that can provide information in a piece of content’s metadata about its origins and whether AIGC models were used to create or edit it. This technology helps TikTok to detect AIGC at scale when content with C2PA metadata is uploaded to the Platform. TikTok was the first video-sharing platform to launch the ability to read Content Credentials from C2PA. Through its use of C2PA, TikTok automatically labels AIGC that has been made on other platforms. From Q3 2024 to Q1 2025, 4,781,019 videos on TikTok were auto-labelled with the AIGC tag of “AI-generated” in the EU.

TikTok has developed models to assist in the detection and moderation of AIGC uploaded to the Platform, including:

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

Finally, TikTok’s capacity to effectively detect AIGC is reinforced by TikTok’s content moderation systems for content violative of TikTok’s Community Guidelines, external partner reports, internal teams who identify instances of violative use of AIGC, manual reviews of trending or popular content and targeted proactive sweeps. TikTok’s approach

has been informed by research conducted by the Trust & Safety Product Policy and Policy Development teams on academic, industry and advocacy literature. TikTok has also consulted a range of external bodies and subject matter experts, such as the Partnership on AI.<sup>1</sup>

## **B. RECOMMENDER SYSTEMS**

### **Definition and cross-module risks**

As explained in more detail in section 4 above, TikTok makes available the FYF, which is the primary means through which users consume video content on the Platform. The FYF uses a personalised recommendation system to help each user discover content, creators and topics, while ensuring that content is interesting and relevant to each user. In determining what is recommended, the FYF's recommender system takes account of user interactions, content information and user information.

Recommendation affects the inherent risk that content that violates TikTok's Community Guidelines or other policies could trend on TikTok. As TikTok's recommender systems can broaden the reach of the content recommended to users, they can potentially increase the risk of rapid and wide dissemination of such content, which could include harmful content.

In addition, given that user's interests and historical engagement are factored into recommendations or recommender system design, they might inadvertently present users with a narrow range of content. Certain types of content, though not violative of TikTok's Community Guidelines, may cause harm to some users if viewed repeatedly ("**Concentrated Content**") by inadvertently reinforcing a negative personal experience. There are also risks to freedom of expression associated with acting against Concentrated Content (e.g. shocking and graphic content).

Finally, recommender systems, including the FYF, may have an impact on a user's extended use of a service. Section 15 addresses risks posed by online engagement for both adult and Younger Users.

### **Cross-risk mitigations**

TikTok takes a range of measures to mitigate the risk that violative content will be amplified or trend as a result of the operation of its recommender systems.

First, TikTok proactively enforces its Community Guidelines to mitigate the risk of illegal and harmful content appearing on users' FYFs. Second, TikTok maintains content eligibility standards for FYF that prioritise safety and are informed by the diversity of community and cultural norms. For example, unverified information awaiting fact-checker review is ineligible for FYF. Third, TikTok integrates age-appropriate design into the FYF, as content created by anyone under 16 years old is ineligible for recommendation and its content classification system rates UGC based on maturity and prevents inappropriate themes from being

---

<sup>1</sup> A non-profit organisation which brings together companies and other stakeholders to establish common codes of practice across the industry.

recommended to Younger Users. Fourth, to mitigate additional risks associated with trending content, TikTok manually reviews all content that reaches a certain level of popularity (by monitoring the volume of video views). TikTok also uses risk management protocols to detect and contain trends that may feature violative content.

To mitigate risks arising from Concentrated Content, TikTok uses dispersion techniques in its FYF. This involves using machine learning models to avoid recommending a series of similar videos on themes that do not violate TikTok's Community Guidelines but that may nevertheless be potentially problematic if viewed repeatedly. In addition, TikTok monitors the risks of Concentrated Content in the FYF and adjusts risk tolerance based on the latest information.

TikTok also offers tools and settings to empower users to influence what they see on their FYF. This includes tools to understand why videos have been recommended; to filter out certain keywords or hashtags to stop seeing certain content; and to "refresh" their FYF, in order to update the content recommended on their feed.

As of March 2025, 2,619,937 users had actively filtered hashtags/keywords in the EU.

Since its Year 2 Risk Assessment, TikTok has introduced additional measures to provide users with more control over their FYF, including the following:

- TikTok has launched its "[Manage Topics](#)" tool, which allows users to customize the content they see in their FYF based on their preferences for a specific number of topics. Through their in-app settings or directly in the FYF, users can move a slider to adjust how much they want to see of each topic.
- TikTok enhanced its [keyword filtering measures](#) through its launch of "Smart Keyword Filtering" in June 2025. Now, users can not only filter keywords and hashtags from the FYF, but they can also filter out similar keywords and variations by turning on the smart filter.

With respect to risks relating to recommender systems and extended online engagement, TikTok encourages the development of digital wellbeing by balancing user agency and encouraging mindful use of the Platform, while also actively taking into account if any targeted measures need to apply to Younger Users given their state of maturity and development.

TikTok implements a number of mitigation measures to address Online Engagement Risks (defined in section 15), including: (i) screentime management tools, (ii) a daily screentime limit for Younger Users that is set to 60 minutes by default, (iii) Family Pairing controls, (iv) sleep reminders, (v) restrictions on push notifications, (vi) digital well-being prompts, (vii) screentime break prompts, (viii) take-a-break videos and (ix) the "Wind Down" feature. TikTok continuously iterates and expands its suite of wellbeing features, taking account of the latest research in the area.

## 7. ILLEGAL CONTENT: RISKS OF ILLEGAL HATE SPEECH CONTENT

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the term “**Hate Speech**” in a manner consistent with EU and EU member state laws, including, but not limited to, EU Framework Decision 2008/913/JHA, Article 1, in respect of offences concerning racism and xenophobia (i.e. targeted speech in any form against a group of persons or a member of such a group defined by reference to protected characteristics, such as, sex, gender, gender identity, race, colour, ethnic origin, language, religion or belief, membership of a national minority, birth, disability or sexual orientation).

TikTok also acknowledges that the protection of other fundamental rights and freedoms, such as the right to non-discrimination contained in Article 21 of the EU Charter of Fundamental Rights (the “**Charter**”) may be undermined by the dissemination of Hate Speech on the Platform.

#### Summary of risks

The risks associated with Hate Speech may arise from users attempting to share or disseminate the following content on or through the Platform, including through video or photo, livestream, comments (in the form of text or symbols), profile information or TikTok Shop features:

- Content claiming individuals or groups with protected attributes are physically, mentally or morally inferior or referring to them as criminals, animals, inanimate objects or other non-human entities;
- Content promoting or justifying violence, exclusion, segregation or discrimination against a protected group;
- Content that includes the use of slurs against others (and not the user themselves);
- Content that denies or minimises the scale of well-documented historical events that harmed protected groups (e.g. denying the existence of the Holocaust or the genocide against the Tutsi in Rwanda);
- Content that targets transgender or non-binary individuals through misgendering or deadnaming;
- Content promoting hateful ideologies, such as conspiratorial statements that target a protected group or claiming supremacy over a protected group;
- Content that depicts harm inflicted upon an individual or a group on the basis of a protected attribute; or
- Content that indirectly demeans protected groups.

There are moderation challenges associated with this type of content as hate speech can be highly localised in terms of language and region and use rapidly evolving alternative vocabulary. The increased availability of AIGC tools can also heighten the risk of content being used to produce fake videos, audio or memes that spread Hate Speech under the guise of humour or using coded language or symbols, making it more difficult to detect.



## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in the inherent risk profile for Hate Speech and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk Score is consistent with TikTok's score in Year 2. TikTok continues to attribute this risk level to intensified global conflicts and challenges associated with potential uses of AIGC, which have amplified the potential for Hate Speech content to appear on the Platform and adversely affect the individuals or groups it targets.

## 3. Mitigation Measures

The mitigations below are organised with reference to sub-sections of Article 35(1) of the DSA, which details measures that VLOPs should consider when determining whether they have implemented reasonable, proportionate and effective risk mitigation measures to address the Systemic Risk identified pursuant to Article 34.

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of the dissemination of Hate Speech on or through the Platform, including (i) blocking certain search results associated with Hate Speech from appearing on the Platform; (ii) measures to mitigate the risk of a livestream being capable of containing Hate Speech; (iii) measures to address the risk that comments or hyperlinks could be used to disseminate Hate Speech (including the “*consider before you comment*” feature, to give users the opportunity to reconsider potentially hateful content before posting); (iv) settings allowing users to control who can Duet and Stitch with their videos each time they post content, which limits the risk that these features can be used to create Hate Speech content; (v) in-app prompts and search guides that re-direct users to authoritative information when they search for certain content associated with Hate Speech and discourage the creation or sharing of hateful content; and (vi) creator tools, including the keyword filter tool, that allows users to filter specific keywords, the unwanted comments filter, which automatically detects and hides spam or generally disliked comments (turned on by default), and “creator care mode” which filters and hides comments a creator is likely to find offensive, based on their prior interactions.

**Article 35(1)(b):** TikTok very clearly prohibits Hate Speech, hateful behaviour or the promotion of hateful ideologies on the Platform through a combination of TikTok's [Terms of Service](#), its [Community Guidelines](#) and [Advertising Policies](#). For example, under its “Hate Speech and Hateful Behavior” policy, TikTok prohibits content promoting any hateful ideology, promoting violence, segregation, discrimination and other harms on the basis of a protected attribute, using a hateful slur associated with a protected attribute or

dehumanizing someone on the basis of their protected attributes. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. In addition, TikTok collaborates with external safety organisations, civil society groups, researchers and fact-checkers to stay on top of evolving trends and hate-speech patterns, to understand local dynamics and to continually refine its detection methods.

To ensure consistent enforcement during high-risk/high-profile periods, dedicated and thematic policy guidance is issued to moderators in connection with crisis events (e.g. elections, armed conflicts, major sporting events).

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. In addition, TikTok’s policy designates content that indirectly demeans protected groups as ineligible for recommendation in the FYF.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not contain hateful behaviour, hate speech, inappropriate glorification or promotion of hateful ideologies or otherwise violate TikTok’s Community Guidelines or Advertising Policies. To ensure robust enforcement of its ad policies, TikTok conducts regular monitoring of potentially violative ads in high-risk areas and uses information gathered to improve its automated moderation measures and mitigate concentrated areas of risk.

**Article 35(1)(f):** TikTok has specialist teams who deal with illegal content, including Hate Speech, and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(h):** TikTok signed the EU Code of Conduct on Hate Speech in September 2020 and participates in regular monitoring exercises by participating NGOs from the European region. The evaluation seeks to understand how quickly and reliably platforms respond to reports on hate speech content. In January 2025, the EU Code of Conduct on Countering Illegal Hate Speech Online + was implemented as a Code of Conduct under the DSA.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Transparency Centre, which has dedicated content explaining [TikTok’s approach to content moderation](#) and [Combating hate](#); (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about TikTok’s content moderation practices and how to report violative content; (iii) TikTok’s Safety Centre, which contains information on [TikTok’s approach to safety](#), as well as a dedicated page on [Countering hate speech & Behavior](#) (updated in July 2025); and (iv) TikTok’s other in-app educational and awareness measures, verified

badges, state-controlled media labels, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** TikTok dedicates significant resources to removing illegal content and seeks to remove it before any user, of any age, is exposed to it. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a summary of TikTok’s controls related to AIGC. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.

TikTok has further enhanced its mitigation measures related to Hate Speech since its Year 2 Report, including:

- TikTok revised its hate policy titles globally, in order to refine their scope and better address the complexity of Hate Speech content, harms and prevalence. TikTok also enhanced the resources available to its moderation teams.
- TikTok introduced new policies and moderation measures to address the challenges posed by complex hate behaviour that often spans multiple features and content types.
- TikTok has also developed a machine learning model [REDACTED].

This model supports earlier detection and strengthens TikTok’s ability to prioritize enforcement.

- TikTok has introduced new user empowerment tools to address content that may not constitute hate speech or otherwise violate the Community Guidelines but that nevertheless may be offensive to some users. These tools allow content creators to filter, review and approve comments before they are publicly visible against their content.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with the European Commission and other signatories regarding the conversion of the 2016 Code on Hate Speech into a DSA Code of Conduct.
- In June 2025, approximately 50 government and civil society organisations, together with a number of TikTok content creators, attended the Anti-Hate Summit at TikTok’s Dublin office in connection with the EU Hate Speech Code.
- Engagement with NGOs regarding Hate Speech, Misinformation and Incitement to Violence. For example, TikTok’s engagements with LGBT+ organisations and agencies and NGOs focused on combating online hate from across Europe support TikTok with policy development, risk detection and reporting and removing Hate Speech from the Platform.

#### **4. Residual Risk**

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of Hate

Speech to be **Medium**. This Residual Risk Score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 829,037 videos under its "Safety and Civility - Hate Speech & Hateful Behaviour" policy; 703,714 of those videos were detected and removed proactively (i.e. before any user report), 371,195 were removed before being viewed by any user, and 500,803 were removed within 24 hours of upload, demonstrating TikTok's swift action to combat Hate Speech on the Platform.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, 125,323 videos were removed, following a user report, for violation of TikTok's "Hate Speech & Hateful Behaviour" policy.

A number of TikTok's other policies capture content containing Hate Speech, including its "Safety and Civility - Bullying and Harassment" policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the risk of Hate Speech on the Platform.

As noted in TikTok's Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok received 22,429 reports from users of allegedly illegal content relating to Illegal Hate Speech. During the same period, it received 5 reports from Trusted Flaggers relating to Illegal Hate Speech. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 111 of which related to Illegal Hate Speech.

External events can have a bearing on the type and volume of Hate Speech TikTok observes on the Platform. As such, TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve its automated and manual moderation processes. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 8. ILLEGAL CONTENT: RISKS OF TERRORIST CONTENT

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the term “**Terrorist Content**” in a manner consistent with EU and EU member state laws. Its scope has been defined having particular regard to Regulation (EU) 2021/784 on addressing the dissemination of Terrorist Content Online (“**TCO Regulation**”), and Directive 2017/541 (EU), Articles 3 to 12. Those laws outline the following illegal activities: (i) Terrorist offences; (ii) Offences relating to a terrorist group; (iii) Public provocation to commit a terrorist offence; (iv) Recruitment for terrorism; (v) Providing training for terrorism; (vi) Receiving training for terrorism; (vii) Travelling for the purpose of terrorism; (viii) organising or otherwise facilitating travelling for the purpose of terrorism; (ix) Terrorist financing; and (x) Other offences related to terrorist activities.

TikTok also has regard to Article 2 of the Charter, which enshrines the right to life, and Article 6, which protects the right to liberty and security of the person, and acknowledges that such rights may be undermined by the dissemination of Terrorist Content on the Platform.

TikTok also recognises that efforts to moderate Terrorist Content must be accurate, balanced and reasonable to ensure that such efforts do not disproportionately impact on other fundamental rights under the Charter, in particular, the rights to freedom of expression and information, data protection and non-discrimination, and freedom of thought, conscience and religion, as well as TikTok’s freedom to conduct a business.

#### Summary of risks

The risks associated with Terrorist Content may arise from users attempting to share or disseminate the following content on or through the Platform, including through video or photo, livestream, comments, profile information or TikTok Shop features:

- Content that praises, promotes, glorifies or supports violent acts or extremist organisations or individuals, or content seeking to raise funds or obtain assistance for such entities;
- Content that encourages participation in, or intends to recruit individuals to, violent extremist organisations; and/or
- Content with names, symbols, logos, flags, slogans, uniforms, gestures, salutes, illustrations, portraits, songs, music, lyrics or other objects meant to represent violent extremist organisations or individuals.

Terrorist Content, including illegal, violent and offensive content, which could cause distress to users, have negative effects on mental health and potentially lead to desensitisation to violence. Some users that are susceptible to mental illness may also be vulnerable to radicalisation pathways.

There are also risks that users will make inauthentic or manipulative use of the Platform to:

- Impersonate others in a deceptive manner: Impersonation may involve terrorist organisations creating false accounts that appear trustworthy, which could be used for recruitment or the promotion of radical or extremist beliefs; or
- Communicate in such a way to evade content moderation: There is a risk that TikTok's content moderation systems may not be familiar with vocabulary related to Terrorist Content that is rapidly evolving, and, for example, uses otherwise benign symbols and vocabulary.

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in the inherent risk profile for Terrorist Content and has considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

TikTok has revised its Inherent Risk Score from Year 2 (Medium), based on a continued increase in global conflicts and crisis events, which result in a heightened inherent risk of terror-related content appearing on the Platform in various forms.

## 3. Mitigation Measures

TikTok clearly prohibits Terrorist Content, as well as certain related content, on the Platform. TikTok undertakes the following mitigation measures related to Terrorist Content risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of the dissemination of Terrorist Content on or through the Platform, including: (i) blocking certain search results associated with Terrorist Content from appearing on the Platform; (ii) measures to mitigate the risk of a livestream being capable of containing Terrorist Content; (iii) measures to prevent the risk of comments or hyperlinks being used to disseminate Terrorist Content; and (iv) search guides to redirect users to credible information to address the proliferation of content from designated violent extremist groups and during crisis events.

**Article 35(1)(b):** TikTok clearly prohibits Terrorist Content, as well as certain related content, on the Platform through a combination of TikTok's [Terms of Service](#) and the [Community Guidelines](#). For example, under its "Violent and Hateful Organizations and Individuals" policy, TikTok prohibits (i) accounts operated by organizations or individuals that promote violence or hateful ideologies, (ii) content or behaviour which provides material support to violent political organizations or promotes violence caused by them and (iii) content or behaviour which promotes (including through praise, celebration or sharing manifestos) or provides material support to hateful organisations, individuals who cause serial or mass violence or promote hateful ideologies, violent criminal organisations or violent extremists. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. In addition, TikTok operates a process to determine if a group is a terrorist group, so that it can implement effective detection and moderation strategies (which involves identifying terrorist groups, [REDACTED] etc.). Further, in order to maintain its detection models, TikTok works with various external parties to keep its keyword lists relevant to Terrorist Content up to date. TikTok also maintains a “blacklist” of hyperlinks, which have been found to be associated with previously detected Terrorist Content, so that it can detect, disrupt and remove such links.

TikTok also works closely with its fact-checking partners to conduct trend analysis and address misinformation fueling extremist content on the Platform. TikTok systematically identifies, categorises and prioritises online trends relating to Terrorist Content on a recurring basis and operates measures to contain these trends.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. TikTok also deploys a model [REDACTED] to mitigate the risk that harmful content appears on or is amplified by the FYF.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not promote or support terrorist organisations or otherwise violate TikTok’s Community Guidelines or Advertising Policies. TikTok also has specific moderation guidance for addressing Terrorist Content and related content within ads.

**Article 35(1)(f):** TikTok has specialist teams who deal with illegal content, including Terrorist Content, and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(i):** TikTok’s awareness-raising measures include (i) TikTok’s Transparency Centre, which has dedicated content explaining [TikTok’s approach to content moderation](#) and [Combating violent extremism content online](#); (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; and (iii) TikTok’s other in-app educational and awareness measures (including measures to empower users with credible information during times of crisis), verified badges, state-controlled media labels, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** TikTok dedicates significant resources to removing illegal content and seeks to remove it before any user, of any age, is exposed to it. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.



TikTok has further enhanced its mitigation measures related to Terrorist Content since its Year 2 Report, including:

- TikTok refined its policies on violent extremist and polarizing content to prioritize moderation enforcement based on the severity and nature of the underlying risk.
- TikTok has been introducing new moderation measures to improve precision in detection of and enforcement against violent extremism.
- TikTok has launched new measures to prevent users from circumventing account bans or restrictions, including bans for violent extremism.
- TikTok maintains search guides for users searching for information on terrorist organisations and crisis events. This year, TikTok launched a search guide for the Örebro Shooting in Sweden to limit the spread of misinformation and extremist content.
- TikTok also updated its dedicated Transparency Centre page, focused on [combating violent extremism](#).

#### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- Membership of Tech Against Terrorism, which supports the tech industry in tackling terrorist exploitation of the internet, whilst respecting human rights. Membership in Tech Against Terrorism provides practical and operational support to the Trust & Safety teams tasked with preventing violent extremists from using the Platform to perpetrate harm. TikTok works with them to strengthen its policies and stay up to date on current trends and scholarship.
- Engagement with counter-terrorism organisations in relation to the complexities of livestreaming and other key issues.
- Participation in the UN Forum on Business and Human Rights and RightsCon, a major global summit on digital human rights, including engagement in dialogue on digital rights and the role of tech companies in upholding human rights.
- TikTok actively participated in the 2025 Referral Action Day, which was coordinated by Europol's European Counter Terrorism Centre, and focused on violent extremist and terrorist propaganda targeting minors.
- TikTok participated in various meetings and initiatives in connection with the EU Internet Forum, including knowledge sharing on violent right-wing extremism and in respect of the TCO Regulation.
- TikTok also continues to engage in outreach with law enforcement authorities in the EU and elsewhere (including specialist counter-terrorism units across the EU) in relation to Terrorist Content and associated risks.
- As members of the Global Internet Forum to Counter Terrorism (GIFCT), TikTok collaborates with industry partners, governments and civil society to take swift, coordinated action against terrorist and violent extremist content. Through GIFCT, TikTok shares best practices, participates in joint incident responses and contributes to transparency and safety initiatives in line with regulatory frameworks such as the DSA.

#### 4. Residual Risk

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of Terrorist Content to be **Medium**. This Residual Risk Score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 517,799 videos under its "Violent and Hateful Organizations and Individuals" policy; 495,308 of those videos were detected and removed proactively (i.e. before any user report), 347,186 were removed before being viewed by any user and 381,084 were removed within 24 hours of publication. TikTok also removed 207,180 videos under its "Violent and Criminal Behavior" policy; 188,501 of those videos were detected and removed proactively (i.e. before any user report), 142,105 were removed before being viewed by any user, and 178,674 were removed within 24 hours of publication, demonstrating TikTok's swift action to combat Terrorist Content on the Platform.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, following a user report, 22,491 videos were removed under TikTok's "Violent and Hateful Organizations and Individuals" policy, while 18,679 were removed under its "Violent and Criminal Behavior" policy.

A number of TikTok's other policies may capture Terrorist Content, including its "Safety and Civility - Hate Speech and Hateful Behavior" policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the risk of Terrorist Content on the Platform.

As noted in TikTok's Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok received 11,486 reports from users of allegedly illegal content relating to Terrorist Offences/Content. During the same period, it received 1 Trusted Flagger report relating to Terrorist Offences/Content. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 367 of which related to Terrorist Offences/Content.

TikTok publishes transparency reports pursuant to the TCO Regulation. As reported in TikTok's 2025 Transparency Report, which covers the period 1 January to 31 December 2024, TikTok received 190 removal orders from competent authorities specifically under the TCO Regulation, and access to the reported content was blocked on the Platform. No administrative or judicial review proceedings were issued and no cases required the reinstatement of content as a result of such proceedings.

In addition, TikTok notes that, in November 2024, the Platform was one of several platforms designated by Coimisiún na Meán under the TCO Regulation, and accordingly, TikTok has

implemented a range of specific measures to protect the Platform against the dissemination of Terrorist Content and to ensure compliance with the TCO Regulation.

External world events can have a bearing on the type and volume of Terrorist Content TikTok observes on the Platform. As such, TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve its automated and manual moderation processes and to improve information and transparency provided to users. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 9. ILLEGAL CONTENT: RISKS OF INTELLECTUAL PROPERTY INFRINGING CONTENT

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the term "**IP-Infringing Content**" to mean content that is created and disseminated in breach of copyright or other intellectual property ("**IP**") rights. TikTok's approach to actioning IP-Infringing content is structured in particular in accordance with relevant EU laws on copyright, including Directive (EU) 2019/790 (the "**Copyright Directive**"). In addition, TikTok notes that Article 17 of the Charter enshrines the right to protection of intellectual property.

#### Summary of risks

The risks associated with IP-Infringing Content may arise from users attempting to share the following content on or through the Platform, including through video or photo, livestream, comments or in profile information or TikTok Shop features on the Platform:

- Content that reproduces the original work of another person or entity without that person's or entity's permission and which does not fall in one of the copyright exceptions. This content may include music works and audio files, artistic works (e.g. photographs, paintings, drawings and other original visual renderings) and audio-visual recordings;
- Content that contains the unauthorised use of a trademark, service mark or logo in connection with goods or services in a way that is likely to cause confusion, deception or mistake about the source or affiliation of the associated goods and/or services;
- Content that infringes on other intellectual property rights, for example, a patent or trade secret;
- Content that advertises or promotes IP-infringing products, including products that can be purchased on TikTok Shop or third-party platforms;
- AIGC content that is IP-Infringing Content. This can be challenging to moderate as it can be difficult for rights holders to demonstrate that their copyright has been infringed by AIGC and therefore difficult for moderators to make an assessment on infringement; or

- Product listings on TikTok Shop, offering IP infringing products for sale including counterfeit products.

There are attendant risks that, in some cases, TikTok may not remove infringing content quickly enough, or may inaccurately remove content that the rightsholder wishes to remain online. For example, TikTok has consistently observed abuses of its IP reporting processes by users attempting to have non-infringing content removed from the Platform (consistent with concerns published widely about abuses relating to takedown processes).

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in the inherent risk profile for IP-Infringing Content and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium</b>

This Inherent Risk Score is consistent with TikTok's score in Year 2.

## 3. Mitigation Measures

TikTok clearly prohibits IP-Infringing Content on the Platform. TikTok undertakes the following mitigation measures related to IP-Infringing Content risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of the dissemination of IP-Infringing Content on or through the Platform, including (i) blocking certain search terms to limit the manual discoverability of IP-Infringing content on the Platform; (ii) measures to mitigate the risk of a livestream being capable of containing IP-Infringing Content; (iii) measures to address the risk of hyperlinks linking to third-party sites containing IP-Infringing Content; (iv) measures to block IP-infringing listings being listed on Shop; (v) measures to detect and remove IP-infringing listings on Shop; and (vi) dedicated reporting forms for rightsholders to make claims to protect their IP.

**Article 35(1)(b):** TikTok clearly prohibits IP-Infringing Content on the Platform through a combination of TikTok's [Terms of Service](#), [Community Guidelines](#) and [Intellectual Property Policy](#). For example, TikTok's "Integrity and Authenticity - Unoriginal Content" Policy prohibits content that violates someone else's copyrights, trademarks or other intellectual property rights and TikTok's "Regulated Goods and Commercial Activities – Trade of Regulated Goods and Services" policy prohibits counterfeit products. TikTok also has a number of TikTok Shop Policies, which mitigate the risk of IP-Infringing Content on TikTok Shop. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 ("About TikTok") for a description of TikTok's content moderation processes. In addition, TikTok's "notice and takedown" process is designed to address IP-Infringing Content effectively. Users can report [copyright](#) and [trademark](#)

infringements (including through UGC, ads and Shop) through dedicated webforms. Additionally, for Shop, TikTok maintains a dedicated portal known as the Intellectual Property Protection Centre (“**IPPC**”), which enables rightsholders to bulk report infringing listings for design, patent, trade mark or copyright infringement. Under TikTok’s “notice and takedown” process, remedies include content removal and account termination. For copyright infringement on UGC specifically, TikTok provides “stay-down” functionality that blocks subsequent attempts by users to upload previously identified copyright-infringing content.

Rightsholders may use TikTok’s blocking webform or other copyright tools, which are designed to ensure the unavailability of protected works on the Platform where rightsholders have provided TikTok with relevant and necessary information. Content protected by these tools cannot be accessed by users in the region(s) in which it is blocked.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. In addition, reproduced or unoriginal content that is imported or uploaded without any new or creative edits (such as content with someone else’s visible watermark or superimposed logo) is ineligible for the FYF, in accordance with TikTok’s Community Guidelines.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not infringe on the intellectual property rights of others or otherwise violate TikTok’s Community Guidelines or Advertising Policies.

**Article 35(1)(f):** TikTok has specialist teams who deal with illegal content, including IP-Infringing Content, and who play a role in policy development, risk management and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(i):** TikTok’s awareness-raising measures include (i) TikTok’s Transparency Centre, which contains TikTok’s bi-yearly reports on intellectual property removal requests; and (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content. The Help Centre also provides broad overviews of [IP](#), [copyright](#) and [trademark and counterfeiting](#) and how TikTok protects these rights on the Platform.

**Article 35(1)(j):** TikTok’s measures to mitigate the risks related to IP-infringing content on the Platform are designed to protect users of all ages.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok also has internal guidelines for processing IP reports made against AIGC. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.

TikTok has further enhanced its mitigation measures related to IP-Infringing Content since its Year 2 Report, including:

- TikTok has expanded its in-house solution for copyright protection, which makes copyright works unavailable in respect of UGC content on the Platform where users have provided TikTok with relevant information.
- TikTok has refined its Advertising Policies to enhance enforcement and better align its practices with user-expectations.
- TikTok expanded its engagement processes with priority rightsholders, known to be targets of impersonation reports, to ensure legitimacy of their reports.
- In October 2024, TikTok launched educational resources regarding the dangers of counterfeit products in its Safety Centre to raise awareness among its user-base.

#### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- With respect to Shop, TikTok actively maintains relationships with key stakeholders including brands, their service providers, industry associations and governmental agencies. TikTok Shop has built a global team which works with stakeholders to combat IP rights issues across all of its global markets.
- IPPC portal engagement strategies: Through outreach, TikTok encourages the use of reporting tools like the IPPC and supports brand owners with the resolution of complex issues. To ensure awareness of its IPPC platform and understand its benefits and how to use it, TikTok conducts events and meetings face-to-face, often through trusted partners. Since the launch of Shop, TikTok has signed memorandums of understanding with several governmental agencies.
- TikTok Shop and the International Olympic Committee ("**IOC**") worked together to ensure that the IOC's intellectual property would be comprehensively protected during a period of heightened interest in the Games and related merchandise. The TikTok Shop IPR team received information from the IOC on their IP portfolio, allowing the implementation of proactive measures aimed at preventing the sale of infringing products globally.
- TikTok Shop partnered with the Anti-Counterfeiting Group ("**ACG**") on a dedicated in-person and online workshop. The event provided ACG members with information on TikTok Shop IP strategy, policies and tools and insights into common pitfalls and tips.
- TikTok Shop provided an Online Takedown Procedures Certificate Program for more than 150 International Trademark Association members comprising brand owners, law firms and legal and brand protection subject matter experts. The program focused on IPPC account creation, establishing IP assets and notice and takedown procedures.
- TikTok Shop supported Unifab, the French brand association, with its counterfeiting consumer awareness campaign on TikTok during Black Friday and end-of-year sales by providing ad credits.



#### 4. Residual Risk

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of IP-Infringing Content to be **Medium**. Taking a conservative approach, TikTok has revised its Residual Risk Score from Year 2 (Low) to primarily take account of the launch of TikTok Shop in the EU and an increase in IP-violative AIGC content.

This Risk Assessment was informed by data from TikTok's Intellectual Property Removal Requests Report for the period 1 July to 31 December 2024, which reports on copyright and trademark removal requests processed by TikTok and content removed as a result. During that period, TikTok received 280,259 copyright removal requests globally, with 194,405 resulting in content removal. TikTok also received 42,407 trademark removal requests globally, with 29,318 resulting in content removal. These figures relate to content not otherwise detected and actioned by TikTok's automated and manual moderation measures. Further, as a result of copyright and trademark requests, TikTok took action against 253,218 accounts.

A number of TikTok's other policies capture IP-Infringing Content, including its "Regulated Goods and Commercial Activities – Trade of Regulated Goods and Services" policy. TikTok's Advertising Policies also prohibit ads that infringe the intellectual property rights of others. As noted in TikTok's Fourth DSA Transparency Report, which covers content in the EU during the period Q3 through Q4 2024, TikTok removed 45,998 videos for infringement of its "Intellectual Property Infringement" Advertising Policy.

This Risk Assessment was also informed by data on TikTok's enforcement of its TikTok Shop Policies in the EU in Q1 of 2025 (TikTok Shop launched in the EU in December 2024). There were a total of 1,235,990 product listings available on TikTok Shop in the EU in March 2025. Over the course of Q1 2025, 18,711 listings were deactivated after being live, including 284 listings deactivated for violations of TikTok Shop Policies relating to IP. In addition, as of March 2025, there were 288,457 active sellers on TikTok Shop in the EU. During Q1 2025, 50 sellers were deactivated, including 1 seller deactivated for a violation of TikTok Shop policies relating to IP.

TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve policy and moderation measures to protect against IP-Infringing Content on the Platform. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.



## 10. ILLEGAL CONTENT: RISKS OF CHILD SEXUAL ABUSE MATERIAL AND CHILD SEXUAL EXPLOITATION AND ABUSE

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the terms child sexual abuse material (“**CSAM**”<sup>2</sup>) and child sexual exploitation and abuse (“**CSEA**”) in a manner consistent with EU and EU member state laws. TikTok understands scope of CSAM and CSEA with particular regard to Directive 2011/93/EU, Articles 3 to 7, in relation to: offences concerning sexual abuse; offences concerning sexual exploitation; offences concerning child pornography; solicitation of children for sexual purposes; and incitement, aiding and abetting, and attempts to commit such offences.

TikTok notes that Article 34 of the United Nations Convention on the Rights of the Child enshrines the right of the child to protection from all forms of sexual exploitation and abuse, and that such risks should be assessed with the best interests of the child as the primary consideration, in accordance with Article 34 of the Charter. TikTok notes that in 2021, the United Nations Committee on the Rights of the Child underlined that these rights must be equally protected in the digital environment.<sup>3</sup>

#### Summary of risks

TikTok considers that the dissemination of CSAM content may involve users attempting to use the Platform to:

- Share, re-share or offer to trade or sell, or direct users of the Platform to obtain or distribute CSAM content;
- Generate and/or share CSAM, including self-generated CSAM;
- Disseminate content that depicts, solicits, glorifies or encourages child abuse imagery including nudity, sexualised minors or sexual activity with minors; or
- Disseminate content that depicts, promotes, normalises or glorifies paedophilia or the sexual assault of a minor.

TikTok considers that CSEA behaviour may involve Adult Users attempting to carry out or participate in CSEA, including using the Platform to:

- Build an emotional relationship with a minor in order to gain the minor’s trust for the purposes of future or ongoing sexual contact, sexual abuse, trafficking or other exploitation;
- Solicit real-world contact between a minor and an adult or between minors with a significant age-difference;
- Solicit minors to connect with an adult on another online platform, website or other digital space for sexual purposes;

---

<sup>2</sup> Although various legal texts still refer to the term “child pornography,” in line with the Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse (known as the “Luxembourg Guidelines”), TikTok considers CSAM to be the more appropriate term.

<sup>3</sup> [UN General Comment No. 25 \(2021\)](#) on Children’s Rights in Relation to the Digital Environment.

- Solicitation of nude imagery or sexual contact, through blackmail or other means of coercion;
- Carry out sextortion or grooming, which can lead to CSAM risks as well as further online and offline child sexual abuse or sex trafficking, or other risks to the safety and wellbeing of users; or
- Post other content that involves CSEA.

TikTok continues to observe risks associated with subtle and deceptive tactics used by predators, including fake profiles, covert behavioural signals and evolving abuse patterns.

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for CSAM and CSEA and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

TikTok has revised its Inherent Risk Score from Year 2 (Medium), based on growing external threats in the digital space relating to sextortion and grooming, and increased challenges that digital platforms are facing posed by AI-generated violative content.

## 3. Mitigation Measures

TikTok very clearly prohibits CSAM and CSEA as well as certain related content and behaviour on the Platform. TikTok undertakes the following mitigation measures related to CSAM and CSEA risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of CSAM and CSEA on or through the Platform, including: (i) measures to mitigate the risk of a livestream being capable of containing CSAM or used to perpetrate CSEA, including restricting the ability to host livestreams to Adult Users; (ii) measures to address the risk of comments or hyperlinks being used to disseminate CSAM or CSEA Content; (iii) moderation and controls related to TikTok Shop features; (iv) age-assurance measures and various settings for 13-17 year olds to limit who can view their content and interact with them on the Platform and what features they can access; (v) settings for users aged 13-15 which make their account private by default, their content ineligible for the FYF, their profile ineligible for recommendation, and restrict certain features, including video downloads (i.e. the ability for others to download their content), direct messaging and duet and stitch; (vi) settings for users aged 16-17, including defaulting video downloads to “off” and providing pro-privacy nudges to encourage these users to consider their privacy settings; (vii) Family Pairing features which allow parents/guardians to manage certain settings relating to the content their teenager sees on the Platform, the time they spend on the Platform and the features they can access on the Platform; and (viii) search interventions to block harmful keywords

and provide deterrence messaging and support resources for individuals with sexual attraction to minors.

**Article 35(1)(b):** TikTok very clearly prohibits CSAM and CSEA as well as certain related content and behaviour on the Platform through a combination of TikTok’s [Terms of Service](#) and the [Community Guidelines](#). For example, under its “Safety and Civility - Youth Sexual and Physical Abuse” policy, TikTok prohibits content or behaviour showing, promoting or engaging in youth sexual or physical abuse or exploitation, including CSAM, paedophilia, grooming behaviour, sextortion, sexual harassment and sexual solicitation. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. As part of TikTok’s content moderation processes, TikTok checks images against known CSAM content [REDACTED] and keywords that are associated with CSAM. Further, TikTok maintains a purpose-built computer vision model, which aims to identify visual material that may contain CSAM or CSEA and route such material to TikTok’s specialist teams for review. TikTok also integrates third party hash lists to enable it to detect known CSAM at the point of upload.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. In addition, TikTok has protocols in place to quickly detect and control trends featuring harmful user content (which can entail banning certain hashtags and/or blocking searches associated with emerging risks).

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not contain minor sexual abuse material involving real or non-real minors, contain inappropriate skin exposure of real or non-real minors or otherwise violate TikTok’s Community Guidelines or Advertising Policies. TikTok also has Advertising Policies prohibiting the appearance of minors in certain ads.

**Article 35(1)(f):** TikTok has specialist teams, which include individuals with law enforcement, intelligence and public safety backgrounds, who deal with illegal content, including CSAM and CSEA-related content, and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures. TikTok has a specialist team which is responsible for escalating child safety issues, such as suspected grooming incidents, and for making reports to the National Centre for Missing & Exploited Children and law enforcement authorities.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(h):** TikTok is committed to the “Voluntary Principles to Counter Child Sexual Exploitation and Abuse” from the WeProtect Global Alliance, of which it is a signatory, in order to help build collaboration across platforms. The WeProtect Global Alliance, which aims to protect children from online sexual exploitation and abuse, fosters collaboration across sectors, including NGOs, safety firms and governments, facilitating the sharing of knowledge and best practices and strengthening the collective response to CSEA.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which includes a page on sexual abuse support with instructions on reporting sexual content involving minors. Users are directed to this page if they search for CSAM-related terms on the Platform. The Safety Centre also contains user-friendly resources to help Younger Users to manage their interactions online, understand the tools available to them and maintain their digital wellbeing; (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; (iii) TikTok’s Transparency Centre, which contains user-friendly articles explaining TikTok’s approach to Keeping People Safe, covering topics like content moderation and protecting teens online and outlining TikTok’s methods to combat child exploitation and abuse; and (iv) TikTok’s other in-app educational and awareness measures, verified badges, state-controlled media labels, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** By their nature, all measures that mitigate risks of CSAM and CSEA are measures to protect the rights of the child. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok’s Edited Media and AIGC policy also prohibits AIGC content that shows certain depictions of minors. TikTok also takes a range of measures to mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities, which could be used to establish trust and exploit victims. TikTok monitors to detect the creation of inauthentic accounts and inauthentic behaviour on the Platform.

TikTok has further enhanced its mitigation measures related to CSAM and CSEA since its Year 2 Report, including:

- TikTok is undergoing a phased launch of [REDACTED]
- TikTok has expanded its Family Pairing features to include new visibility tools, which allow parents/guardians to see who their teen is following on TikTok and who follows them, as well as blocked accounts.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with members of TikTok’s Safety Advisory Councils on sexual abuse.
- Engagement in outreach with law enforcement authorities in the EU and elsewhere (including specialist child safety units) in relation to CSEA and CSAM and associated risks, in particular with regard to high-priority child safety matters.
- TikTok meets regularly with NCMEC and regularly receives information on emerging trends and feedback from law enforcement.
- Policy consultations with experts from organisations focused on victim support and understanding indications of potential predatory behaviour.
- Policy consultation with a nonprofit organisation committed to preventing child sexual abuse.

- TikTok is an active member of a number of expert child safety groups, including the Technology Coalition, an industry group dedicated to fighting child sexual exploitation and abuse online, WeProtect Global Alliance (described above) and the Internet Watch Foundation, an organisation committed to identifying and removing online records of CSAM.
- TikTok has ongoing partnerships with other leading online safety organisations, to help ensure that TikTok’s policies and features continue to promote a safe and welcome environment for its users.
- TikTok attended a number of conferences and events focused on countering child exploitation and abuse including the WeProtect Global Summit, TrustCon and the INHOPE Summit.
- TikTok also hosted a workshop on combatting grooming and sextortion which involved knowledge-building with external researchers.

#### 4. Residual Risk

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of CSAM and CSEA to be **Medium**. This Residual Risk Score is consistent with TikTok’s score in Year 2.

This Risk Assessment was informed by data on TikTok’s enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 1,507,542 videos labelled under its “Safety and Civility – Youth Sexual and Physical Abuse” policy; 1,488,738 of those videos were detected and removed proactively (i.e. before any user report), 1,382,920 were removed before being viewed by any user and 1,457,609 were removed within 24 hours of publication, demonstrating TikTok’s swift action to combat CSAM and CSEA.

The rate of proactive removals is a positive indicator of the efficacy of TikTok’s content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, 18,804 videos were removed under TikTok’s “Safety and Civility – Youth Sexual and Physical Abuse” policy following a user report.

A number of TikTok’s other policies capture CSAM and CSEA, including its “Sensitive and Mature Themes – Sexually Suggestive Content” policy. TikTok’s policy approach works in combination with its other mitigation measures to mitigate the risk of CSAM and CSEA on the Platform.

As noted in TikTok’s Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok received 14,325 reports from users of allegedly illegal content related to child sexual exploitation. During the same period, it received 14 reports from Trusted Flaggers related to child sexual exploitation. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 1 of which related to child sexual exploitation.

TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve its automated and manual moderation processes and enhance its in-app safety features to protect against CSEA and CSAM on the Platform. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 11. ILLEGAL CONTENT: RISKS OF GENDER-BASED VIOLENCE CONTENT

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands that neither the DSA nor EU law contains a single or comprehensive definition of gender-based violence. TikTok therefore understands gender-based violence to be the perpetration, support or incitement of “any type of harm [...] against a person or group of people because of their factual or perceived sex, gender [...] and/or gender identity” (“GBV”).<sup>4</sup>

TikTok has regard to Directive (EU) 2024/1385 on combating violence against women and domestic violence, which defines “violence against women” as “all acts of gender-based violence directed against a woman or a girl because she is a woman or a girl or that affect women or girls disproportionately, that result in or are likely to result in physical, sexual, psychological or economic harm or suffering, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life.” This definition is consistent with and encompassed by TikTok's understanding of GBV.

In addition, TikTok notes that Art. 21 of the Charter prohibits discrimination based on sex and sexual orientation, and that Art. 23 enshrines the right to equality between men and women. TikTok further notes that GBV, in particular violence against women and domestic violence, is a violation of fundamental rights such as the right to human dignity, the right to life and integrity of the person, the prohibition of inhuman or degrading treatment or punishment, the right to respect for private and family life, the right to liberty and security, the right to protection of personal data, the right to non-discrimination, including on the grounds of sex, and the rights of the child as enshrined in the Charter.<sup>5</sup>

---

<sup>4</sup> The Council of Europe's [definition](#) is based on the [Explanatory Report](#) to the Convention on preventing and combating violence against women and domestic violence; note that TikTok considers violative content related to hate speech against people due to their sexual orientation in the separate Hate Speech section of this Report.

<sup>5</sup> See [Directive \(EU\) 2024/1385 of the European Parliament and of the Council on combating violence against women and domestic violence](#), Recital (3); Recital (6) (“Violence against women and domestic violence can be exacerbated where it intersects with discrimination based on a combination of sex and any other ground or grounds of discrimination as referred to in Article 21 of the Charter”).



### Summary of risks

The risks relating to GBV may arise from users attempting to share or disseminate content depicting or involving the following types of behaviour on or through the Platform, including through video or photo, livestream, comments, profile information or TikTok Shop features:

- Non-consensual sexual acts that are real or fictional including rape, molestation and non-consensual touching;
- Image-based sexual abuse (“**IBSA**”), including AI-generated IBSA, and sextortion;
- Sexual harassment;
- Gender-based hate speech, including the promotion of violence, exclusion, segregation, discrimination and other harms on the basis of gender, gender identity or sex or certain hateful ideologies, including male supremacy or misogyny;
- Violent behaviour where gender is a relevant factor, including intimate partner violence or threatening or expressing a desire to cause physical injury to a person or a group on the basis of gender; and
- Harassment and bullying where gender is a relevant factor, such as degrading someone or expressing disgust on the basis of their personal characteristics or circumstances, such as their physical appearance, intellect, personality traits and hygiene (where gender is a relevant factor) (together, “**GBV Content**”).

TikTok is aware of the risk that users may use fake accounts to post harmful content or harass, silence or discredit victims of GBV.

## **2. Inherent Risk**

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for GBV content, and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

TikTok has revised its Inherent Risk Score from Year 2 (Medium), based on risks relating to the use of AI-generated content, including content generated off-Platform by so-called “nudification apps.” Sextortion is also a growing concern globally and there has been an increase globally in the emergence of harmful gender narratives, including misogynistic content.

## **3. Mitigation Measures**

TikTok very clearly prohibits GBV Content, as well as certain related content, on the Platform. TikTok undertakes the following mitigation measures related to GBV Content:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of the dissemination of GBV Content on or through the Platform, including (i) blocking search results for certain high-risk terms (for example, terms associated with misogyny); (ii)



measures to mitigate the risk of a livestream being capable of containing GBV Content; (iii) measures to address the risk of comments or hyperlinks being used to disseminate GBV Content (including the “consider before you comment” prompt, to remind users to reconsider before posting potentially harmful comments); (iv) defaulting Duet and Stitch features to “off” which provides users with control each time they post a video and limits the opportunity for content to be used to create GBV content; (v) defaulting the filter for spam and offensive or harmful comments to “on” for all users; and (vi) user tools, which allow creators to filter, review and approve comments before they are publicly visible under their content, including a filtering tool for keywords containing user-selected terms, the “Unwanted Comments Filter,” which automatically detects and hides spam or generally disliked comments, and the Creator Care mode which hides comments a creator is likely to find offensive, based on prior interactions.

**Article 35(1)(b):** TikTok very clearly prohibits GBV Content, as well as certain related content, on the Platform through a combination of TikTok’s [Terms of Service](#) and the [Community Guidelines](#). For example, under its “Safety and Civility - Adult Sexual and Physical Abuse” policy, TikTok prohibits showing, promoting or engaging in sexual harassment, non-consensual sex acts, image-based sexual abuse, physical abuse and sextortion. Further, under its “Hate Speech and Hateful Behavior” policy, TikTok prohibits the promotion of hateful ideologies, including misogyny, and content or behaviour that dehumanises someone on the basis of a protected attribute. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. In addition, in order to maintain its detection models, TikTok works with various external parties to keep its keyword lists relevant to GBV Content up to date. TikTok also collaborates with leading image abuse prevention organisations, which enables victims of non-consensual intimate image abuse to request their content be removed from any platform that finds it, and the NCMEC’s “Take it Down” initiative, which is a similar service that enables the hashing and automated removals of image-based sexual abuse content. TikTok also implements moderation strategies to recall potentially violative content for human moderation.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not display content that is derogatory, inflammatory, demeaning, distasteful, disrespectful, profane or which contains coarse language or insulting actions or violate any other existing Community Guidelines or Advertising Policies.

**Article 35(1)(f):** TikTok has specialist teams who deal with illegal content, including GBV content, and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted

subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(h):** TikTok signed the EU Code of Conduct on Hate Speech in September 2020 and continues to be evaluated on a yearly basis by participating NGOs from the European region. The evaluation seeks to understand how quickly and reliably platforms respond to reports on hate speech content. In January 2025, the EU Code of Conduct on Countering Illegal Hate Speech Online was implemented as a Code of Conduct under the DSA.

**Article 35(1)(i):** TikTok's awareness-raising measures include (i) TikTok's Transparency Centre, which has dedicated content on Keeping People Safe, [TikTok's approach to content moderation](#), [Combating Hate](#) and [Upholding Rights](#); (ii) TikTok's Help Centre, containing "how to" explanations to allow users to learn about its content moderation practices and how to report violative content; (iii) TikTok's Safety Centre, which includes a Sexual abuse support page, containing resources on how to access help, links to trusted external reporting platforms, helplines and clear reporting channels for victims of image-based sexual abuse; and (iv) TikTok's other in-app educational and awareness measures, verified badges, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** TikTok dedicates significant resources to removing illegal content and seeks to remove it before any user, of any age, is exposed to it. Additional Youth Safety measures are set out in the "Youth Safety and Online Engagement" section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok's controls related to AIGC. In addition, TikTok's Community Guidelines on Adult Sexual and Physical Abuse prohibit Image-based sexual abuse, whether the images are real or altered. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities.

TikTok has further enhanced its mitigation measures related to GBV Content since its Year 2 Report, including:

- In an update to TikTok's Community Guidelines, TikTok indicated clearly that it prohibits image-based sexual abuse regardless of whether the image is real or altered (by AI or other digital alterations). TikTok has updated its policies to better address harassment, exploitation, bullying and sextortion networks.
- TikTok has introduced targeted detection and enforcement tools relating to hate accounts to better detect patterns of gender-based hate that may not be obvious in isolation and to speed up enforcement.
- TikTok expanded the use of specialised moderation queues to better detect and enforce against GBV. These are operated by trained moderators who receive targeted guidance on identifying gender-based harms while respecting freedom of expression in line with TikTok's policies. [REDACTED]
- TikTok updated its machine learning models to incorporate insights from spikes in GBV observed [REDACTED].

This has improved TikTok's ability to detect and act on nuanced gendered harms.

- TikTok introduced a two-tier comment moderation framework for violative and non-violative comments. TikTok removes violative comments through the standard enforcement process, while creators are provided with a range of self-moderation tools to manage non-violative, but potentially harmful, comments.
- TikTok has conducted research and development into measures to combat harmful gender narratives.
- TikTok has introduced account-level policies to address sextortion risks.

#### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with members of TikTok's Safety Advisory Councils on sexual abuse.
- Policy consultations with experts in relation to adult sexual abuse and domestic violence.
- Policy consultations with experts from victim support organisations in relation to detecting predators.
- Participation in the NO MORE Week Tech Conference, including engagement in discussions on technology-facilitated GBV and TikTok's related policies.
- Participation in the [REDACTED] that meets quarterly to discuss solutions to technology-facilitated GBV on online platforms.
- TikTok hosted a roundtable on gendered harassment and the impact of misogynistic content on minors with a number of organisations who work to prevent sexual harassment and abuse.

#### **4. Residual Risk**

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of GBV Content to be **Medium**. This Residual Risk score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 944,915 videos under its "Safety and Civility - Adult Sexual and Physical Abuse" policy; 801,745 of those videos were detected and removed proactively (i.e. before any user report), 346,214 were removed before being viewed by any user and 756,825 were removed within 24 hours of publication, demonstrating TikTok's swift action to combat GBV.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, 143,170 videos were removed under TikTok's "Safety and Civility - Adult Sexual and Physical Abuse" policy following a user report.

A number of TikTok's other policies capture GBV, including its "Safety and Civility - Hate Speech and Hateful Behavior" policy and its "Safety and Civility - Harassment and Bullying"

policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the risk of GBV on the Platform.

As noted in TikTok's Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok received 24,247 reports from users of allegedly illegal content related to non-consensual sharing of private or intimate images. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 17 of which related to non-consensual sharing of private or intimate images.

External events can have a bearing on the type and volume of GBV Content TikTok observes on the Platform. As such, TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve detection of and enforcement against GBV and has continued to invest in researching and developing new solutions to target GBV. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 12. ILLEGAL CONTENT: RISKS OF ILLEGAL AND UNSAFE PRODUCTS

### 1. Description of the Risk and how it may manifest on the Platform

As explained above, this is a new Risk Module for Year 3, which TikTok has introduced to continue to build on its risk assessment process.

TikTok understands "**Illegal and Unsafe Products**" to include items prohibited by law in certain countries or regions, banned goods and substances, products hazardous to adults and children, products that have been recalled by product safety authorities due to product safety concerns and other goods prohibited under TikTok's Community Guidelines.

#### Summary of risks

Risks regarding the promotion or sale of Illegal and Unsafe products may manifest on the Platform in the following ways, including through TikTok Shop, video or photo, livestream, comments, profile information and Search:

- Illegal or Unsafe Products being sold on TikTok Shop in the EU;
- Illegal or Unsafe Products being promoted on the Platform, including through: (i) "**Shop Content**," i.e. shoppable videos and livestreams that can be uploaded to the Platform or streamed by Sellers and Creators and which contain a link to product description pages ("**Product Listing**") or a Seller's product inventory from where individual Product Listings can be accessed ("**Product Showcase**"); (ii) links to third-party websites; and (iii) other content promoting Illegal or Unsafe Products; and

- Illegal and unsafe products being discovered by users on the Platform.

## 2. Inherent Risk

TikTok has considered both severity and probability in assessing the inherent risk profile for Illegal and Unsafe Products. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

To provide context for this year's (new) risk level, TikTok notes the following:

- The sale of Illegal and Unsafe Products, including via an online platform, can cause (severe) physical harm to users, including minors.
- Without any controls in place, including age-based controls and policies and procedures that limit the Platform from being used to promote and sell Illegal and Unsafe Products, the likelihood of such promotion and sales taking place is likely.

## 3. Mitigation Measures

TikTok very clearly prohibits Illegal and Unsafe Products, as well as certain related content, on the Platform. TikTok undertakes the following mitigation measures related to Illegal and Unsafe Product risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the risk of the sale of Illegal or Unsafe Products on or through the Platform, including: (i) seller onboarding that includes verification of sellers' personal and corporate identity and address, applicable regulatory checks and use of information from third-party risk vendors; (ii) "block lists" to prevent sellers previously blocked for fraudulent behaviour or repeated policy violations from rejoining; (iii) creator controls that includes ID validation and verification of age checks before creators can share Shop Content; (iv) blocking certain search results associated with Illegal and Unsafe Products; (v) measures to mitigate the risk of a livestream being capable of promoting Illegal or Unsafe Products; and (vi) measures to mitigate the risk of comments or hyperlinks promoting or selling Illegal or Unsafe Products.

**Article 35(1)(b):** TikTok very clearly prohibits Illegal and Unsafe Products, as well as certain related content, on the Platform through a combination of TikTok's [Terms of Service](#), the [Community Guidelines](#) and Shop terms and policies. For example, under its "Regulated Goods and Commercial Activities" policy, TikTok prohibits facilitation of the trade of and the marketing of firearms or explosive weapons as well as the marketing of regulated, prohibited or high-risk goods and services, such as alcohol and tobacco products, illegal drugs, prescription and over-the-counter drugs, weight loss or muscle gain products and other regulated substances. TikTok's Shop policies also prohibit the sale of dangerous or unsafe goods, including drugs, knives, gambling products and firearms. In addition, TikTok's

policies, including its Branded Content Policy, require users to disclose when they promote a brand, product or service in exchange for payment or other incentive. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. In addition, TikTok has detection and reporting mechanisms to flag potential violative content within Product Listings on Shop and takes steps to identify recalled and illegal products on the Platform, and where necessary, take action in relation to such products. TikTok also reviews buyer feedback and complaints to identify potentially violative product listings.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. In the context of TikTok Shop, TikTok’s diversification rules operate to avoid users being recommended a series of Shop Content in the FYF.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not promote or advertise dangerous products or otherwise violate TikTok’s Community Guidelines or Advertising Policies.

**Article 35(1)(f):** TikTok has specialist teams who deal with illegal content, including content related to Illegal and Unsafe Products.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers. TikTok reports on Article 16 illegal content reports received from Trusted Flaggers in its DSA Transparency Reports.

**Article 35(1)(i):** TikTok’s awareness-raising measures include (i) TikTok’s Transparency Centre, which has dedicated content explaining [TikTok’s approach to content moderation](#) and [Upholding human rights](#), which links to the [TikTok Shop Business Partner Code of Conduct](#) and [the TikTok Shop Safety Report](#); (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; (iii) TikTok Shop Academy, which is a centralized hub, providing content to educate TikTok’s community on its policies, including those on illegal activity, prohibited products and restricted products; and (iv) TikTok’s other in-app educational and awareness measures, verified badges, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** TikTok dedicates significant resources to removing illegal content and seeks to remove it before any user, of any age, is exposed to it. In addition, users must be 18 years of age or older to use TikTok Shop. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement relevant to this risk includes:



- TikTok participates as an active member of numerous national trade associations focused on digital policy, including eCommerce, as well as EU-level organisations, including Digital Europe and DOT Europe based in Brussels.
- TikTok has become a member of [REDACTED]  
[REDACTED] Similarly, TikTok is a member of [REDACTED]
- In connection with the launch of Shop in the EU, TikTok engaged minor safety and youth wellbeing-focused NGOs [REDACTED]
- TikTok actively engages with other organisations (such as consumer organisations and product safety experts) to further develop the increasing engagement with external stakeholders to maintain a safe and secure Platform.

#### 4. Residual Risk

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual risk of the promotion or sale of Illegal or Unsafe Products to be **Medium**.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 11,792,444 videos under its "Regulated Goods and Commercial Activities" policy; 11,523,078 of those videos were detected and removed proactively (i.e. before any user report), 9,424,572 were removed before being viewed by any user, and 10,998,736 were removed within 24 hours of publication, demonstrating TikTok's swift action to combat Illegal and Unsafe Products on the Platform.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, 269,366 videos were removed under TikTok's "Regulated Goods and Commercial Activities" policy following a user report.

This Risk Assessment was also informed by data on TikTok's enforcement of its TikTok Shop Policies in the EU in Q1 of 2025 (TikTok Shop launched in the EU (in Ireland and Spain) in December 2024). Over the course of Q1 2025, the total number of product listings on TikTok Shop in the EU grew to 1,235,990. In Q1 2025, 18,711 live product listings were deactivated due to a violation of TikTok Shop Policies. Moreover, of the 1,235,990 EU product listings, 35 were reported for allegedly relating to Illegal Content, of which 14 were removed from the Platform on the basis that they violated TikTok Shop Policies. During that same period, of 288,457 active sellers in March 2025, 50 seller accounts were deactivated on the basis of breach of TikTok Shop Policies; and of 12,748 active publishing Creators in March 2025, 483 Creators had their e-commerce permissions revoked.

With respect to the promotion of goods and services outside of Shop, as noted in TikTok's Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok



received 19,753 reports from users of allegedly illegal content related to Illegal goods/services. During the same period, it received 3 reports from Trusted Flaggers related to illegal goods/services. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 7 of which related to Illegal goods/services.

TikTok's policy approach works in combination with its other mitigation measures to mitigate the risk of the promotion or sale of Illegal and Unsafe Products on the Platform.

### 13. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO AGE-INAPPROPRIATE CONTENT

#### 1. Description of the Risk and how it may manifest on the Platform

TikTok is a Platform accessible to minors. TikTok therefore gives careful and appropriate consideration to the risks encompassed by Article 28(1) of the DSA, which requires TikTok to put in place appropriate and proportionate measures to ensure a high level of privacy, safety and security for minors on its service. As explained below, TikTok's minimum age requirement in the EU is 13 years old (the "**Minimum Age Requirement**"). TikTok uses the term "Underage Users" to refer to minors under the age of 13 who may attempt to access the Platform and the term "**Younger Users**" to refer to TikTok users with a declared age of 13 to 17 years old.

TikTok acknowledges that a balance must be found between the design of the measures to address risks to Younger Users and the Fundamental Rights of users of any age. Those rights include the right to freedom of expression, which includes the right of users to express themselves freely on the Platform, receive and impart information and ideas to aid their individual learning and development and to exercise autonomy. Accordingly, TikTok seeks to avoid unnecessary and disproportionate impacts on users in the design of its Platform and when implementing the specific Youth Safety measures summarised in this Report.

#### Summary of risks

There is a risk that Younger Users might access, view or be exposed to content on the Platform that is not age-appropriate ("**Age-inappropriate Content**" and the "**Age-inappropriate Content Risk**"). Related risks of such content may involve negative effects on physical or mental well-being.

TikTok considers that Age-inappropriate Content generally falls into the following categories of harmful content:

- (i) "**Violative Content**": Content that violates TikTok's Community Guidelines, is prohibited for all audiences and is removed from the Platform once identified (in each section of the Community Guidelines, such content is labelled "NOT allowed");

- (ii) **“Age-restricted Content”**: Content that does not violate the Community Guidelines, per se, but that TikTok considers is only appropriate for Adult Users due to its complexity and overt maturity (in TikTok’s Community Guidelines, such content is labelled “RESTRICTED (18 years and older)”);
- (iii) Concentrated Content, as defined above.

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for Age-inappropriate Content and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk Score is consistent with TikTok’s score in Year 2.

## 3. Mitigation Measures

TikTok undertakes the following mitigation measures related to the Age-inappropriate Content Risk:

**Article 35(1)(a)**: TikTok implements a range of measures to prevent and mitigate the Age-inappropriate Content Risk including: (i) blocking certain search results associated with high-risk words from appearing on the Platform; (ii) measures to mitigate the risk of a livestream being capable of containing harmful content; (iii) measures to address the risk of hyperlinks being used to disseminate harmful content; (iv) defaulting the filter for spam and offensive or harmful comments to “on” for all users; (v) Family Pairing features, which allow parents/guardians to manage certain settings relating to the content their Younger User sees on the Platform and the features they can access on the Platform (including by enabling “Restricted Mode” and filtering certain content); (vi) audience control settings, which allows creators to hide their content from Younger Users; and (vii) warning labels and mask layers, which TikTok applies to certain videos to provide users with information and context about a video, so users can decide in advance if they would like to watch it.

**Article 35(1)(b)**: TikTok clearly sets a Minimum Age Requirement, whereby users must be 13 years or older to have an account and to use the Platform, and this is reflected across a range of user-facing documentation and other resources including TikTok’s [Terms of Service](#), [Privacy Policy](#) and the [Community Guidelines](#). TikTok’s Community Guidelines also detail the types of content classed as Age-Restricted Content. For example, TikTok’s “Youth Safety and Well-being” policy sets out a range of content restricted to Adult Users, including content related to disordered eating and body image, dangerous activities and challenges, body exposure and sexually suggestive content, among other things. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for a description of TikTok’s content moderation processes. In addition, TikTok uses content classification measures (Content Levels) to organize content and prevent Age-Restricted Content from reaching Younger Users.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. Of particular importance in the context of Younger Users, TikTok makes certain content that may be unsuitable for a broad audience ineligible for the FYF. Further, TikTok’s dispersion models mitigate the risk that users will be exposed to Concentrated Content.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not violate the Community Guidelines or TikTok’s Advertising Policies, which include a range of measures to protect Younger Users from Age-inappropriate Content. Having regard to Art. 28(2) of the DSA, TikTok has developed systems to prevent personalised ads being served to Younger Users in the EU. Instead, ads are based only on generic data points, for example, country and language. TikTok’s Advertising Policies restrict certain industries (e.g. Healthcare and Pharmaceuticals, Financial Services and Products and Weight Management) from advertising to Younger Users.

**Article 35(1)(f):** TikTok has specialist teams who deal with youth safety and Age-Restricted Content and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures. TikTok conducts regular sampling and reviewing to assess the effectiveness of its mitigation measures in ensuring an age-appropriate experience for Younger Users. This process is key to detecting emerging risks impacting Younger Users.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which contains TikTok’s dedicated [Teen Safety Centre](#), and which has information and resources for Younger Users on the tools available to them on the Platform and how to manage their interactions on the Platform. The Safety Centre also contains TikTok’s Digital Well-being guide and its Guardian’s Guide (which provides information for parents and guardians on how TikTok works and the tools available to them to manage their teen’s use of TikTok); (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; and (iii) TikTok’s other in-app educational and awareness measures, warning labels, verified badges, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** Ensuring that Younger Users have a safe and age-appropriate experience on the Platform continues to be one of TikTok’s highest priorities. In its approach to safeguarding minors in their use of the Platform, TikTok has regard to the rights of the child under the Charter of Fundamental Rights of the European Union and the United Nation Convention on the Rights of the Child including, among other things, the rights of the child to: non-discrimination (including digital inclusion); development and preservation of identity; respect for the views of the child; freedom of expression; protection of privacy; access to

appropriate information; education; and leisure, recreation and cultural activities. TikTok also notes in particular:

- Article 24 of the Charter on the rights of the child states that “[in] all actions relating to children, whether taken by public authorities or private institutions, the child’s best interests must be a primary consideration”;
- Article 13 of the Convention enshrines the right of children to freedom of expression, including “freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of the child’s choice.”; and
- Article 17 of the Convention which protects the rights of children to “access appropriate information,” and specifically encourages: “... the mass media to disseminate information and material of social and cultural benefit to the child and in accordance with the spirit of article 29 [of the Convention]”; “.... International co-operation in the production, exchange and dissemination of such information and material from a diversity of cultural, national and international sources”; and “... the production and dissemination of children’s books.”

The mitigation measures set out in this section are designed to respect, protect and uphold the Rights of the Child on the Platform.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform.

TikTok has further enhanced its mitigation measures related to the Age-inappropriate Content Risk since its Year 2 Report, including:

- Earlier in 2025, TikTok almost doubled the size of its Youth Council from 15 to 28 teens, now from 15 countries, strengthening the diversity of voices on the council. In the past year TikTok has actioned insights from its Youth Council on key projects, including in relation to: (i) the Teen Safety Centre; (ii) the use of appearance effects by Younger Users (including implementing age restrictions and information interventions); and (iii) launching the Digital Safety Partnerships for Families guide, to help families to talk about online experiences and how to set positive digital boundaries.
- TikTok has expanded its Family Pairing features to include new visibility tools, which allow parents/guardians to see who their teen is following on TikTok and who follows them, as well as blocked accounts.
- TikTok recently expanded audience controls to accounts (previously this feature was only available at the content level) so that creators have the option to limit their account and content to viewers at least 18 years old.
- See also mitigation updates referred to in sections 14 and 15.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with TikTok’s Youth Council.
- Engagement with ESAC, which includes leading external subject matter experts on various youth safety and wellbeing issues. This expertise is supplemented by the expertise of TikTok’s US SAC, APAC SAC, Brazil SAC, and Middle East and North Africa SAC.
- Hosting a roundtable on gendered harassment and the impact of misogynistic content on minors with a number of organisations who work to prevent sexual harassment and abuse.
- Ongoing partnerships with leading safety organisations to help ensure that TikTok’s policies and features continue to promote a safe and welcoming environment for its community.

#### 4. Residual Risk

Following its assessment of the effectiveness, reasonableness and proportionality of relevant mitigations (detailed in 3. above), TikTok has assessed the residual Age-inappropriate Content Risk to be **Medium**. This Residual Risk score is consistent with TikTok’s score in Year 2.

This Risk Assessment was informed by data on TikTok’s enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok age-restricted 5,497,995 videos under a number of its policies, making them visible only to Adult Users.

TikTok’s Community Guidelines also prohibit a range of content that is not appropriate for Younger Users from being present on the Platform entirely. For example, 7,720,919 videos were removed under the Community Guidelines and were also deemed unsuitable for Younger Users under the “Youth Safety and Wellbeing Policy.”

TikTok’s policy approach works in combination with its other mitigation measures to mitigate the Age-inappropriate Content Risk. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 aimed at mitigating the Age-inappropriate Content Risk and ensuring that Younger Users have an age-appropriate experience on the Platform. This Risk Assessment has taken account of these measures in determining the effectiveness, reasonableness and proportionality of TikTok’s mitigation measures.

## 14. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO AGE-ASSURANCE

### 1. Description of the Risk and how it may manifest on the Platform

While Younger Users enjoy the right to freedom of expression, access to information and other fundamental rights associated with use of the Platform, TikTok considers it reasonable and proportionate to prohibit minors under the age of 13 from accessing the Platform and to permit Younger Users to access the Platform, while limiting their access to certain features, content and experiences.

**“Age Assurance”** refers to the tools and technologies employed by TikTok to enforce the Minimum Age Requirement and to identify accounts belonging to Younger Users in order to enforce appropriate restrictions on their access to the Platform. TikTok’s approach is consistent with relevant international legal frameworks relating to the protection of children, including the United Nations Convention on Rights of the Child, which recognise the unique vulnerabilities of children and teenagers, as well as their rights to access the digital environment.

#### Summary of risks:

TikTok recognises the following Age Assurance risks:

- Underage Users may mis-state their age at account registration to gain access to the Platform; and
- Individuals between 13 to 17 years of age may mis-state their age at account registration and, as a result, may not receive an age-appropriate experience on the Platform.

(together, the **“Mis-Stated Age Risk”**)

The Mis-Stated Age Risk also includes the risk of inaccurate age determinations at other stages of the Age Assurance process, such as during moderation and appeals. [REDACTED]

## **2. Inherent Risk**

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for Age-Assurance and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium
Overall probability	Very Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk Score is consistent with TikTok’s score in Year 2.

## **3. Mitigation Measures**

TikTok undertakes the following mitigation measures related to Mis-Stated Age Risk:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate the Mis-Stated Age Risk. In order to use the TikTok app, the app must first be downloaded from an app store. The TikTok app is rated 12+ in the Apple App Store and “Parental Guidance Recommended” in the Google Play Store, which empowers parents and guardians to prevent those under the Minimum Age Requirement from downloading the app should they wish to do so using available parental controls on their device. TikTok also uses a neutral

age gate, where the account registration page does not highlight the Minimum Age Requirement (in order not to sway a user's response). If a prospective user enters a birth date indicating they are under 13, they will receive a non-eligibility notification and are prevented from re-attempting to register with an alternative (older) age [REDACTED].

**Article 35(1)(b):** TikTok clearly sets a Minimum Age Requirement, whereby users must be 13 years or older to have an account and to use the Platform, and this is reflected across a range of user-facing documentation and other resources including TikTok's [Terms of Service](#), [Privacy Policy](#) and the [Community Guidelines](#). TikTok's Community Guidelines also detail the types of content classed as Age-restricted Content. TikTok's Terms, Community Guidelines and policies also set out a number of age restrictions on certain features. For example, TikTok's Community Guidelines state that users must be 16 or over for their content to be eligible for the FYF; TikTok's Terms of Service state that users must be 18 years of age or older to host a livestream, use livestream features and send, receive or interact with virtual items; users must be 16 years of age or older to use the direct messaging feature; and TikTok Shop's Terms of Use state that TikTok Shop is for users 18 years of age or older.

**Article 35(1)(c):** Please see section 4 ("About TikTok") for a description of TikTok's content moderation processes. TikTok uses a range of measures to detect if a user has inaccurately stated their age at account registration, including automated moderation, which uses a range of signals to predict a user's age, and reporting channels, where users can report a "suspected underage user," if they believe the user is too young to use TikTok or to use a certain feature (such as LIVE). Once detected, such accounts are sent to a specialist moderation process for review by trained moderators.

TikTok takes a holistic approach to moderating users that mis-state their age and deploys multi-layered strategies to ensure a high level of age assurance. If an account with a potentially mis-stated age is detected, moderators will review it to determine whether they believe the account holder is under 13 and if they determine this is the case, the account will be banned. If a moderator reviews an account and determines that the account holder is under the minimum age to use an age-restricted feature, their access to such feature will be restricted and their account experience will be adjusted.

TikTok operates an appeals process, where users can use proof of age to appeal a determination that they do not meet the Minimum Age Requirement by: (i) providing a selfie with ID; (ii) providing a photo with a trusted adult; (iii) using credit card authorisation (if an Adult User); or (iv) using AI facial age estimation via a third-party service provider. In order to mitigate the risk that AIGC could be used to manipulate the appeals process, the methods include requirements that would be difficult for AIGC to replicate, including secure live image capture or requiring the user to hold up a piece of paper detailing certain information, when taking a live selfie.

**Article 35(1)(f):** TikTok's has specialist teams who deal with Youth Safety and who play a role in policy development, risk prevention and containment and reinforcing TikTok's risk detection measures.



**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(h):** TikTok continues to engage with industry bodies, authorities, regulators and external experts to better understand and address the Mis-Stated Age Risk.

**Article 35(1)(j):** Please refer to the equivalent entry in section 13 for information on TikTok's approach to protecting the Rights of the Child on the Platform. TikTok considers it necessary to strike a balance between the safety and the rights and freedoms of the child.

TikTok also acknowledges that a balance must be found between the design of measures to address the Mis-Stated Age Risk and the fundamental rights of legitimate users of any age to receive information and to freedom of expression, including the right to express themselves freely on the Platform. Accordingly, TikTok has to strike the right balance and avoid unreasonable and disproportionate impacts on such legitimate users (i.e. potentially resulting in preventing them from accessing the Platform, such as by creating unreasonable burdens during sign up or excessive removal of suspected underage accounts in error, and related privacy impacts arising in the context of age assurance strategies, that could function as a deterrent to legitimate users).

TikTok has further enhanced its mitigation measures related to the Mis-Stated Age Risk since its Year 2 Report, including:

- Extending proactive age confirmation measures to users who are hosting a LIVE for the first time in the EU.
- Deploying an improved machine learning model to predict whether a user is within a certain age range.

#### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- ESAC includes leading external subject matter experts on various youth safety and wellbeing issues. This expertise is supplemented by the expertise of TikTok's US SAC, APAC SAC, Brazil SAC, and Middle East and North Africa SAC.
- TikTok helped launch, and supports and serves on the informal steering committee of the "Multi-Stakeholder Dialogue on Age Assurance," working with key strategic partners to foster meaningful dialogue to establish best practices to address current challenges in assessing age. As part of this project, TikTok co-hosted a multi-stakeholder dialogue on age assurance, focused on law and regulation. TikTok also participates in weekly calls concerning the direction of the project, and, over the past year, has participated in several other events.
- TikTok has also engaged with its Youth Council regarding the impact of its strategies to protect, respect and uphold the Rights of the Child on the Platform.

#### **4. Residual Risk**

Following an assessment of the effectiveness, reasonableness, and proportionality of TikTok's mitigations relevant to the systemic risk related to Age Assurance (detailed in 3.

above), TikTok has assessed residual risk to be **Medium**. This Residual Risk Score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok detected and removed 3,562,215 suspected underage accounts from the Platform, 2,520,752 of which were removed proactively (i.e. before any user report).

The rate of proactive removals is a positive indicator of the efficacy of TikTok's moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners, through its Community Partner Channel. When TikTok receives such reports, it diligently acts to investigate and remove suspected underage accounts. From Q3 2024 through Q1 2025, TikTok removed 1,041,463 accounts from the Platform following a user report.

TikTok is developing new ways to mitigate this risk as part of TikTok's efforts to ensure that Underage Users are not able to have an account on the Platform.

TikTok continues to engage with a range of external experts on its Age Assurance measures and, as set out above, has implemented a number of measures to improve its age assurance and detection measures since its Year 2 report. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 15. YOUTH SAFETY AND ONLINE ENGAGEMENT: RISKS RELATED TO ONLINE ENGAGEMENT

### 1. Description of the Risk and how it may manifest on the Platform

TikTok gives careful consideration to risks to users arising from online engagement, including, as required by Article 34(1)(d) of the DSA, risks of serious negative consequences to users' physical and mental wellbeing in connection with their use of the Platform (the "**Online Engagement Risks**"). While TikTok does not believe that use of the Platform is inherently harmful, it recognises that absent mitigations in place some users could at times and in specific circumstances face negative mental and physical health impacts from engagement with content on online platforms.

With respect to Younger Users, TikTok understands these risks may arise from their creation and publication of content, Younger Users' interactions with other users or their content, and other potential impacts on Younger Users' wellbeing arising from their use of the Platform. There are also risks to the wellbeing of Younger Users arising from potential exposure to violative, illegal or Age-inappropriate Content. Those risks are dealt with in sections 7-14 and 16-18 of this Report.

With respect to Adult Users, wellbeing risks may arise from the way some Adult Users use the Platform. There are also risks to the wellbeing of Adult Users arising from potential exposure to violative or illegal content. Those risks are dealt with in sections 7-12 and 16-18 of this Report.

### Summary of risks

Given the absence of consensus among health professionals and academics on how social media use may adversely affect user health, TikTok engages experts and monitors associated risks. TikTok has considered Online Engagement Risks that could arise in the context of online platforms, which may include:

- Younger Users creating and publishing content without fully comprehending or anticipating the potential consequences of doing so (including sharing sensitive personal information or content, illegal content, content that violates TikTok's Community Guidelines or inappropriate AIGC), which could lead to a loss of privacy, embarrassment, or distress.
- Younger Users proactively engaging with content posted by other users, including through comments or by using the Stitch or Duet tools (for users with a declared age of over 16), in an inappropriate manner.
- Other users interacting with Younger Users or their content in an inappropriate manner, which could cause embarrassment or distress to Younger Users or lead to more serious forms of harm. For example, the Platform could potentially be used for bullying behaviour or bad actors, such as child predators, groomers, cyberbullies, or scammers may attempt to use the Platform to discover or interact with Younger Users.
- Impacts on the mental wellbeing of Younger Users, which could, depending on the specific circumstances and/or choices of the individual, involve: (i) extended use of the Platform; (ii) exposure to Concentrated Content, which may be fine when viewed occasionally, but may be problematic if viewed repeatedly; (iii) social comparison; or (iv) risks arising from engagement with dangerous online trends or challenges (as well as related risks of physical harm).
- Impacts on the mental wellbeing of Adult Users, which could, depending on the specific circumstances and/or choices of the individual, involve for example: (i) extended use of the Platform; (ii) exposure to certain types of Concentrated Content that, while not violative of TikTok's Community Guidelines, may cause harm to some users by inadvertently reinforcing a negative personal experience; or (iii) social comparison, which may arise, for example, through the use of filters, beauty effects, or AIGC that promote unrealistic standards of appearance.

## **2. Inherent Risk**

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for Online Engagement Risks and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk Score is consistent with TikTok's score in Year 2.

### 3. Mitigation Measures

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate Online Engagement Risks including: (i) screentime management tools (including the screentime management dashboard) and digital wellbeing tools (such as digital wellbeing prompts and the wind down feature) to help users manage their time spent on the Platform; (ii) restrictions on push notifications for Younger Users; (iii) Family Pairing features which allow parents/guardians to manage certain settings relating to the content their Younger User sees on the Platform, the time they spend on the Platform and the features they can access on the Platform; (iv) account and content-level privacy settings to give users control over how others can view their profile, content and following lists and interact with them and their content, including pro-privacy default settings for Younger Users; (v) measures, including age restrictions, to mitigate the risk that LIVE, Duet and Stitch could be used to share harmful information or engage in harmful conduct that may affect user wellbeing; and (vi) measures to mitigate the risk that direct messaging could be used to share harmful content or conduct harmful interactions, including restriction on accessing the feature for certain Younger Users.

**Article 35(1)(b):** TikTok mitigates Online Engagement Risks through a combination of TikTok's [Terms of Service](#) and the [Community Guidelines](#). TikTok's Community Guidelines contain dedicated policies on "Mental and Behavioural Health" and "Youth Safety and Well-Being," as well as other policies that are designed to protect the well-being of users. For example, TikTok's "Mental and Behavioural Health" policy prohibits or age-restricts a range of content related to suicide and self-harm, disordered eating and body image and dangerous activities and challenges. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 ("About TikTok") for a description of TikTok's content moderation processes. In addition, TikTok uses a range of systems to identify and limit potentially dangerous trends and challenges on the Platform.

**Article 35(1)(d):** Please see section 6 for a description of TikTok's controls related to recommender systems. TikTok uses dispersion techniques to reduce the risks associated with Concentrated Content, by dispersing content containing certain themes, including borderline extreme dieting and fitness, borderline negative affect, adult nudity and sexual activity and mental health narratives, which may not be problematic when viewed

occasionally, but may become so if viewed repeatedly. TikTok also makes certain content that may be unsuitable for a broad audience ineligible for the FYF.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not violate the Community Guidelines or TikTok’s Advertising Policies. Having regard to Art. 28(2) of the DSA, TikTok has developed systems to prevent personalised ads being served to Younger Users in the EU. Instead, ads are based only on generic data points, for example, country and language.

**Article 35(1)(f):** TikTok has specialist teams who deal with youth safety and wellbeing and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(h):** TikTok has a global Youth Council, through which it actively engages young people to hear their insights and understand how TikTok can better respond to their safety and wellbeing needs on the Platform. TikTok is also involved in the industry-wide engagement on Online Engagement Risks, and engages with industry bodies, NGOs and external experts on Online Engagement Risks.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which contains TikTok’s dedicated [Teen Safety Centre](#), TikTok’s [Digital well-being guide](#), and TikTok’s [Safety Centre portal on Online Challenges](#) as well as resources on a range of topics including bullying, tragic events, gambling, scams, eating disorders and online challenges; and (ii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content.

**Article 35(1)(j):** The measures which mitigate the Online Engagement Risks address the rights of the child. Additional Youth Safety measures are set out in sections 13 and 14 of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. TikTok prohibits a range of edited or AIGC content that is misleading or that contains certain depictions of minors, which mitigates the risk of [REDACTED]. Restrictions on the use of AIGC also lower the risks of potentially negative effects from a proliferation of content depicting unrealistic beauty standards. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities.

TikTok has further enhanced its mitigation measures related to Online Engagement Risks since its Year 2 Report, including:

- TikTok introduced a new in-app guided meditation feature to encourage users to switch off at night (the “Wind Down feature”). This feature is available to all users

and is turned on by default for Younger Users after 10pm. If the user is on TikTok after 10pm, their FYF will be interrupted and a full-screen takeover is initiated, with calming music to help users relax and be mindful of the time. If the user decides to spend additional time on TikTok after the first reminder, TikTok shows a second full-screen prompt. TikTok designed these features to reflect best practices in behavioural change theory by providing positive nudges that can help users develop balanced long-term habits.

- TikTok added a new feature to Family Pairing, the “Time Away” feature, which allows parents or guardians to block their teens from using the TikTok app at a certain time. Parents/guardians can also set a recurring schedule to best suit their family life. If plans change, teens can request extra time, but parents/guardians make the final decision on whether to allow additional time on the Platform. TikTok has also added new visibility tools, which allow parents/guardians to see who their teen is following on TikTok and who follows them, as well as blocked accounts.
- TikTok recently updated its Teen Safety Centre, which provides teens with information about the different features and safety tools available on TikTok.
- TikTok also updated: (i) its Well-being Guide, which includes tips on how teens can assess their digital well-being and how they can use TikTok’s tools to support this; and (ii) its Guardian’s Guide, which provides parents and guardians with information and resources to help them understand how TikTok works and the tools TikTok provides to help them take an active role in ensuring their child’s online safety and help them manage their child’s content, privacy, and well-being settings on the Platform.
- TikTok introduced a policy on Harmful Body Idealization, which covers content that promotes an idealized body aesthetic when associated with potentially harmful weight management behaviours, such as disordered eating and anabolic steroid use.
- TikTok launched its “beauty effects strategy” following consultations with its Youth Council and in response to research carried out by Internet Matters with teens and parents, supported by TikTok. In the initial phase of this strategy, TikTok developed a risk framework related to beauty effects on the Platform and has used that framework to implement age restrictions on different beauty filters and effects, according to their risk level.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with TikTok’s Youth Council on a number of issues, including the use of beauty effects and filters by Younger Users on the Platform, the Teen Safety Centre, AIGC, parental controls, screentime management and launching the Digital Safety Partnerships for Families guide.
- Engagement in the Better Internet for Kids Focus Group on Cyberbullying, as part of a set of targeted consultations to inform the drafting of the European Commission’s Action Plan against cyberbullying.
- Engagement with experts on research projects focused on improving practices to keep children safe from technology-assisted sexual exploitation and abuse, including online grooming.



- Policy consultations with experts from victim support organisations in relation to detecting predators.
- Engagement with ESAC, which includes leading external subject matter experts on various youth safety and wellbeing issues. This expertise is supplemented by the expertise of TikTok's US SAC, APAC SAC, Brazil SAC, and Middle East and North Africa SAC.
- Membership in various support and advocacy organisations aimed at protecting children from online abuse and bullying.
- Consultations with experts in relation to Platform design, including TikTok's screentime management tools.

#### 4. Residual Risk

Following an assessment of the effectiveness, reasonableness, and proportionality of TikTok's mitigations relevant to the systemic risk related to Online Engagement (detailed in 3. above), TikTok has assessed residual risk to be **Medium** in Year 3. This Residual Risk score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 274,088 videos under its "Mental and Behavioral Health - Suicide and Self-Harm" policy; 253,322 of those videos were detected and removed proactively (i.e. before any user report), 201,136 were removed before being viewed by any user and 241,865 were removed within 24 hours of publication. TikTok also removed 6,449,646 videos under its "Mental and Behavioral Health - Dangerous Activity and Challenges" policy; 6,392,321 of those videos were detected and removed proactively: 5,669,569 were removed before being viewed by any user and 5,897,024 were removed within 24 hours of publication. This demonstrates TikTok's swift action to combat violative content that could potentially impact user wellbeing on the Platform.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, following a user report, TikTok removed 20,766 videos under its "Mental and Behavioral Health - Suicide and Self-Harm" policy and 57,325 videos under its "Mental and Behavioral Health - Dangerous Activity and Challenges" policy.

A number of TikTok's other policies capture content that impacts the Online Engagement Risks, including TikTok's "Youth Safety and Well-Being" policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the Online Engagement Risk on the Platform.

TikTok remains engaged in industry-wide conversations about Online Engagement Risks and is continuously refining its approach to addressing these risks. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to promote wellbeing on



the Platform, protect users from harmful content, safeguard Younger Users and equip users with information to help them have a positive online experience. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 16. CIVIC INTEGRITY AND MISINFORMATION: RISKS TO ELECTIONS AND CIVIC INTEGRITY

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the risk to be the actual and foreseeable negative effects on election processes and on civic integrity arising from the dissemination of verifiably false or misleading content related to an election or other civic events in the EU (such as referenda or a census) (together, "**Election Misinformation**").

Articles 39 and 40 of the Charter enshrine the rights of EU citizens to vote and to stand as a candidate at elections to the European Parliament and at municipal elections, respectively, and such fundamental rights may be undermined by Election Misinformation.

#### Summary of risks

Election Misinformation risks may arise from attempts to share or disseminate content on or through the Platform, whether as photo, short video, comment, livestream, comments, or profile information, including in the following ways:

- Misleading information about how to vote or register to vote or the qualifications to vote or run in an election;
- Misleading information about the date of an election or other civic process (e.g. stating that an election is on a later date than it is scheduled for);
- Misleading information about how to participate in a census or eligibility requirements for participating in a census;
- Content that advances false claims related to the technical eligibility requirements for current political candidates and sitting elected government officials to serve in office or run in an election;
- Content that falsely makes claims of election fraud (such as voting machines being tampered with to favour a candidate or political party) or that that an election has been or will be rigged, with the implication that the results cannot be trusted;
- Content that makes false claims about (including denial of) legitimate election outcomes;
- Content that promotes or provides instruction on electoral interference, including voter intimidation, or calls for disrupting legitimate election outcomes through violence;
- Misinformation or/and conspiracy theories about candidates, candidate impersonation and related issues that may impact civic integrity;
- False claims misrepresenting authoritative sources of civic information (e.g. the text of a parliamentary bill);

- Content which otherwise undermines users' trust in civil processes and procedures through unverified claims;
- Undisclosed political branded content, where an individual may have been paid or otherwise incentivised to promote an election related message, but does not disclose this as required by TikTok's terms; and
- Misleading AIGC or edited media featuring public figures (such as a government official, politician, business leader or celebrity) may impact civic integrity if mis-used for political endorsements or other purposes. There are also risks associated with edited media and AIGC that falsely appear to come from a reputable news organisation.

These risks could be amplified by inauthentic use of, and engagement on, the Platform, including through covert influence operations, "hack-and-leak" operations, commercial disinformation (including disinformation as-a-service sold by third-party vendors or misleading marketing campaigns presented as UGC), violative use of the gifting feature or by fake engagement and spam activity that attempts to disrupt election processes or negatively affect civic discourse. Geopolitical factors can also intensify these risks, including political divisions during election cycles.

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for elections and civic integrity risks, and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk score is consistent with TikTok's score in Year 2.

## 3. Mitigation Measures

TikTok implements the following mitigation measures related to elections and civic integrity risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate election and civic integrity risks on or through the Platform, including: (i) blocking certain search results associated with Election Misinformation from appearing on the Platform; (ii) search interventions, including search guides to redirect users to reliable information when searching for issues that may be susceptible to misinformation risks; (iii) measures to mitigate the risk of a livestream being capable of containing Election Misinformation; (iv) election labels and search banners to link users to authoritative information on elections, as well as warning labels, to indicate where the accuracy of information cannot be verified; (v) displaying the verified badge alongside verified accounts, which includes government, politician and political party ("GPPPA") accounts; (vi) measures to address the risk that

hyperlinks shared on the Platform can be used to spread Election Misinformation; (vii) measures to address the risk of harmful misinformation in comments or commentary on TikTok, including users controls over the comments and Duet and Stitch features; and (viii) the launch of in-app Election Centres for specific elections, to connect users with reliable information about the elections and voting.

**Article 35(1)(b):** TikTok prohibits Election Misinformation and prohibits or restricts a range of related content and behaviour on the Platform through a combination of TikTok's Terms of Service, Community Guidelines and policies. For example, under its "Integrity and Authenticity - Civic and Election Integrity" policy, TikTok prohibits election misinformation, including on how, when and where to vote, eligibility requirements of voters, laws and processes that govern elections and civic processes and election results. In addition TikTok's policies, including its Branded Content Policy, require users to disclose when they promote a brand, product or service in exchange for payment or other incentive. TikTok takes a range of measures to generate awareness and inform users about its terms and policies. TikTok also performs scenario-planning prior to elections to inform its policy approach.

**Article 35(1)(c):** Please see section 4 ("About TikTok") for descriptions of TikTok's content moderation processes and fact-checking programme. In addition, TikTok brings together cross-functional teams ahead of certain elections to respond to incidents, manage escalations and monitor real-time trends. Further, TikTok operates an Elections Integrity programme, part of which involves the detection and assessment of Election Misinformation content and trends proximate to election cycles, which includes real time monitoring of a range of signals and sampling content to understand risks and trends. In the context of a number of EU elections, TikTok engaged with external parties, including civil society organisations and fact-checkers, as part of the Code of Conduct on Disinformation Rapid Response System.

**Article 35(1)(d):** Please see section 6 for a description of TikTok's controls related to recommender systems. Further, following review by TikTok's fact-checking teams, unverified content may be designated as ineligible for recommendation in the FYF.

**Article 35(1)(e):** Please see section 4 ("About TikTok") for a description of TikTok's advertising systems and policies. TikTok prohibits paid political advertising including when using promotional tools on the Platform such as [Promote](#), and political branded [content](#). TikTok turns off the Promote advertising function as default for accounts labelled as GPPAs (except for governments which have a limited exception to advertise).

**Article 35(1)(f):** TikTok has specialist teams who deal with harmful content, including Election Misinformation, and who play a key role in policy development, risk prevention and containment and reinforcing TikTok's risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(h):** TikTok has continued its active participation as a signatory to the EU Code of Practice on Disinformation, contributing to the Permanent Taskforce and subgroups. In

July 2025, the Code of Practice on Disinformation was integrated into the framework of the DSA as a code of conduct. TikTok also continues to engage with various industry bodies, NGOs, external experts and government and regulatory authorities on risks related to elections and civic integrity.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which includes a Harmful Misinformation Guide; (ii) TikTok’s Transparency Centre, which has dedicated content explaining TikTok’s approach to content moderation and keeping people safe and information about the [covert influence operations](#) TikTok has disrupted on the Platform; (iii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; (iv) dedicated in-app Election Centres in the run up to key elections to provide users with reliable information on voting and other resources; and (iv) TikTok’s other in-app educational and awareness measures, verified badges, state-controlled media labels, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** TikTok’s measures to mitigate the risks related to Elections and Civic Integrity on the Platform are designed to protect users of all ages. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. In particular, TikTok’s Edited Media and AIGC policy prohibits misleading AIGC that falsely shows public figures in sensitive contexts, including taking a position on a political issue or a matter of public importance (such as an election) or being politically endorsed or condemned by an individual or group. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.

TikTok has further enhanced its mitigation measures related to Elections and Civic Integrity since its Year 2 Report, including:

- Launching the Global Elections Integrity Hub in TikTok’s Transparency Centre as part of TikTok’s ongoing commitment to safeguarding electoral processes and promoting trust in civic discourse on TikTok. This dedicated resource provides deeper insight into how TikTok approaches election-related content globally, including TikTok’s approach to misinformation, advertising, and enforcement.
- Launching dedicated Election Centres for elections across Europe including Germany, Poland, Portugal and Romania.
- Enhancing TikTok’s fact-checking programme to enable better insights and enforcement.
- Increasing transparency efforts around inauthentic behaviour and impersonation in the elections context. For example, TikTok dynamically disclosed metrics on its efforts to remove inauthentic behaviour related to the Romanian elections through its elections integrity newsroom and provided similar updates on various elections in the Global Elections Integrity Hub.
- Updating TikTok’s political advertising policy which prohibits paid political advertising and political branded content.

### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- Partnerships with more than 20 IFCN-accredited fact-checking organizations to assess the accuracy of content.
- Engagement with reputable fact-checking organisations and local media outlets ahead of elections in Europe as part of TikTok's efforts to protect the integrity of the Platform during elections and provide reliable information in TikTok's Election Centres.
- As part of the Elections Integrity Programme, TikTok's teams undertake various initiatives, including an External Election Speaker Series in the lead-up to elections with suitably qualified external local/regional experts who share their insights and market expertise with TikTok's internal cross-functional teams on key issues involved in and impacting on the election, and potential risks that may impact or manifest on the Platform.
- Participation in RightsCon, a major global summit on digital human rights. TikTok's teams attended sessions covering content moderation, misinformation and disinformation.
- Participation in the United Nations "AI for Good Global Summit." TikTok met with established partners to discuss the C2PA initiative, connected with potential new partners and engaged with several AI creators. As an official industry partner for the summit, TikTok also hosted a panel discussion.
- Engagement with relevant national electoral authorities and/or Digital Services Coordinators in the run up to and during elections.

### **4. Residual Risk**

Following an assessment of the effectiveness, reasonableness and proportionality of TikTok's mitigations relevant to the systemic risk to elections and civic integrity (detailed in 3. above), TikTok has assessed residual risk to be **Medium** in Year 3. This Residual Risk score is consistent with TikTok's score in Year 2.

The last 12 months have seen a number of elections in the EU, including at member state level. In response to new and evolving threats and actions by bad actors to misuse online platforms to disrupt elections and civic integrity in the EU, TikTok has continued to strengthen its mitigations. This includes active risk monitoring during election cycles, working with fact-checking partners and media literacy campaigns.

During elections in Europe, TikTok protects the integrity of its Platform by:

- [Preventing harmful content;](#)
- [Connecting people to reliable facts;](#)
- [Combating deceptive behaviour and defending platform integrity;](#)
- [Partnering with experts;](#) and
- [Preparing for rapidly unfolding events.](#)

TikTok's measures to safeguard the integrity of the Platform include TikTok's Elections Taskforce, who work diligently to protect TikTok's community and make TikTok a safe place for TikTok's users to express their views and reach reliable information, and TikTok's Fact-Checking Programme (described in section 4 of this Report), which plays an important role in protecting the Platform from election misinformation.

In addition, in January 2025 TikTok launched the [Global Elections Integrity Hub](#) ("**Elections Hub**") to keep users informed of TikTok's efforts and to provide its community with regular updates on TikTok's ongoing work to protect TikTok and its users during elections. In the Elections Hub, TikTok provides links to its policies on Civic and Election Integrity, Misinformation, Hate Speech and Hateful Behavior, AI-generated content, News and GPPAs, Public interest content exceptions, Political advertising and paid political promotion and Violent and Criminal Behavior and reports on the volume of content it has removed from the Platform under those policies. TikTok also provides updates on its efforts to monitor and take action against inauthentic and harmful content in the context of specific elections, including details of covert influence networks it has disrupted during election periods, as well as the content, accounts and inauthentic engagement it has prevented or removed in connection with those elections. The Elections Hub also contains information on TikTok's Election Centres, which are described below.

TikTok launches dedicated in-app **Election Centres** for specific elections, where its users can find reliable facts about elections and voting. The Election Centres contain links to authoritative election information from trusted sources and resources developed in connection with local fact-checking partners. TikTok directs users to the Centres via labels on election-related content or searches in-app. This year, TikTok launched a number of Election Centres for elections in Europe.

In addition, TikTok launched media-literacy campaigns in relation to the following: 1. Croatia Presidential Election; 2. German Federal Election; 3. Greenland General Election; 4. Finland local and municipal elections; 5. Romania Presidential Elections; 6. Albania General Elections; 7. Croatia Local Elections; 8. Portugal Legislative Elections; 9. Poland Presidential Elections and; 10. Latvia Local and Municipal elections.

TikTok has also recently (July 2025) updated its [Politics, Governments, and Elections](#) Advertising Policy to improve transparency and clarity on the rules related to government and elections advertising, as well as to better support election integrity, while maintaining its core prohibition on political advertising. As stated in the Policy, TikTok does not allow political content to feature in advertising or in branded content (e.g. a user promoting content for a third party in exchange for payment or any other incentive, where TikTok is not remunerated to promote that content, "Branded Content"). In addition, to assist in enforcing its Advertising Policies, TikTok classifies certain political accounts as GPPAs and applies restrictions to such accounts, including disabling their access to advertising features such as Promote (except for governments who can advertise in limited circumstances).

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 61,340 videos under its "Integrity and Authenticity - Civic and Election Integrity" policy;



59,015 of those videos were detected and removed proactively (i.e. before any user report), 49,349 were removed before being viewed by any user, and 54,428 were removed within 24 hours from publication, demonstrating TikTok's swift action to combat content that threatens civic and election integrity

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, following a user report, TikTok removed 2,325 videos under its "Integrity and Authenticity - Civic and Election Integrity" policy.

TikTok regularly reports on its efforts to detect and disrupt Covert Influence Operations on the Platform through its Transparency Centre. Globally, TikTok removed 55 separate Covert influence Operations from July 2024 to March 2025, resulting in more than 81 thousand accounts being banned.

A number of TikTok's other policies capture content that may threaten elections and civic integrity, including its "Integrity and Authenticity - Misinformation" policy and its "Integrity and Authenticity – Edited Media and AI-Generated Content (AIGC)" policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the risks to elections and civic integrity on the Platform.

In the 2025 Global Elections Hub TikTok posts information regarding its work with local electoral authorities and fact-checking partners, and its efforts to detect and address inauthentic behaviour and remove content that violates its Community Guidelines, and data regarding engagement with elections Centres.

External events can have a bearing on the type and volume of content that threatens civic and election integrity on the Platform. As such, TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve transparency and provide users with accurate information. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 17. CIVIC INTEGRITY AND MISINFORMATION: HARMFUL MISINFORMATION

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands the risk to be the risk of actual and foreseeable negative effects arising from the dissemination of harmful misinformation, which includes:

- verifiably false medical advice which may cause imminent, negative public health effects ("**Medical Misinformation**"). Verifiably false medical advice is advice which



has been disproved by recognized medical authorities, such as the World Health Organization or an EU member state's national health service or Ministry of Health; and

- harmful misinformation and other harmful and/or illegal content that may pose **"Public Security Risks,"** i.e. risks which may relate to: armed conflicts and emerging conflicts; acts of terrorism; natural and manmade disasters (such as floods, earthquakes, hurricanes, fires, landslides, and environmental or industrial accidents); and other emergency situations that may induce panic, including in relation to current/unfolding events, such as civil unrest (including protests or riots).

TikTok notes that Article 35 of the Charter enshrines for everyone the right of access to preventive health care and a high level of human health protection as part of EU policies and activities, which may be undermined by Medical Misinformation.

TikTok notes that Article 6 of the Charter enshrines for everyone the right to liberty and security of person, and Article 12 of the Charter enshrines the right to freedom of peaceful assembly and to freedom of association at all levels (in particular in political, trade union and civic matters), and that protection of these rights in particular could be restricted or jeopardised in the context of Public Security Risks.

TikTok also recognises that individuals retain other rights and freedoms under the Charter, including freedom of expression and information, and TikTok's freedom to conduct a business, which must be balanced in a proportionate manner in the context of addressing Public Security Risks.

#### Summary of risks

These risks may arise from attempts to share or disseminate the following content on or through the Platform, whether as photo, short video, comment, livestream, comments, or profile information:

Medical Misinformation, including:

- Misinformation claims related to medical or health conditions that could potentially result in serious harm. For example, misinformation claims relating to cancer, heart disease, HIV/AIDS, COVID-19, Ebola, abortion, birth control, pregnancy, serious disabilities, birth defects, strokes, tuberculosis, Zika and certain chronic illnesses such as Alzheimer's or diabetes, and other similar conditions;
- Misinformation or misleading content around the treatment or prevention of injuries, conditions or illnesses that are not immediate or life-threatening;
- Misinformation promoting or encouraging the use of alternative prevention or treatment methods which are medically unsupported or unproven or which could dissuade individuals from seeking appropriate treatment or medical care;
- Inaccurate medical or health advice that may cause imminent harm, such as discouraging people from getting appropriate medical care for a life threatening disease or other Medical Misinformation regarding holistic/homoeopathic remedies (e.g. statements that drinking or inhaling cleaning/corrosive substances could act to prevent or treat a disease or that drinking or eating a herbal remedy can treat cancer

or another life threatening illnesses) that could themselves cause serious injury or illness;

- Content which misrepresents authoritative sources, such as selectively citing scientific data to support misleading conclusions;
- Medical Misinformation regarding vaccines; and
- Other misinformation that poses a threat to public health.

While not all Medical Misinformation leads to immediate or severe public health consequences, some claims may still pose moderate risks, requiring a proportionate response.

Harmful misinformation posing Public Security Risks, including:

- Misinformation making verifiably false and harmful claims regarding natural and manmade disasters (such as floods, earthquakes, hurricanes, fires, landslides, environmental or industrial accidents);
- Misinformation making verifiably false and harmful claims regarding unfolding shooting events or mass murders;
- Misinformation making verifiably false and harmful claims regarding public demonstrations or protests;
- Repurposing old video content, making verifiably false and harmful claims that the event or video is new/current and likely to trigger societal panic (e.g. misleadingly repurposing footage of a bombing or armed attack out of context);
- Misinformation making verifiably false and harmful claims that basic necessities (e.g. food, water.) or services (e.g. banks, cash machines) are no longer available in a particular location, causing hoarding;
- Stating dangerous conspiracy theories that are violent or hateful, such as making a violent call to action, having links to previous violence, denying well-documented violent events, or causing prejudice towards a group with a protected attribute;
- Incitement to violence or criminal acts, such as property damage; and
- Climate change misinformation, i.e. misinformation which denies the existence of climate change or the human activities contributing to it, disputes the scientific consensus on climate change or spreads false claims about climate change solutions.

TikTok recognises that these risks may be heightened in the context of rapidly evolving events, as information may evolve and become outdated quickly, even where content was not originally misinformation when posted.

Misinformation can also arise as a result of, or be influenced by, AIGC. AIGC can sow confusion, particularly if users do not understand that the content is either completely generated or significantly edited by AI and contains realistic appearing scenes or people. TikTok is closely monitoring misinformation risks associated with AIGC, including in the context of high-profile events, emergencies and crises.

## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for harmful misinformation, and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Possible
<b>Inherent risk</b>	<b>Medium</b>

This Inherent Risk score is consistent with TikTok's score in Year 2 related to harmful misinformation posing Public Security Risks (the Inherent Risk score for Medical Misinformation in Year 2 was low).

## 3. Mitigation Measures

TikTok undertakes the following mitigation measures related to Medical Misinformation and harmful misinformation posing Public Security risks:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate Medical Misinformation and misinformation posing Public Security Risks on or through the Platform, including: (i) blocking certain search results associated with harmful misinformation from appearing on the Platform; (ii) search interventions, including search guides to redirect users to reliable information when searching for issues that may be susceptible to misinformation risks; (iii) measures to mitigate the risk of a livestream being capable of containing harmful misinformation; (iv) content labels and prompts to provide users with additional context about certain content, highlight that certain content is unverified, or to label certain accounts as verified or linked to state-controlled media; (v) measures to address the risk that hyperlinks shared on the Platform can be used to spread harmful misinformation; and (vi) measures to address the risk of harmful misinformation in comments or commentary on TikTok, including user controls over Comment, Duet and Stitch features.

**Article 35(1)(b):** TikTok prohibits a range of harmful misinformation, including misinformation that poses a risk to public safety and health misinformation, on the Platform through a combination of TikTok's Terms of Service and the Community Guidelines. For example, under its "Integrity and Authenticity - Misinformation" policy, TikTok prohibits misinformation that poses a risk to public safety or may induce panic about a crisis event or emergency, health misinformation, climate change misinformation and conspiracy theories that name and attack individual people or that are violent or hateful. TikTok takes a range of measures to generate awareness and inform users about its terms and policies. To address the unique risks posed by harmful misinformation during crises, TikTok's policy response is grounded in continuous assessment of how crises evolve, how misinformation and conspiracy theories emerge in real time, and the potential harm they can cause both on and off the Platform.

**Article 35(1)(c):** Please see section 4 (“About TikTok”) for descriptions of TikTok’s content moderation processes and TikTok’s fact-checking programme. TikTok maintains a repository of fact-checked information to help its moderation teams accurately and efficiently assess the veracity of suspected harmful misinformation. TikTok also collaborates with independent fact-checking partners and climate experts to stay aligned with the most recent scientific consensus to identify misinformation trends in relation to climate change.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems. Further, following review by TikTok’s fact-checking teams, unverified content may be designated as ineligible for recommendation in the FYF.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not contain inaccurate, misleading or false content that may cause significant harm to individuals or society, including dangerous misinformation, medical misinformation or content that undermines public health or safety or otherwise violates TikTok’s Community Guidelines or Advertising Policies.

**Article 35(1)(f):** TikTok has specialist teams who deal with harmful content, including harmful misinformation, and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(h):** TikTok has continued its active participation as a signatory to the EU Code of Practice on Disinformation, engaging in the Permanent Taskforce and subgroups, and contributing to the development of shared standards for tackling disinformation. In July 2025, the Code of Practice on Disinformation was integrated into the framework of the DSA as a code of conduct.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which includes a [Harmful Misinformation Guide](#); (ii) TikTok’s Transparency Centre, which has dedicated content explaining TikTok’s approach to content moderation, combating harmful misinformation and countering deceptive behaviour; (iii) TikTok’s Help Centre, containing “how to” explanations to allow users to learn about its content moderation practices and how to report violative content; and (iv) TikTok’s other in-app educational and awareness measures, verified badges, state-controlled media labels, unverified content labels and AI-generated content labels. In addition, TikTok partners with fact-checkers, expert organisations and public health authorities to deliver media literacy campaigns including on topics such as climate misinformation, health disinformation and conflict-related narratives. TikTok created a digital literacy hub in its app with content aimed at equipping people with skills to evaluate content they consume.

**Article 35(1)(j):** TikTok’s measures to mitigate the risks of harmful misinformation on the Platform are designed to protect users of all ages. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC. In particular, TikTok’s Edited Media and AIGC policy prohibits content that shares or shows fake authoritative sources or crisis events or falsely shows public figures in certain contexts. TikTok also takes a range of measures to prevent and mitigate the risk that bad actors may intentionally make inauthentic use of the Platform, including by crafting inauthentic identities and by conducting covert influence operations.

TikTok has further enhanced its mitigation measures related to harmful misinformation since its Year 2 Report, including:

- Introducing new misinformation policies, including granular ad policies for medical misinformation, dangerous conspiracy theories, and manipulated media.
- Launching new search guides for several natural disasters, including climate change and floods and wildfires in Europe, which direct users to authoritative sources of information.
- Renewing its partnership with UN Verified Climate Champions, a trusted messenger program to counter climate misinformation and disinformation. TikTok has worked with content creators to raise awareness about climate change and drive realistic action.

#### Relevant stakeholder engagement

TikTok’s stakeholder engagement since Year 2 that is relevant to this risk includes:

- Partnerships with more than 20 IFCN-accredited fact-checking organizations to assess the accuracy of content.
- Engagement with TikTok’s Safety Advisory Councils, which are an important source of expert advice and outside perspectives across its safety work.
- Regular attendance at conferences and events on misinformation issues, which in the last 12 months included the GlobalFact12 summit, the EU DisinfoLab’s annual conference, the Cambridge Disinformation Summit, and the Content Authenticity Initiative Summit.
- Participation in RightsCon, a major global summit on digital human rights. TikTok’s teams attended sessions covering content moderation, misinformation and disinformation.
- Participation in the United Nations “AI for Good Global Summit.” TikTok met with established partners to discuss the C2PA initiative, connected with potential new partners and engaged with several AI creators. As an official industry partner for the summit, TikTok also hosted a panel discussion.
- Partnership with Verified Climate Champions, a joint initiative of the United Nations and Purpose, which targets climate misinformation and disinformation.

#### 4. Residual Risk

Following an assessment of the effectiveness, reasonableness and proportionality of TikTok's mitigations relevant to the systemic risk of harmful misinformation (detailed in 3. above), TikTok has assessed residual risk to be **Medium** in Year 3. This Residual Risk score is consistent with TikTok's score in Year 2 for harmful misinformation posing Public Security Risks (the Residual Risk score for Medical Misinformation in Year 2 was low).

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. During that time, TikTok removed 337,952 videos under its "Integrity and Authenticity - Misinformation" policy; 330,334 of those videos were detected and removed proactively (i.e. before any user report), 263,391 were removed before being viewed by any user, and 311,624 were removed within 24 hours from publication, demonstrating TikTok's swift action to combat content that threatens civic integrity.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, following an user report, TikTok removed 7,618 videos under its "Integrity and Authenticity - Misinformation" policy.

A number of TikTok's other policies capture content that may pose harmful misinformation risks, including its "Integrity and Authenticity – Edited Media and AI-Generated Content (AIGC)" policy. TikTok's policy approach works in combination with its other mitigation measures to mitigate the risks to civic integrity on the Platform.

TikTok also uses "notice tags" to provide users with reliable information in connection with topics at risk for misinformation. For example, from Q3 2024 through Q1 2025, TikTok's video notice tags for M-Pox received 430,628,865 views.

External events can have a bearing on the type and volume of harmful misinformation content that TikTok sees on the Platform. As such, TikTok continuously works to refine its policies, improve its detection and enforcement mechanisms and keep pace with evolving trends. As noted above, TikTok has implemented a range of new mitigation measures since Year 2 to improve its policy approach and connect users with reliable information. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.

## 18. FUNDAMENTAL RIGHTS: RISKS TO FUNDAMENTAL RIGHTS

### 1. Description of the Risk and how it may manifest on the Platform

TikTok understands risks to the exercise of “**Fundamental Rights**” on its Platform to comprise in particular the rights set out below, as enshrined in the Charter.

In its consideration of any actual or foreseeable negative effects on the exercise of fundamental rights protected under the Charter, TikTok has determined the following fundamental rights as most relevant to its Platform: (1) the right to human dignity; (2) the right to non-discrimination; (3) the right to freedom of expression; (4) the right to private and family life; (5) the right to the protection of personal data; and (6) the right to a high level of consumer protection. Various other fundamental rights are considered in this Report in the context of other risks.

As noted in section 4, TikTok’s mission is to inspire creativity and bring joy. At the heart of this mission, the Platform therefore seeks to encourage and foster opportunities for exercise of various Fundamental Rights, most notably the rights to freedom of expression and information. However, fundamental rights are not absolute and there is a necessity for careful balancing of Fundamental Rights if and when they come into conflict. For example, in exercising their right to freedom of expression, a user may choose to post on the Platform discriminatory content that undermines the human dignity and right to non-discrimination of the impacted individual(s) or group. In this case, it would be appropriate for TikTok to strike a balance by restricting that user’s freedom of expression, in line with its Community Guidelines, in order to limit the potential harm to others.

#### Summary of risks

Risks to Fundamental Rights may include:

- **Content risks:** The risk that content uploaded by TikTok’s users may undermine Fundamental Rights. For example, content that is discriminatory or dehumanising, that seeks to scam or defraud users or to otherwise undermine Fundamental Rights.
- **Conduct risks:** The risk of users misunderstanding or ignoring TikTok’s Terms of Service, Community Guidelines or Privacy Policy, and carrying out behaviour that is prohibited on the Platform.
- **Moderation risks:** The risk that TikTok’s content moderation systems and human moderators may: (i) over-moderate by taking action against speech that does not violate TikTok’s Community Guidelines; or (ii) under-moderate, by failing to recognise evolving discriminatory language or other violating content. In addition, there is a risk that users may be negatively impacted by bias, including through human and automated processes.



## 2. Inherent Risk

TikTok has assessed whether there have been changes since Year 2 in its inherent risk profile for risks related to Fundamental Rights and considered both severity and probability in reaching that view. TikTok has concluded that:

Overall severity	Medium-High
Overall probability	Likely
<b>Inherent risk</b>	<b>Medium-High</b>

This Inherent Risk score is consistent with TikTok's score in Year 2.

## 3. Mitigation Measures

TikTok undertakes the following mitigation measures related to risks to Fundamental Rights:

**Article 35(1)(a):** TikTok implements a range of measures to prevent and mitigate risks to the exercise of Fundamental Rights on or through the Platform, including: (i) measures to address the risk that comments on the Platform could undermine Fundamental Rights, including a filter for spam and offensive comments; (ii) measures to address the risk of a livestream being capable of containing content that undermines Fundamental Rights; (iii) blocking certain search results associated with harmful content that could undermine Fundamental Rights; (iv) measures to address the risk that hyperlinks shared on the Platform can be used to undermine Fundamental Rights; and (v) controls which mitigate the risk that TikTok Shop can be used to market or sell content or sell content in a manner that undermines Fundamental Rights, in particular consumer protection rights.

**Article 35(1)(b):** TikTok's Community Guidelines include TikTok's Community Principles, which set out TikTok's commitment to Fundamental Rights. TikTok also prohibits content undermining the Fundamental Rights on the Platform through a combination of TikTok's Terms of Service and Community Guidelines. For example, TikTok's Terms of Service prohibit content on the Platform which infringes anyone else's rights or contains or promotes violence or discrimination based on a range of protected characteristics. Further, TikTok's Privacy Policy outlines how the Platform collects, uses and shares personal information. It covers legal bases for processing, users' rights and choices, data security and retention, global data transfers, and how policy updates are communicated. In addition, TikTok's policies, including its Branded Content Policy, require users to disclose when they promote a brand, product or service in exchange for payment or other incentive. TikTok takes a range of measures to generate awareness and inform users about its terms and policies.

**Article 35(1)(c):** Please see section 4 ("About TikTok") for a description of TikTok's content moderation processes. TikTok's automated and manual review measures are designed to prevent the spread of harmful or illegal content that could infringe on Fundamental Rights, such as the right to non-discrimination and the right to human dignity, while also supporting

the right to freedom of expression by creating a safe and inclusive environment where users can express themselves freely without fear of suppression or harm.

**Article 35(1)(d):** Please see section 6 for a description of TikTok’s controls related to recommender systems.

**Article 35(1)(e):** Please see section 4 (“About TikTok”) for a description of TikTok’s advertising systems and policies. Ads on TikTok must not target individuals or promote discrimination based on protected characteristics, mislead consumers or otherwise violate TikTok’s Community Guidelines or Advertising Policies.

**Article 35(1)(f):** TikTok has specialist teams who deal with Fundamental Rights risks and who play a role in policy development, risk prevention and containment and reinforcing TikTok’s risk detection measures.

**Article 35(1)(g):** TikTok is committed to engaging with EU Trusted Flaggers. TikTok continues to take action on reports received from its Community Partner Channel, trusted subject matter experts and Trusted Flaggers.

**Article 35(1)(h):** TikTok engages with external experts, NGOs, and civil society organisations to identify and mitigate risks to Fundamental Rights. TikTok also participates in the EU Code of Conduct on Countering Illegal Hate Speech Online and collaborates with other platforms and the European Commission to improve industry-wide standards. In January 2025, the EU Code of Conduct on Countering Illegal Hate Speech Online + was implemented as a Code of Conduct under the DSA.

**Article 35(1)(i):** TikTok’s awareness-raising measures include: (i) TikTok’s Safety Centre, which includes resources on user safety and privacy; (ii) TikTok’s Transparency Centre, which has dedicated content explaining TikTok’s approach to content moderation and human rights; (iii) TikTok’s Help Centre, containing “how to” explanations on TikTok’s content moderation practices and how to report violative content as well as a range of resources on security and privacy; (iv) TikTok’s Privacy Centre, which contains resources and information on data processing and privacy tools on TikTok; (v) TikTok’s Newsroom, where TikTok shares updates on topics and product developments related to Fundamental Rights; and (vi) TikTok’s other in-app educational and awareness measures, verified badges, state-controlled media labels, unverified content labels and AI-generated content labels.

**Article 35(1)(j):** By their nature, the measures designed to mitigate risks to the exercise of Fundamental Rights also contribute to protecting the rights of the child. Additional Youth Safety measures are set out in the “Youth Safety and Online Engagement” section of this Report.

**Article 35(1)(k):** Please see section 6 for a description of TikTok’s controls related to AIGC.

TikTok has further enhanced its mitigation measures related to Fundamental Rights risks since its Year 2 Report, including:

- TikTok has expanded its Community Partner Channel by onboarding additional human rights organisations.
- TikTok has revised its freedom of expression assessments, which it uses to assess and improve updates to its Community Guidelines.
- TikTok has launched a refactored policy framework to improve clarity and consistency and reduce the risk of over or under-enforcement that could impact Fundamental Rights.
- TikTok has rolled out an implicit bias training programme for its content moderation team.
- TikTok has also introduced new measures to assess the fairness and safety of AI models and features, to mitigate risks to Fundamental Rights.
- TikTok has formed a Human Rights Coordination Group, a cross-functional group including individuals with a range of expertise, responsible for lending coherence to TikTok's human rights work by putting forward and ensuring the implementation of company-wide initiatives on human rights and anticipating human rights risks to the company and bringing them to the attention of leadership.

#### Relevant stakeholder engagement

TikTok's stakeholder engagement since Year 2 that is relevant to this risk includes:

- Engagement with human rights organisations in relation to the experience of human rights defenders on the Platform.
- Engagement with TikTok's Safety Advisory Councils, which are an important source of expert advice and outside perspectives across its safety work.
- Participation in the UN Forum on Business and Human Rights in Geneva. The forum provided a platform for dialogue with global stakeholders on corporate accountability, digital rights, and the role of tech companies in upholding human rights.
- Participation in RightsCon, a major global summit on digital human rights. TikTok's teams attended 60 sessions covering key topics such as content moderation, misinformation and disinformation, digital rights, and civil society collaboration. The event provided valuable insights into global digital rights trends, regulatory challenges, and strategies to enhance TikTok's future engagement and positioning.
- Participation in the United Nations "AI for Good Global Summit." TikTok met with established partners to discuss the C2PA initiative, connected with potential new partners and engaged with several AI creators. As an official industry partner for the summit, TikTok also hosted a panel discussion.

#### **4. Residual Risk**

Following an assessment of the effectiveness, reasonableness, and proportionality of TikTok's mitigations relevant to the systemic risk to Fundamental Rights (detailed in 3. above), TikTok has assessed residual risk to be **Medium** in Year 3. This Residual Risk score is consistent with TikTok's score in Year 2.

This Risk Assessment was informed by data on TikTok's enforcement of its Community Guidelines in the EU from Q3 2024 through Q1 2025. Systemic risks to Fundamental Rights

are observed across each of the Risk Modules in this Risk Assessment. The data assessed in connection with those Risk Modules was consulted in assessing the effectiveness of TikTok's mitigation measures related to risks to Fundamental Rights. In addition, TikTok reviewed data related to a range of policy categories which implicate Fundamental Rights.

From Q3 2024 through Q1 2025: (i) TikTok removed 587,855 videos under its "Privacy and Security" policy; 576,321 of those videos were detected and removed proactively (i.e. before any user report), 435,980 were removed before being viewed by any user, and 538,717 were removed within 24 hours from publication; (ii) TikTok removed 11,792,444 videos under its "Regulated Goods and Commercial Activities" policy; 11,523,078 of those videos were detected and removed proactively (i.e. before any user report), 9,424,572 were removed before being viewed by any user, and 10,998,736 were removed within 24 hours from publication; (iii) TikTok removed 5,650,697 videos under its "Safety and Civility" policy; 5,166,642 of those videos were detected and removed proactively (i.e. before any user report), 3,816,439 were removed before being viewed by any user, and 4,770,771 were removed within 24 hours from publication; and (iv) TikTok removed 7,200,632 videos under its "Mental and Behavioural Health" policy; 7,115,839 of those videos were detected and removed proactively (i.e. before any user report), 6,325,882 were removed before being viewed by any user, and 6,609,470 were removed within 24 hours from publication.

The rate of proactive removals is a positive indicator of the efficacy of TikTok's content moderation systems and processes. TikTok also receives reports from users, Trusted Flaggers and partners (through its Community Partner Channel). When TikTok receives such reports, it diligently acts to investigate and remove violative content. From Q3 2024 through Q1 2025, following a user report, TikTok removed 11,543 videos under its "Privacy and Security" policy, 269,366 videos under its "Regulated Goods and Commercial Activities" policy, 484,055 videos under its "Safety and Civility" policy and 84,793 videos under its "Mental and Behavioural Health" policy.

TikTok's policy approach works in combination with its other mitigation measures to mitigate the risks to elections and civic integrity on the Platform.

As noted in TikTok's Fourth DSA Transparency Report, which covers the period July to December 2024, TikTok received 9,110 reports from users of allegedly illegal content relating to consumer-related offences, 30,007 related to illegal privacy-related violations and 16,191 related to harassment or threats. During the same period, it received 26 reports from Trusted Flaggers related to illegal privacy-related violations and 5 related to harassment or threats. In addition, during the same period, TikTok received 701 requests from government authorities in the European Union to remove content, 2 of which related to illegal privacy-related violations and 44 of which related to harassment or threats.

As protecting Fundamental Rights on the Platform requires a careful balancing exercise of the rights at play, TikTok continuously works to improve its policy approach and remain alive to evolving trends. As noted above, TikTok remains engaged with external experts and has onboarded additional human rights organisations to its Community Partner Channel to enable it to keep up with industry understanding and consensus related to Fundamental Rights. Since Year 2, TikTok has also improved its policy approach and training measures and has implemented new fairness measures when testing AI models to minimise negative

impacts to Fundamental Rights on the Platform. This Risk Assessment has taken account of these factors in determining the effectiveness, reasonableness and proportionality of TikTok's mitigation measures.